Water Resources Research

RESEARCH ARTICLE
10.1029/2023WR036926

Satellite Observations Reveal Widespread Color Variations in Global Lakes Since the 1980s

Xiaoyi Shen[1,2], Chang-Qing Ke[2], Zheng Duan[3], Yu Cai[2], Haili Li[2], and Yao Xiao[2]

[1]School of Earth Sciences and Engineering, Hohai University, Nanjing, China, [2]Jiangsu Provincial Key Laboratory of Geographic Information Science and Technology, Key Laboratory for Land Satellite Remote Sensing Applications of Ministry of Natural Resources, School of Geography and Ocean Science, Nanjing University, Nanjing, China, [3]Department of Physical Geography and Ecosystem Science, Lund University, Lund, Sweden

Key Points:

- Fifty-Eight percent of global lakes shifted significantly toward shorter visible wavelengths from 1984 to 2021 based on ~32 million satellite observations
- Only 14% of the lakes across the globe have steady color patterns, indicating that most lakes are in an unstable state
- Landscape, population, and lake properties are associated with lake color changes, and basin vegetation condition is the most dominant

Correspondence to:
C.-Q. Ke,
kecq@nju.edu.cn

Citation:
Shen, X., Ke, C.-Q., Duan, Z., Cai, Y., Li, H., & Xiao, Y. (2025). Satellite observations reveal widespread color variations in global lakes since the 1980s. *Water Resources Research*, *61*, e2023WR036926. https://doi.org/10.1029/2023WR036926

Received 14 DEC 2023
Accepted 11 DEC 2024

Abstract The color of lakes is an essential indicator of the local ecological state, and the corresponding changes can reflect the physical and biochemical processes of lakes. However, worldwide changes in lake color and their drivers remain largely unknown. Here, we analyze the long-term color distributions and changes of 67,579 lakes worldwide from 1984 to 2021 by utilizing 32 million consistent satellite observations. Blue lakes (<495 nm) were primarily located in high-latitude and high-elevation areas. Green lakes (495–560 nm) were more prevalent in densely populated middle-latitude regions, while most red and yellow colors (≥560 nm) were located in the Southern Hemisphere. Our findings reveal distinct temporal patterns of lake color changes, with the majority of global lakes shifted toward shorter wavelengths. This phenomenon is more common in Warm Temperate and Boreal zones. Lake color changes are closely linked to basin vegetation conditions, population, water volume change, and lake area. Our study provides essential references for monitoring the ecological status of global lakes, further supporting the sustainable development of water resources in the future.

Plain Language Summary Lakes integrate multiple basin-scale climatic and anthropic processes and lake ecological state often characterizes as "sentinel" of climate change. Lake color is thought to be the closest variable to lake ecosystem properties, the variation of lake colors therefore reflects both short- and long-term climate fluctuations. However, worldwide changes in lake color and their drivers still remain largely unknown. Here we detected the color variation in global lake colors by building a lake color variability data set for 1984–2021. Most of these trends are attributed to basin vegetation conditions, variation trends of lake water volume, basin population and lake area. Considering the importance of these lakes for ecosystem services and water supply, the consequences of lake color changes are both locally and globally important.

1. Introduction

The ecological states of lakes, such as cyanobacterial blooms (Huisman et al., 2018), productivity (Kuhn et al., 2020), carbon storage (Bogard & del Giorgio, 2016; Buffam et al., 2011; Denfeld et al., 2018), and water quality (Topp, Pavelsky, Stanley, et al., 2021), are influenced by terrestrial processes and in turn affect human water security (Cao et al., 2023). Global lake ecology has undergone widespread changes linked to climate change and human activities (Adrian et al., 2009). Most of these changes can be closely linked to lake color; it is thought to be the closest variable to lake ecosystem properties (Yang et al., 2022) and is also the essential climate variable in the current global climate observing system (Kuhn & Butman, 2021). Therefore, dynamic monitoring of lake color is of great scientific significance for revealing the lake and regional ecological state.

Although in situ or field observations have provided insights into the color variations of some lakes (Kutser, 2012; Weyhenmeyer et al., 2016; Williamson et al., 2015), these measurements are often limited in spatial coverage and observation frequency, making large-scale observations challenging. Satellite remote sensing provides a promising approach for global-scale and regular monitoring of lake color (Kuhn & Butman, 2021). Recent studies have used multiple satellites to investigate lake colors; however, these studies are either limited to single specific regions or use inconsistent methods to derive lake color changes and identify their driving factors. For example, lake greenness declines were found for Arctic-boreal lakes, and the change is assumed to be related to the air temperature and precipitation (Kuhn & Butman, 2021); 68% of lakes in China shifted toward blue, and the trends and driving factors are spatially heterogeneous (Cao et al., 2023). In a recent advancement, Yang et al. (2022) investigated the colors of 85,360 lakes globally using satellite data from Landsat 8 and analyzed the relationship between lake color and climate conditions and lake morphology. Despite the significance of this endeavor, their

study was constrained by a relatively short observation period (2013–2020) and limited discussion on color changes. Surprisingly, global lakes have garnered comparatively less attention despite the recent emphasis on preserving the lake environment and safeguarding water quality in the context of sustainable development, and there is still a lack of understanding of global lake color trends, shifts and their drivers.

To address this knowledge gap, we estimated the annual colors for lakes and reservoirs worldwide from 1984 to 2021 using consistent satellite observations from Landsat 5, 7, and 8. We derived lake color trends for over 67,000 lakes larger than 1 km^2, representing 39% of the current total lake area across the globe. Small lakes were not included as the surface reflectance of them are more easily influenced by the surrounding land. We analyze lake color trends across various conditions of lake morphology and ecological zoning. Finally, a machine learning method was used to identify drivers of global lake color changes using data sets of climate, landscape, and human activity. Our study provides crucial insights into global lake ecology and has important implications for lake environment monitor.

2. Data and Methods

2.1. Data

2.1.1. Landsat Data

To obtain the longest and consistent satellite observations of lake colors, we used the Landsat Collection 1 Tier 1 data set (30 m spatial resolution) from 1984 to 2021 on the Google Earth Engine (GEE) platform. This data set includes data from Landsat 5, 7, and 8 and utilizes the highest quality images corrected for radiometry and georegistration (Yang et al., 2022). Uniform quality over time and across instruments has been maintained for this data set, ensuring its reliability for long-term analysis (USGS, 2019). This data set has been atmospherically corrected and can be directly used for inland water bodies, such as chlorophyll-a (Cao et al., 2020), suspended sediments (Dekker et al., 2001), primary productivity (Kuhn et al., 2020), and colored dissolved organic matter (CDOM) (Olmanson et al., 2020).

2.1.2. Auxiliary Data

We derived the colors of lakes from the Hydrolake data set (Messager et al., 2016), which contains records of over 1.4 million lakes across the globe. However, as we mentioned above, lakes with areas smaller than 1 km^2 were not considered. Consequently, approximately 180 thousand lakes were finally detected, and their locations and attributes (area, depth, and elevation) were recorded. Additionally, lake Basin information was derived from Sikder et al. (2023), climate zone data were sourced from Beck et al. (2018).

Climate, landscape and anthropogenic factors were used to detect the drivers of global lake color changes. Specifically, we used monthly air temperature, total precipitation, and wind speed data from the ERA5 data set, spanning from 1984 to 2021, at a spatial resolution of 10 km. The MODIS-based 16-day normalized difference vegetation index (NDVI) data set at 1 km from 2000 to 2021 was used and converted to annual-averaged NDVI at the lake basin scale to represent the local landscape conditions. Similarly, the annual world gridded population data at a spatial resolution of 1 km, which is available for the years from 2000 to 2021, from Oak Ridge National Laboratory (2021), were used to indicate anthropogenic factors, and basin total populations were used. A data set of annual lake water volume from 1992 to 2020 in Yao et al. (2023) was also used. All these data sets were converted to the annual scale.

2.2. Methods

2.2.1. Lake Color Derivation

The main processing steps are briefly summarized in Figure 1. We used the visible dominant wavelength (λ_d) in the lake centroid to represent the surface color of individual lakes (Cao et al., 2023). In particular, those lakes whose centroids are located in non-water areas or surrounding lands have been manually amended. The lake λ_d was estimated by converting the satellite reflectance in blue, green, and red bands to color wavelengths in the chromaticity color space, which can be perceived by humans (Wang et al., 2015). This indicator can reveal subtle changes in color that may not be captured by satellite reflectance (Topp, Pavelsky, Dugan, et al., 2021). The calculation process is summarized as follows. First, to reduce the possible influences from the lake bottom on lake

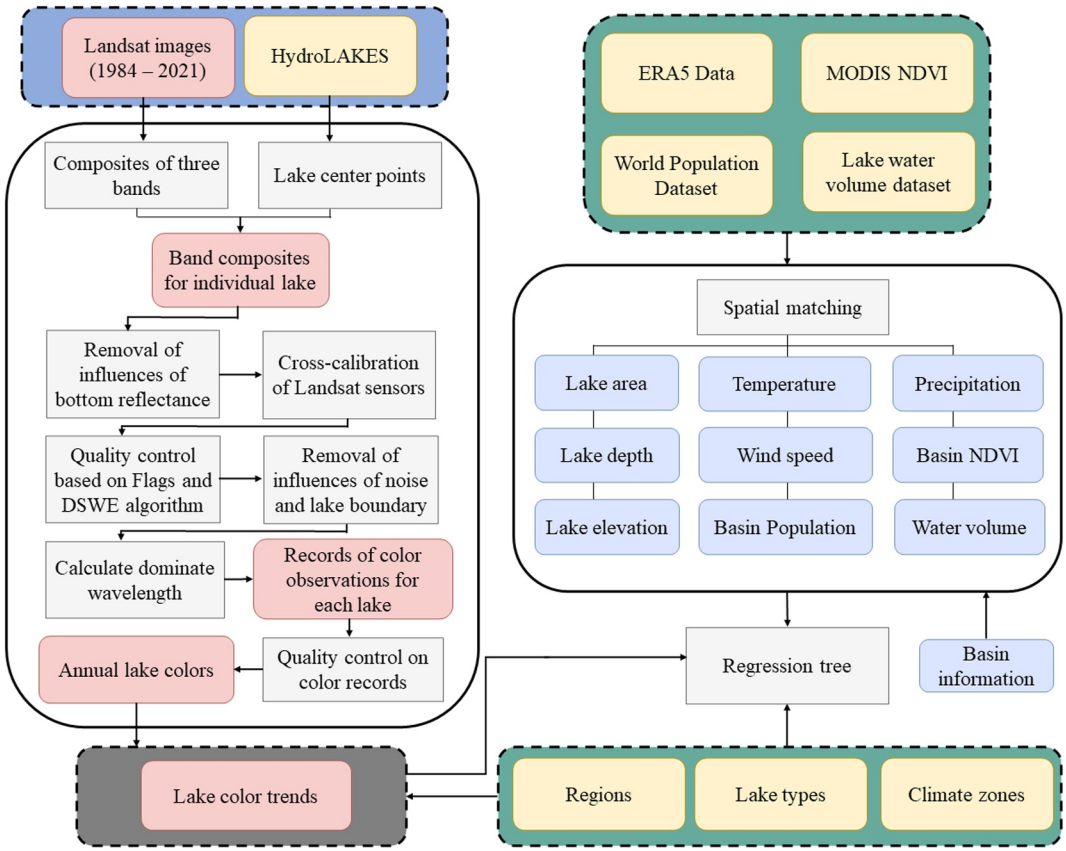

Figure 1. The main process of data processing, including data preprocessing and spatiotemporal matching of multisource data, calculation of lake color, generation of long-term lake color data set and analysis method of factors influencing lake color.

surface reflectance, only lakes with a mean depth larger than 2.2 m were kept, as in this situation the lakes were assumed to not likely to be affected by lake bottom reflectance (Yang et al., 2022). Second, to ensure the consistency of surface reflectance across Landsat sensors, the surface reflectance from Landsat 5 and 8 were both cross-calibrated to match with those from Landsat 7. In detail, the surface reflectance observed on the same day over the same lake location from Landsat 5–7 and 7–8 were collected, and a second-order polynomial regression was used to the 1%–99% percentiles of surface reflectance records for individual bands (Table 1). These empirical equations were then applied to the surface reflectance from Landsat 5 and 8 for consistent observations. More discussion about the performance of cross-calibration can be seen in Section 4.1. Third, to ensure the reliability of satellite observations, we identified and eliminated observations with clouds, shadows, and snow/ice based on the quality flag in Landsat data (Zhu et al., 2015). This was performed for the lake centroid and all pixels within a radius of 4 pixels around it, which also tends to reduce the influence of boundary land and vegetation. Fourth, we used the dynamic surface water extent algorithm (Jones, 2019) to assess the reliability of water pixels by detecting the presence of aquatic vegetation. Only observations with at least nine water pixels with high confidence were retained (Topp, Pavelsky, Dugan, et al., 2021). Fifth, to further reduce potential noise and lake boundary effects, we derived the median surface reflectance value of all pixels, including surface reflectance from the red, green, and blue bands. Last, following Wang et al. (2015), we converted the λ_d from the satellite surface reflectance observations.

On average, lakes have valid color observations every month, with relatively more observations in June, July, and August (the average number of monthly observations is 1.8, compared to 1.45 in other months, Figure 2b). To obtain reliable annual lake colors, only annual lake color records with more than six valid color observations were retained. More than six observations per year can guarantee one observation every 2 months on average, ensuring nearly uniform color sampling. In fact, the average annual number of observations is 24, and the number of

Table 1
The Coefficients of Equations to Correct the Surface Reflectance From Landsat 5, and 8 to Those From Landsat 7

Band	*a*	*b*	*c*	R^2
Correct Landsat 5 (to 7)				
Blue	0.9785	0.7697	0.0059580	0.8938
Green	0.2029	0.8911	0.0014080	0.9520
Red	0.07243	0.9670	−0.0003055	0.9535
Correct Landsat 8 (to 7)				
Blue	0.8750	0.6774	0.02072	0.7186
Green	0.2960	0.8007	0.01571	0.8817
Red	0.3832	0.8207	0.01300	0.9055

Note. The correction equation is $\rho_c = a\rho^2 + b\rho + c$, ρ_c is the corrected surface reflectance, ρ is the original surface reflectance, and a, b, and c are equation coefficients.

observations in mid-latitudes is higher (Figure 2a). The purpose of ensuring six observations is also to account for the limited number of satellite observations in the early period. To ensure the reliability of the trend calculation, only lakes with 20 or more valid annual color observations between 1984 and 2021 were used (Cao et al., 2023; Kuhn & Butman, 2021). More discussions about these thresholds can be seen in Section 4.2. Consequently, a total of 67,579 lakes were finally included. These lakes are widely distributed across the globe, ensuring their good spatial representation.

2.2.2. Statistical Analysis

Based on the annual λ_d record from 1984 to 2021, Sen's slope was calculated as the variation trend of lake color by using the Mann–Kendall (MK) trend analysis. MK test was then used to examine the significance of estimated trends, and a significant trend was identified with $p < 0.05$.

Prior investigations have demonstrated the close relationship between lake color changes and climatic conditions and lake characteristics (Cao et al., 2023; Hopkins et al., 2021; Hou et al., 2017; Kuhn & Butman, 2021; Paerl & Huisman, 2008). To comprehensively explore the drivers of global lake color changes as much as possible, we collected data sets of climate (temperature, precipitation, and wind speed), landscape (vegetation growth status, indicated by basin NDVI), human activity (total population), and lake properties (area, depth, elevation and water volume) for each lake. NDVI and total population were calculated at the basin scale; lake Basin, that is, referring to the entire drainage area encompassing each lake network, was derived from Sikder et al. (2023). These factors were chosen based on their global scale and proven links with lake color in previous studies. A machine learning method named regression tree analysis was applied to detect their potential

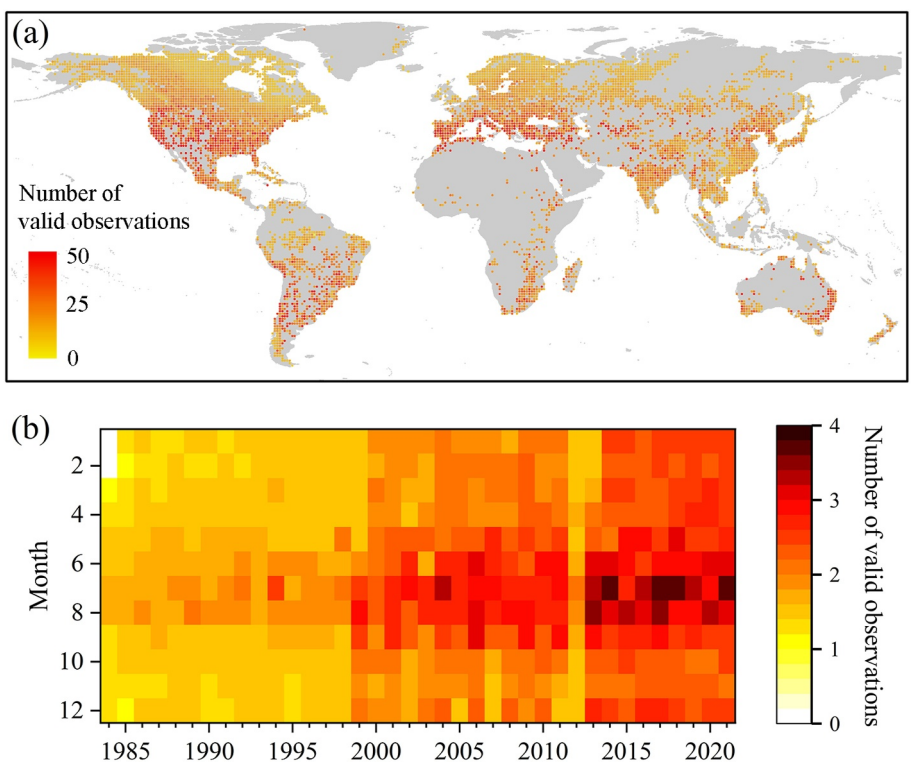

Figure 2. (a) Spatial distribution of average annual valid satellite observation counts for global lakes. The average count of valid observations for all lakes within each 1° × 1° grid was calculated and shown. (b) The temporal distribution of the average monthly valid satellite observation count.

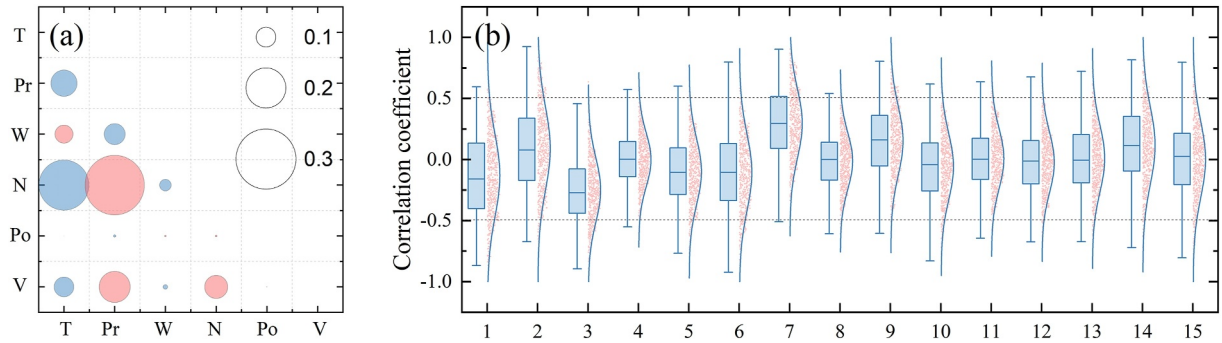

Figure 3. (a) Average correlation coefficients between temperature (T), precipitation (Pr), wind speed (W), normalized difference vegetation index (NDVI) (N), total population (Po), and lake water volume (V). Positive values are shown in red and negative values are shown in blue. (b) Distribution of correlation coefficients among different variables. Index 1: temperature versus precipitation, index 2: temperature versus wind speed, index 3: temperature versus NDVI, index 4: temperature versus total population, index 5: temperature versus lake water volume, index 6: precipitation versus wind speed, index 7: precipitation versus NDVI, index 8: precipitation versus total population, index 9: precipitation versus lake water volume, index 10: wind speed versus NDVI, index 11: wind speed versus total population, index 12: wind speed versus lake water volume, index 13: NDVI versus total population, index 14: NDVI versus lake water volume, and index 15: total population versus lake water volume.

relationship. Prior to the analysis, we calculated the correlation coefficients between various variables, including air temperature, precipitation, wind speed, basin NDVI, population, and lake water volume, to ensure that these variables were independent of each other. The results show that no apparent correlations can be found between them with most correlation coefficients between −0.5 and 0.5 (Figure 3), and all these factors can be independently used for analysis. In the regression tree analysis, we used all the variables mentioned above, including the trend variations and averaged values of climate and human activity factors and lake morphology (area, depth, elevation, and water volume), due to their potential connections with lake colors. Specifically, 15 factors were used as inputs for the regression tree analysis. These included variation trends and time-averaged values of various environmental and human activity variables (i.e., temperature, precipitation, wind speed, basin NDVI, water volume, and total population), as well as three lake properties: area, depth, and elevation.

To ensure the temporal consistency of the input data and analysis reliability, only the lakes with all 15 valid and spatiotemporally matched parameters were used, and the time series should be 16 years or longer. In total, 566 lakes were included, and the depth of the regression tree was optimized according to the minimized cross-validation error (Sharma et al., 2012).

3. Results

3.1. Distributions and Long-Term Change Trends of Global Lake Color Since the 1980s

Spatially, visual representations of global lake color (i.e., denominated as the period-median dominant wavelength λ_d) over the past 38 years reveal a predominantly blue and green color spectrum with distinct spatial clustering patterns (Figure 4). Blue lakes (<495 nm) were primarily found in high-latitude and high-elevation areas, which are usually characterized by ice cover in winter. Usually, there are cold temperatures in winter, which also affect the summer lake environment (Hampton et al., 2017) by influencing phytoplankton growth, reducing nutrients and suspending solid inputs (Yang et al., 2022). Green lakes (495–560 nm) were more prevalent in densely populated middle-latitude regions, likely due to more phytoplankton (Heathcote & Downing, 2012). Some lakes exhibited red and yellow colors (≥560 nm), most of which were located in the Southern Hemisphere.

Approximately 60% of lakes experienced significant color changes ($p < 0.05$), with a decrease in λ_d found in 96% of these cases. That is to say, we observed a significantly decreased λ_d in 58% of global lakes over the past four decades. The temporal changes of lake color present apparent spatial heterogeneity. Hotspots of significantly decreasing λ_d are found in high-latitude regions, including North America, Northern Europe, and Western Canada; whereas in the middle-latitude regions, some lakes experienced significantly increased λ_d (Figure 5a). Additionally, there is no apparent difference in the average trends of λ_d between different longitudes and latitudes (Figures 5b and 5c).

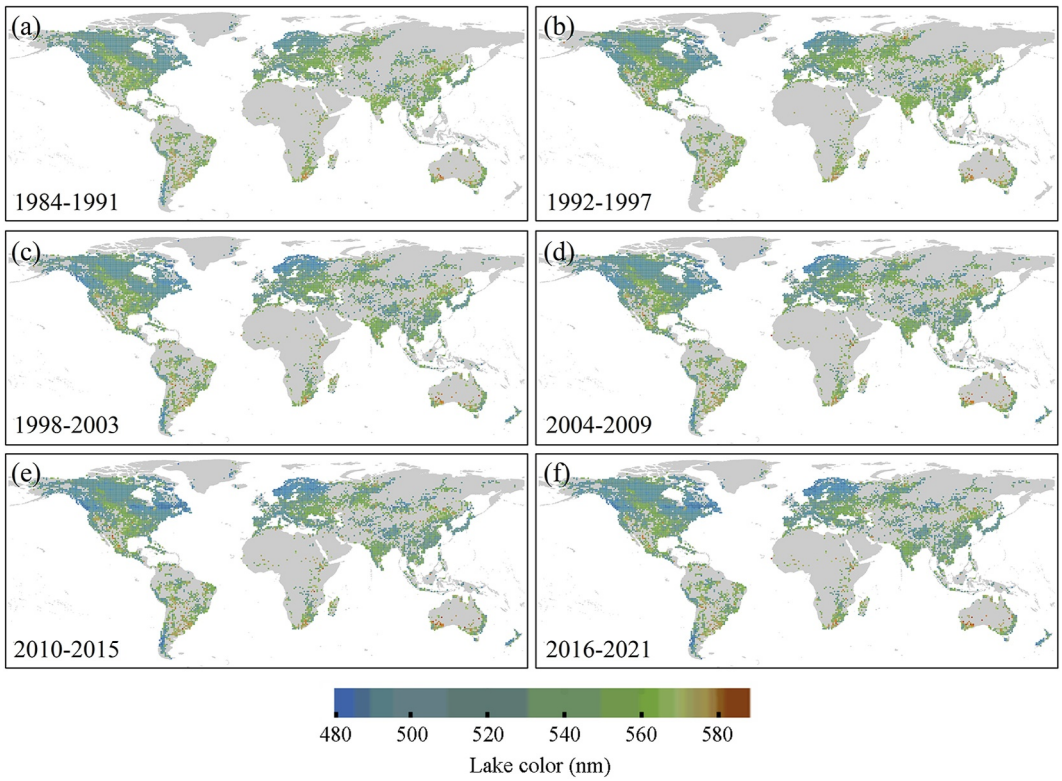

Figure 4. Spatial distribution of global lake color during individual time periods from 1984 to 2021. For better visualization, the lake colors are shown in grids of 1° × 1°. The median dominant wavelength λ_d of all lakes within each grid cell was calculated and transferred into the Forel-Ule Index and finally shown in the corresponding color. The calculation of Forel-Ule Index and the corresponding color can be referred to Wernand and Van der Woerd (2010).

Globally, the lake showed an overall decline in λ_d with an average value of −0.39 nm/yr from 1984 to 2021 (Figure 6a), most of them ranging from −2 to 1 nm/yr. Similarly, an overall decreasing trend of λ_d can also be found for lakes with different types and in individual continents and climate zones (Figures 6b–6d). Relatively lower λ_d trends can be found in North America and Europe (Figure 6e), South America and Oceania have a higher proportion of increasing λ_d trends. Similar pattern of λ_d trend can be found for natural lakes and reservoirs (Figure 6f). More lakes with decreasing λ_d trend can be found in Warm Temperate and Boreal zones, while lakes with increasing λ_d trend are usually presented in Equatorial and Arid regions (Figure 6g). Notably, lake morphology is also found to be related to changes in color. Smaller and shallower lakes are more likely to experience a decreasing λ_d trend (Figures 6h–6j).

3.2. Drivers of Global Lake Color Changes

A machine learning method (i.e., regression tree analysis) was applied to unravel the complex and nonlinear relationship between lake color changes and various factors, including climate, landscape, human activity, and lake morphology ($R^2 = 0.78$, Figure 7a). This method has been applied to the changes in lake surface temperature, stratification phenology, and dissolved organic carbon (O'Reilly et al., 2015; Toming et al., 2020; Woolway et al., 2021), which is suitable for revealing complex, nonlinear and multicollinearity relationships (O'Reilly et al., 2015). Here, 566 lakes were used for analysis (see Methods). These lakes are widely distributed across different regions and climate zones, encompassing various types and attribute characteristics. Compared with the 67,579 lakes analyzed in Section 3.1, the two data sets exhibit similar numerical distributions of λ_d trends (the averaged and standard deviation values of the 566 lakes are −0.25 ± 0.44 nm/yr, while the numbers of the 67,579 lakes are −0.39 ± 0.42 nm/yr). Thus, these selected lakes demonstrate good representativeness in terms of spatial distribution, attribute characteristics, and color trends.

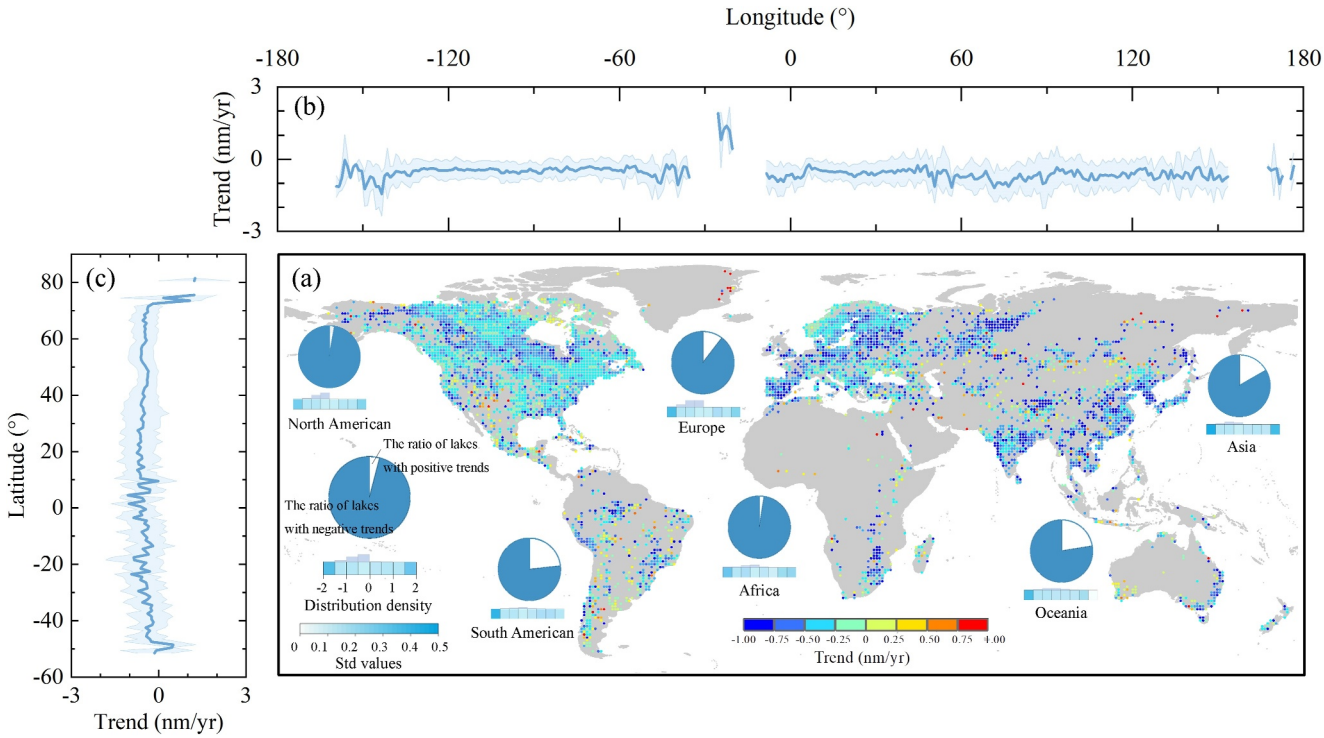

Figure 5. (a) Spatial distribution of the dominant wavelength λ_d trends for global lakes from 1984 to 2021. For better visualization, the lake colors are shown in grids of $1° \times 1°$, and the median color trend values for all lakes within each grid were used. Additionally, we display the numerical proportions of lakes showing different color trend patterns (i.e., positive and negative) at the continental scale. The distribution density of trend values, along with the standard deviation of trends within each value range, is also presented. Panels (b, c) show the average trend variation related to longitude and latitude, respectively. It should be noted that only lakes with significant λ_d trends were presented here.

The basin-wide NDVI was found to be the most dominant factor controlling lake color changes (Figure 7b). Usually, lakes with high clarity often have high vegetation coverage in the basin, impacting the concentration of detrital particles and promoting the lake to be less green (Pi et al., 2020). In regions with high NDVI (basin NDVI larger than or equal to 0.57), the lake λ_d tends to be much smaller with more water volume loss (<-0.23 Gt/yr, Figure 7a). This suggests that a large reduction in water volume would promote a further decrease in lake λ_d, while an increase in water volume may resist this change.

In low-NDVI areas (basin NDVI less than 0.57), population and lake area are more strongly related to lake color changes (Figure 7a). In general, in densely populated areas, the changes in lake λ_d are more apparent, indicating the direct influence of human activities on lake color changes. In less populated areas, changes in lake λ_d are often also affected by lake size. Large lakes have smaller changes in λ_d than smaller lakes, suggesting that lakes with more water are more resilient to color changes. This suggests that small lakes are more susceptible to color change and therefore need to be managed with more care in the future. Large lakes are more resilient to color change, but when they do change surface color, it suggests that they have been deeply affected.

We acknowledge that the above analysis only employs several global-scale factors to assess the drivers of lake color change, and the actual processes of change are often more complex, requiring consideration of the physical and chemical properties of the lakes. Here, we do not delve into the specific mechanisms driving color changes in individual lakes. Instead, we analyze the consistent influences of major climatic and anthropogenic factors on lakes worldwide.

4. Discussions

4.1. The Consistency of Reflectance Observations Across Landsat Sensors

Landsat data from 1984 to 2021 were used to analyze the variations of lake color around the world. The lake color observations of these 38 years were obtained by using three Landsat satellites. Due to the initial differences

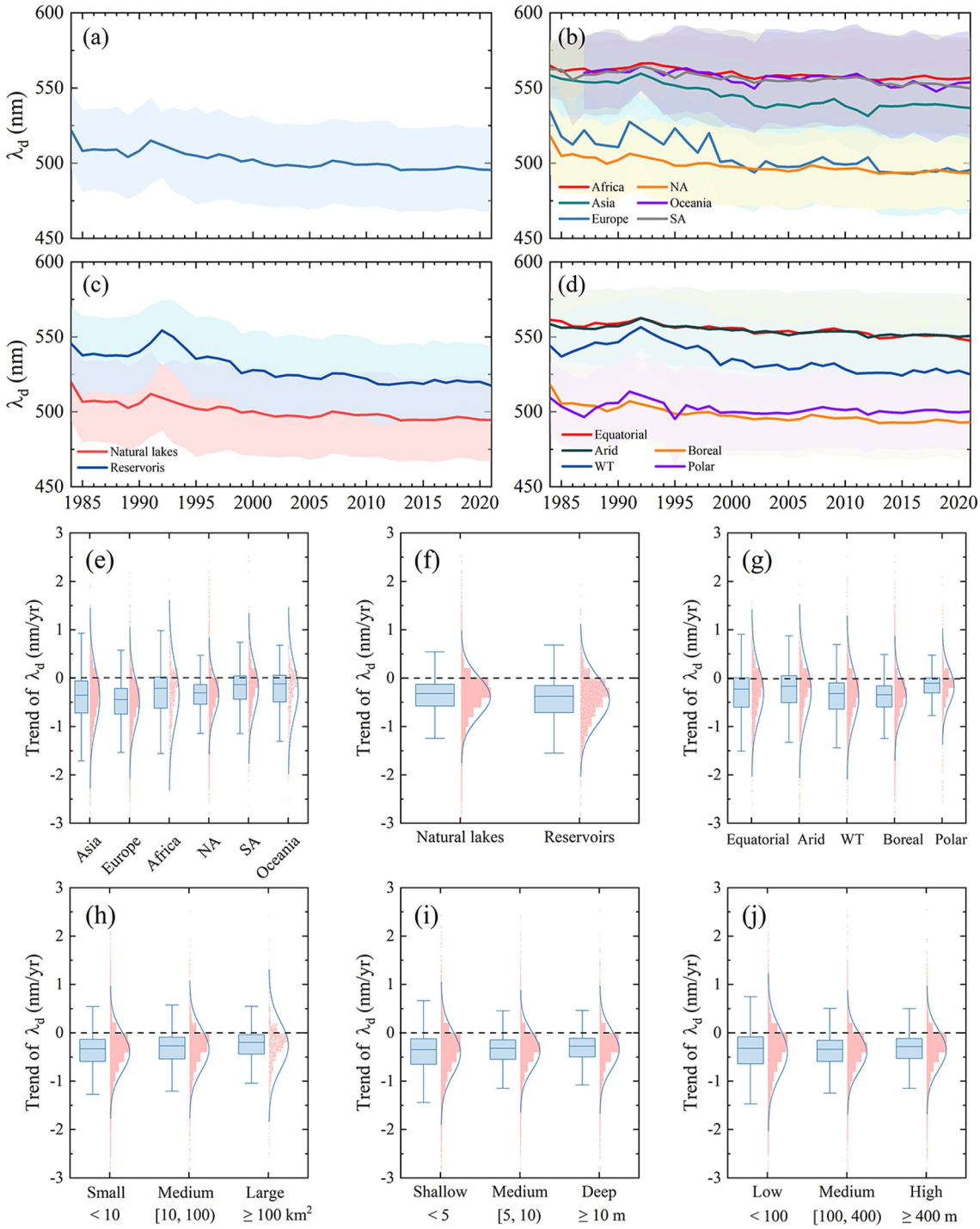

Figure 6. Time series of annual median λ_d for all lakes (a) and lakes across various spatial scales or lake attributes, including continents (b), lake types (where human-controlled natural water bodies were also classified as reservoirs, (c)), and climate zones ((d), Beck et al. (2018)). The bold lines represent the median values, and the standard deviations are shown in shading. (e–j) Comparison of λ_d trends for all lakes across the abovementioned spatial scales or lake attributes and lake surface area, depth, and elevation. NA: North America, SA: South America, and WT: Warm Temperature.

between them, obtaining lake color observations with high consistency is a prerequisite for the reliable analysis of color variation. Considering this, the biases across different Landsat sensors were corrected by using an empirical method as shown in Section 2.2.1, and Figure 8 shows the differences in surface reflectance across Landsat sensors. It can be found that, before cross-calibration, Landsat 5 has higher consistency with Landsat 7 compared

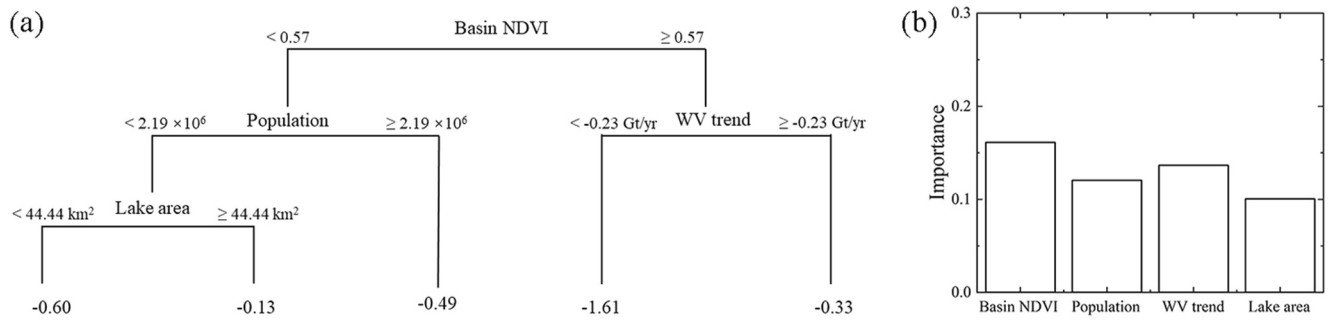

Figure 7. A regression tree to reveal the impact of climate, landscape, and lake morphology on (a) lake color change and (b) their importance. The numbers at each node represent the classification criteria, and the final number represents the average color trend of all lakes under that classification criterion.

with Landsat 8, but the differences in the blue and green bands seem to be more apparent. Similar phenomena can also be found between Landsat 7 and 8. Overall, the consistency of the data has been improved after cross-calibration, no matter which indicator is used (Table 2). The average surface reflectance of the red, green, and blue bands of Landsat 5 after cross-calibration is 0.06, 0.08, and 0.09, while the average reflectance of the red, green, and blue bands of Landsat 8 after correction is 0.05, 0.07, and 0.06. The average absolute difference after cross-calibration is less than 0.01 for all three bands of these two satellites, and thus the consistency of the surface reflectance is acceptable.

In particular, the surface reflectance of Landsat 8 is generally lower than that of Landsat 7 across all bands, particularly for reflectance values below 0.025 (Figure 8). This difference may be attributed to the higher radiometric resolution of Landsat 8, which enables the detection of smaller reflectance values. A similar observation was also reported by Maciel et al. (2023), which compared the reflectance of Landsat 7 and Landsat 8 using in situ measurements from over 1,000 inland lakes. To ensure consistency among observations from multiple Landsat satellites, correction equations were developed to align Landsat 5 and Landsat 8 reflectance with that of Landsat 7. The analysis above has validated the reliability of these equations. After the calibration of

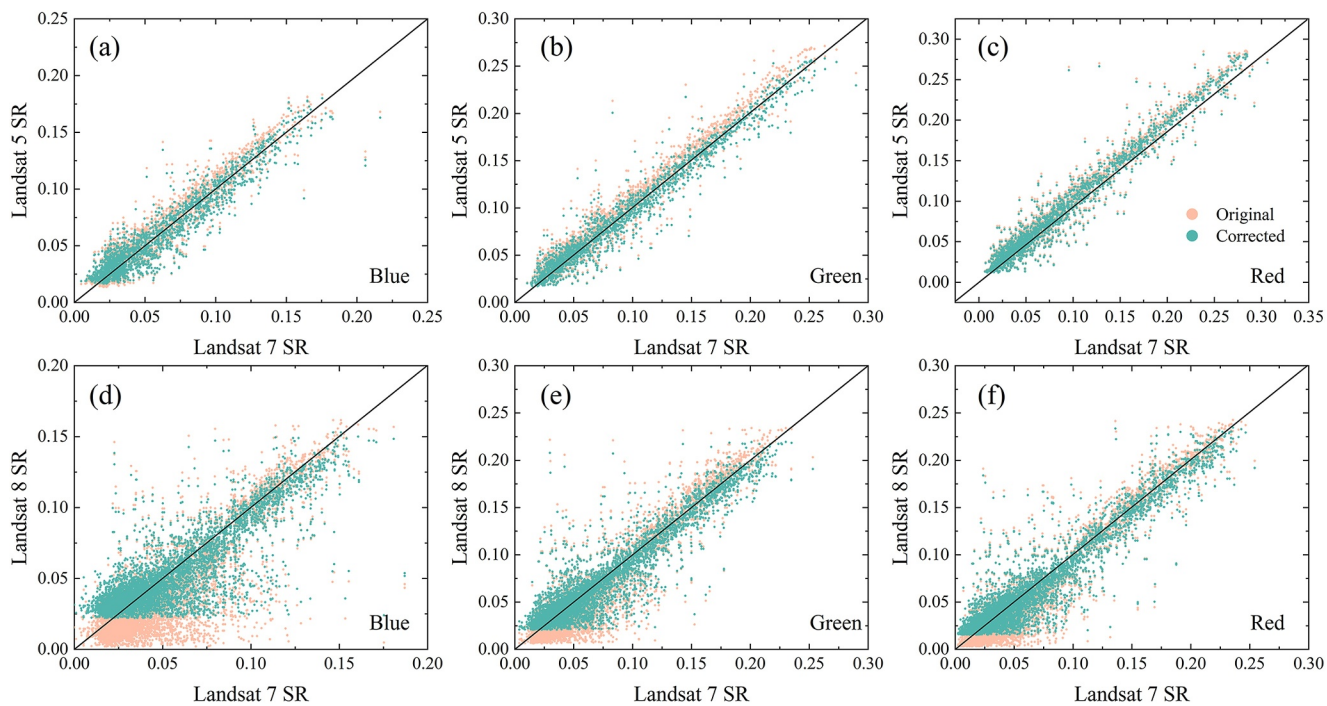

Figure 8. Scatter plot of the surface reflectance from (a–c) Landsat 7 and Landsat 5, (d, e) Landsat 7 and Landsat 8. Original and corrected surface reflectance are shown in orange and cyan, respectively. SR: surface reflectance.

Table 2
Comparsion of the Corrected Surface Reflectance From Landsat 5, 8 to Those From Landsat 7

Band	Mean difference (nm)		Average absolute difference (nm)		RMSE (nm)		*R*	
Correct Landsat 5 (to 7)								
Blue	−0.0031	-4.5137×10^{-7}	0.0096	0.0088	0.0128	0.0120	0.9446	0.9454
Green	−0.0063	-4.9302×10^{-6}	0.0107	0.0084	0.0147	0.0127	0.9756	0.9757
Red	−0.0023	-4.4672×10^{-7}	0.0094	0.0090	0.0145	0.0143	0.9765	0.9765
Correct Landsat 8 (to 7)								
Blue	0.0104	5.0567×10^{-6}	0.0148	0.0114	0.0210	0.0168	0.8470	0.8477
Green	0.0056	1.2053×10^{-6}	0.0127	0.0107	0.0190	0.0165	0.9389	0.9390
Red	0.0054	5.5741×10^{-6}	0.0119	0.0105	0.0182	0.0163	0.9513	0.9516

Note. RMSE: Root Mean Squared Error and *R*: correlation coefficient.

Landsat 8 data, across all three bands, there is a lack of small reflectance values. We attribute these to differences in the observational performance of the two satellites. Landsat 7 is less capable of detecting low reflectance values in water bodies compared to Landsat 8. By applying the correction equations, the smaller reflectance values captured by Landsat 8 are adjusted to a level consistent with Landsat 7, thereby reducing their apparent magnitude. This phenomenon has also been observed in other studies (e.g., Topp, Pavelsky, Dugan, et al., 2021; Topp, Pavelsky, Stanley, et al., 2021). Nonetheless, consistent surface reflectance is crucial for ensuring the reliability of long-term series analyses.

To further illustrate the necessity and effectiveness of cross-calibration, we randomly selected six lakes around the world and displayed the variation of annual λ_d before and after cross-calibration. The results show that before calibration, there are certain differences between the λ_d derived from different satellites, often showing consistent overestimation or underestimation (Figure 9). These differences are bound to affect the calculation of the λ_d trend, while the time series of corrected λ_d is more consistent and the obtained trend is thus more reliable.

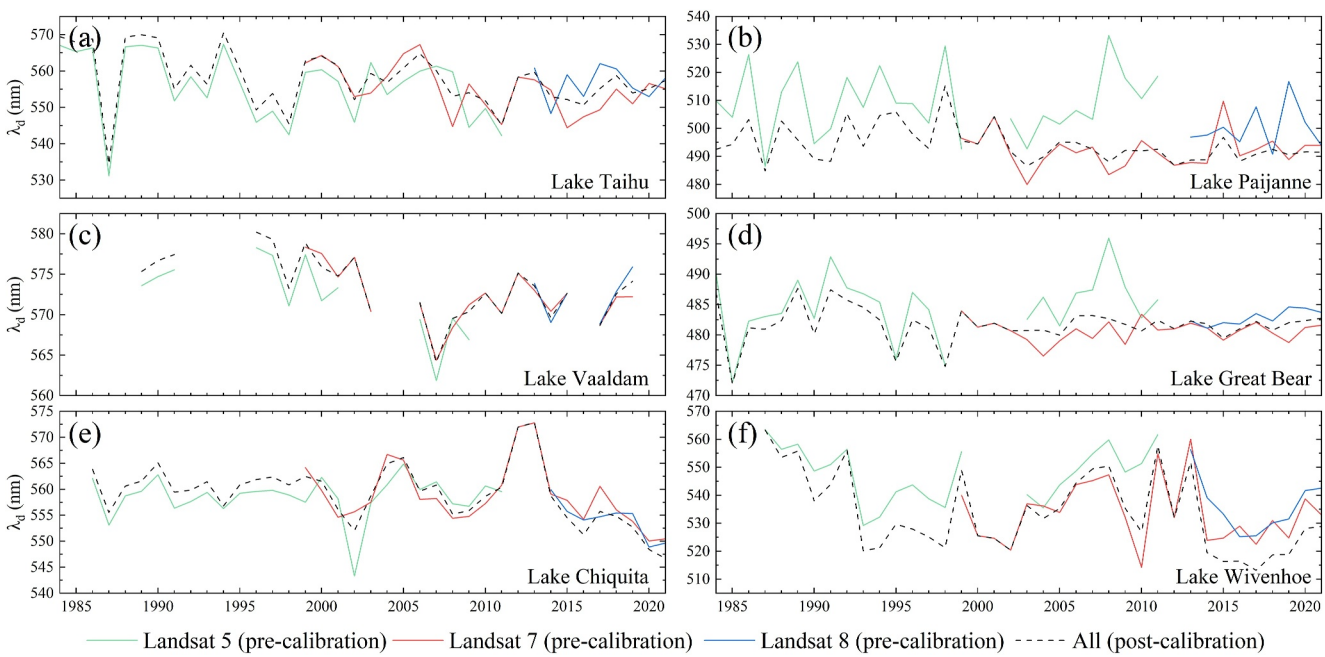

Figure 9. The variation of annual λ_d in (a) Lake Taihu (Asia), (b) Lake Paijanne (Europe), (c) Lake Vaaldam (Africa), (d) Lake Great Bear (North America), (e) Lake Chiquite (South America), and (f) Lake Wivenhoe (Oceania) from 1984 to 2021 derived from pre- and post-calibration of the times series of Landsat images. The data before calibration is shown in color (green for Landsat 5, red for Landsat 7, and blue for Landsat 8), while the data after calibration is shown in black.

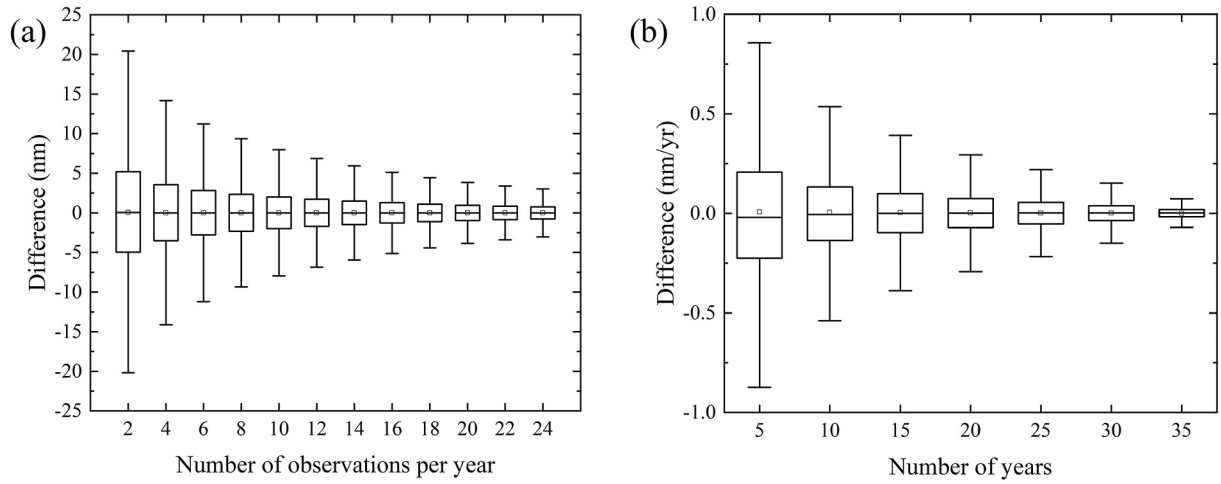

Figure 10. (a) The difference of annual λ_d derived by using all observations and different numbers of observations per year. (b) The difference of λ_d trends derived by using observations from all years and these from different numbers of years.

4.2. The Effect of the Missing Data on Trend Analysis of Lake Color

Missing data is another factor that affects the trend analysis of lake color. Missing data usually caused by not meeting quality control standards (e.g., heavy cloud cover) and then the relevant data are excluded. Due to these issues, the observation frequency of lake color within 1 year is uneven and data for some months are usually lacking, affecting the derivation of annual λ_d. In order to discuss the possible impact of this on the analysis of color trends, lakes with complete 12 months of λ_d observations in a certain year or multiple years were collected as samples, and finally obtained 6,635 lakes with a total of 30,559 years of observations. In individual years, for each lake, $2i$ observations were randomly selected ($i = 1, 2, \ldots, 12$), and the average λ_d obtained from each randomly selected sample was calculated and compared with the average λ_d obtained from all observations. The above random process was repeated 100 times. The results show that most differences range between −20 and 20 nm; with the number of samples increasing, the difference between the calculated λ_d value and the final value becomes smaller (Figure 10a). However, it is important to note that most lakes have missing data. It would further reduce the number of lakes that can be analyzed if much more observations within 1 year are needed. For example, if 15 observations per year are needed, only 8,894 lakes can be kept (Figure 11a). Ultimately, a compromise threshold was chosen, which required six observations per year, which would generally ensure an observation every 2 months. The average absolute difference is 5.18 nm in this case, which is about 1% compared to the 475–580 nm range of lake λ_d. Therefore, the impact on trend calculations can be ignored and also more lakes may be retained.

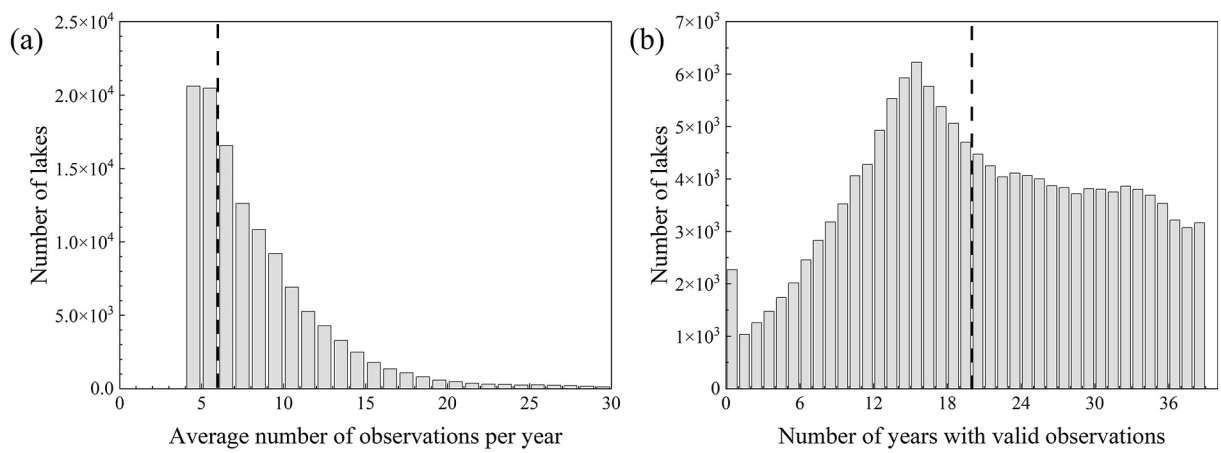

Figure 11. (a) The number of lakes with different average numbers of observations per year. (b) The number of lakes with different numbers of years with valid observations. The black lines in panels (a, b) indicate the thresholds used for quality control.

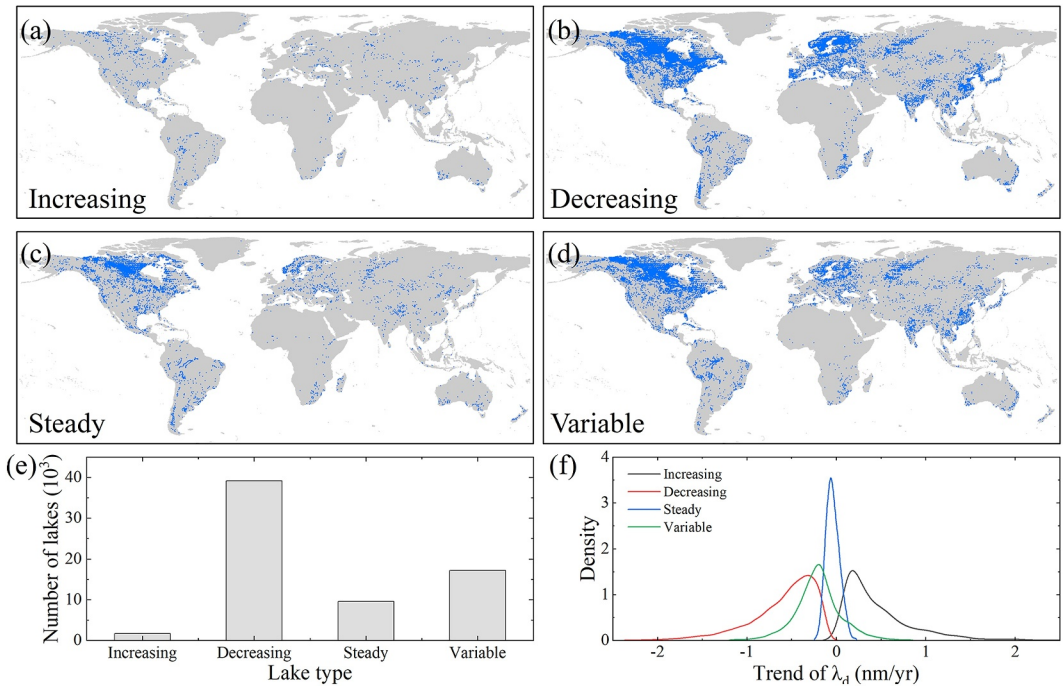

Figure 12. The spatial locations of lakes with an (a) increasing trend, (b) decreasing trend, (c) steady state, and (d) variable state. (e) The number of different types of lakes around the world. (f) The distribution of λ_d trend for different types of lakes around the world.

Similarly, we also conducted the trend analysis on the intra-annual scale. Lakes with 38 years of valid lake color observations were selected, resulting in 3,163 lakes in total. Then, we randomly selected $5i$ observations ($i = 1, 2, \ldots 7$), and compared the color trend obtained from each randomly selected sample with the trend obtained by using all observations. The random process was also repeated 100 times. Most differences range between -1 and 1 nm/yr; with the number of samples increases, the trend value becomes closer to the final value (Figure 10b). Lakes with at least 20 years of observations were used, in this case, more than 60,000 lakes could be used for analysis (Figure 11b). The average trend value deviation in this case was 0.1 nm/yr, which is about 25% of the average trend of all lakes of 0.44 nm/yr.

4.3. Additional Analysis of the Variation Trends of Lake Color

It can be seen from the above that there are some uncertainties in the estimated trend of lake λ_d. Here, instead of focusing on the numerical value, we divide the variation trends into four categories: significant increasing, significant decreasing, steady, and variable states. For lakes that λ_d have not changed significantly when the variation coefficient is less than or equal to 1%, the lake situation is deemed to be steady, otherwise, it is considered as a variable state. Compared to lakes with significant increasing trends, other types of lakes have more widespread distributions (Figures 12a–12d). We found that most lakes were in a state of significant decrease in λ_d ($n = 39{,}137$), followed by lakes in a state of color change ($n = 17{,}167$) and a stable state ($n = 9{,}613$, Figure 12e); a small number of lakes were in a state of significant increase ($n = 1{,}662$), indicating that most lakes are in an unstable state. Although there is no significant trend of lakes in a steady state, most of them also have decreasing trends (Figure 12f).

4.4. The Uncertainty of λ_d as a Representative of Lake Color

In this study, we use dominant wavelength λ_d as an indicator of lake color and analyze its long-term changes. We acknowledge that this indicator has certain uncertainties. First, λ_d is calculated by the reflectance of three visible bands from Landsat satellites. Compared with hyperspectral observations, their spectral resolution is relatively coarse (Woerd & Wernand, 2015), and it may not be quite sensitive to slight changes in the spectrum. Second,

chromaticity is known to be perceptually nonlinear (Gardner et al., 2021), two colors may have similar λ_d values but are perceived slightly differently. Third, the relationship between λ_d and other water quality parameters, such as Secchi depth (SD), colored dissolved organic matter (CDOM), and chlorophyll (CHL), is complex (Burket et al., 2023). For a given λ_d value, it may be caused by different combinations of SD, CDOM, and CHL. Therefore, λ_d cannot provide a quite clear inference of some water quality parameters. In addition, λ_d has boundedness, and its range is often distributed between 472 and 572 nm. For example, dominant wavelengths will saturate to an upper limit of around 572 nm regardless of the water quality constituent present, so long as they reach a certain high concentration (Bukata et al., 1997). Therefore, λ_d has a limited capacity to determine the quantified concentrations of the various optical water quality variables (Malthus et al., 2020). Due to the coarse spectral resolution and perceptually nonlinear chromaticity, more slight color changes may not be captured. Additionally, 98% of the lakes analyzed in the present study have λ_d between 480 and 570 nm, so the trend analysis may be less affected by the boundedness of λ_d. Although λ_d will be affected by the above factors, it has still been successfully and widely used in many studies (Cao et al., 2023; Gardner et al., 2021; Topp, Pavelsky, Dugan, et al., 2021; Topp, Pavelsky, Stanley, et al., 2021; Yang et al., 2022).

4.5. Future Implications

Here, lake color is chosen as the monitoring indicator to represent the lake ecological state because of its good representativeness, and it can be easily obtained through remote sensing technology. This study presents, for the first time, consistent spatiotemporal patterns of global lake color changes over the past four decades. Global lake color is currently undergoing significant shifts, and the divergent shifts in lake color can be attributed to a combination of climate, landscape dynamics, anthropogenic activities, and lake morphometry factors. The changes in lake color also indicate possible ecosystem shifts, such as decreased colored dissolved material or plankton. Given the ongoing global climate change and human impacts, future trends in lake color are expected to change significantly. Our findings provide crucial insights into shifts in lake color, serve as invaluable references for assessing changes in lake ecological statuses and environmental conditions, and help to improve the understanding of the complex interactions between climate, anthropogenic impacts, and lake colors. Furthermore, these findings can guide policy-makers in formulating more informed, practical policies related to lake and environmental conservation, as color changes are related to lake attributes and geochemical processes, including lake surface temperature (Wetzel et al., 2000), lake stratification (Read & Rose, 2013), carbon storage, productivity (Kuhn & Butman, 2021), and phytoplankton and food chains (Leech et al., 2018).

5. Conclusions

Using a comprehensive data set derived from almost 40 years of continuous Landsat observations, this study investigates the color changes in 67,579 lakes worldwide at various spatial and temporal scales. The findings reveal distinct global lake color changes, with a prevailing decrease in dominant wavelength. Lakes of different continents, climate zones, types, and attributes show varied color change trends. By integrating long-term lake color observations with climate, landscape, and human activity data, we identified that basin NDVI, population, water volume change, and lake area were assumed to affect the variation of lake colors.

Data Availability Statement

All data used in this study are publicly available. Landsat Collection 1 Tier 1 data set (including Landsat-5/7/8) from 1984 to 2021, MODIS NDVI data set (i.e., MOD13A2 V6.1 product) from 2000 to 2021 and ERA5-Land data set of monthly air temperature, total precipitation, and wind speed data from 1984 to 2021 were derived through Google Earth Engine (GEE) code editor at https://code.earthengine.google.com. Please enter the keywords "USGS Landsat 5/7/8 Surface Reflectance Tier 1," "ERA5-Land Monthly Aggregated-ECMWF Climate Reanalysis," and "MOD13A2.061 Terra Vegetation Indices 16-Day Global 1 km" to retrieve the corresponding data set; Lake boundary and basin information were obtained at Messager et al. (2016) and Sikder et al. (2023), respectively; World gridded population data from 2000 to 2021 was derived at Oak Ridge National Laboratory (2021); Annual lake water volume data set from 1992 to 2020 was derived at Yao et al. (2023). The code scripts required to produce the lake color data can be derived from Topp et al. (2020). The lake color data set used in this study is available from Shen (2024).

Acknowledgments
This work is supported by the National Natural Science Foundation of China (42206174) and the Natural Science Foundation of Jiangsu Province (BK20210193).

References

Adrian, R., O'Reilly, C. M., Zagarese, H., Baines, S. B., Hessen, D. O., Keller, W., et al. (2009). Lakes as sentinels of climate change. *Limnology and Oceanography*, *54*(6), 2283–2297. https://doi.org/10.4319/lo.2009.54.6_part_2.2283

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific Data*, *5*(1), 1–12. https://doi.org/10.1038/sdata.2018.214

Bogard, M. J., & del Giorgio, P. A. (2016). The role of metabolism in modulating CO_2 fluxes in boreal lakes. *Global Biogeochemical Cycles*, *30*(10), 1509–1525. https://doi.org/10.1002/2016GB005463

Buffam, I., Turner, M. G., Desai, A. R., Hanson, P. C., Rusak, J. A., Lottig, N. R., et al. (2011). Integrating aquatic and terrestrial components to construct a complete carbon budget for a north temperate lake district. *Global Change Biology*, *17*(2), 1193–1211. https://doi.org/10.1111/j.1365-2486.2010.02313.x

Bukata, R. P., Jerome, J. H., Kondratyev, K. Y., Pozdnyakov, D. V., & Kotykhov, A. A. (1997). Modelling the radiometric color of inland waters: Implications to a) remote sensing and b) limnological color scales. *Journal of Great Lakes Research*, *23*(3), 254–269. https://doi.org/10.1016/S0380-1330(97)70910-9

Burket, M. O., Olmanson, L. G., & Brezonik, P. L. (2023). Comparison of two water color algorithms: Implications for the remote sensing of water bodies with moderate to high CDOM or chlorophyll levels. *Sensors*, *23*(3), 1071. https://doi.org/10.3390/s23031071

Cao, Z., Ma, R., Duan, H., Pahlevan, N., Melack, J., Shen, M., & Xue, K. (2020). A machine learning approach to estimate chlorophyll-a from Landsat-8 measurements in inland lakes. *Remote Sensing of Environment*, *248*, 111974. https://doi.org/10.1016/j.rse.2020.111974

Cao, Z., Melack, J. M., Liu, M., Kutser, T., Duan, H., & Ma, R. (2023). Shifts, trends, and drivers of lake color across China since the 1980s. *Geophysical Research Letters*, *50*(8), e2023GL103225. https://doi.org/10.1029/2023GL103225

Dekker, A. G., Vos, R., & Peters, S. W. (2001). Comparison of remote sensing data, model results and in situ data for total suspended matter (TSM) in the southern Frisian lakes. *Science of the Total Environment*, *268*(1–3), 197–214. https://doi.org/10.1016/S0048-9697(00)00679-3

Denfeld, B. A., Klaus, M., Laudon, H., Sponseller, R. A., & Karlsson, J. (2018). Carbon dioxide and methane dynamics in a small boreal lake during winter and spring melt events. *Journal of Geophysical Research: Biogeosciences*, *123*(8), 2527–2540. https://doi.org/10.1029/2018JG004622

Gardner, J. R., Yang, X., Topp, S. N., Ross, M. R. V., Altenau, E. H., & Pavelsky, T. M. (2021). The color of rivers. *Geophysical Research Letters*, *48*(1), e2020GL088946. https://doi.org/10.1029/2020GL088946

Hampton, S. E., Galloway, A. W., Powers, S. M., Ozersky, T., Woo, K. H., Batt, R. D., et al. (2017). Ecology under lake ice. *Ecology Letters*, *20*(1), 98–111. https://doi.org/10.1111/ele.12699

Heathcote, A. J., & Downing, J. A. (2012). Impacts of eutrophication on carbon burial in freshwater lakes in an intensively agricultural landscape. *Ecosystems*, *15*(1), 60–70. https://doi.org/10.1007/s10021-011-9488-9

Hopkins, J. E., Palmer, M. R., Poulton, A. J., Hickman, A. E., & Sharples, J. (2021). Control of a phytoplankton bloom by wind-driven vertical mixing and light availability. *Limnology and Oceanography*, *66*(5), 1926–1949. https://doi.org/10.1002/lno.11734

Hou, X., Feng, L., Duan, H., Chen, X., Sun, D., & Shi, K. (2017). Fifteen-year monitoring of the turbidity dynamics in large lakes and reservoirs in the middle and lower basin of the Yangtze River, China. *Remote Sensing of Environment*, *190*, 107–121. https://doi.org/10.1016/j.rse.2016.12.006

Huisman, J., Codd, G. A., Paerl, H. W., Ibelings, B. W., Verspagen, J. M., & Visser, P. M. (2018). Cyanobacterial blooms. *Nature Reviews Microbiology*, *16*(8), 471–483. https://doi.org/10.1038/s41579-018-0040-1

Jones, J. W. (2019). Improved automated detection of subpixel-scale inundation—Revised dynamic surface water extent (DSWE) partial surface water tests. *Remote Sensing*, *11*(4), 374. https://doi.org/10.3390/rs11040374

Kuhn, C., Bogard, M., Johnston, S. E., John, A., Vermote, E., Spencer, R., et al. (2020). Satellite and airborne remote sensing of gross primary productivity in boreal Alaskan lakes. *Environmental Research Letters*, *15*(10), 105001. https://doi.org/10.1088/1748-9326/aba46f

Kuhn, C., & Butman, D. (2021). Declining greenness in Arctic-boreal lakes. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(15), e2021219118. https://doi.org/10.1073/pnas.2021219118

Kutser, T. (2012). The possibility of using the Landsat image archive for monitoring long time trends in coloured dissolved organic matter concentration in lake waters. *Remote Sensing of Environment*, *123*, 334–338. https://doi.org/10.1016/j.rse.2012.04.004

Leech, D. M., Pollard, A. I., Labou, S. G., & Hampton, S. E. (2018). Fewer blue lakes and more murky lakes across the continental US: Implications for planktonic food webs. *Limnology and Oceanography*, *63*(6), 2661–2680. https://doi.org/10.1002/lno.10967

Maciel, D. A., Pahlevan, N., Barbosa, C. C. F., de Novo, E. M. L. D. M., Paulino, R. S., Martins, V. S., et al. (2023). Validity of the Landsat surface reflectance archive for aquatic science: Implications for cloud-based analysis. *Limnology and Oceanography Letters*, *8*(6), 850–858. https://doi.org/10.1002/lol2.10344

Malthus, T. J., Ohmsen, R., & van der Woerd, H. J. (2020). An evaluation of citizen science smartphone apps for inland water quality assessment. *Remote Sensing*, *12*(10), 1578. https://doi.org/10.3390/rs12101578

Messager, M. L., Lehner, B., Grill, G., Nedeva, I., & Schmitt, O. (2016). Estimating the volume and age of water stored in global lakes using a geo-statistical approach. *Nature Communications*, *7*(1), 13603. https://doi.org/10.1038/ncomms13603

Oak Ridge National Laboratory. (2021). LandScan Global-global population distribution data (2000-2021) [Dataset]. Retrieved from https://landscan.ornl.gov/

Olmanson, L. G., Page, B. P., Finlay, J. C., Brezonik, P. L., Bauer, M. E., Griffin, C. G., & Hozalski, R. M. (2020). Regional measurements and spatial/temporal analysis of CDOM in 10,000+ optically variable Minnesota lakes using Landsat 8 imagery. *Science of the Total Environment*, *724*, 138141. https://doi.org/10.1016/j.scitotenv.2020.138141

O'Reilly, C. M., Sharma, S., Gray, D. K., Hampton, S. E., Read, J. S., Rowley, R. J., et al. (2015). Rapid and highly variable warming of lake surface waters around the globe. *Geophysical Research Letters*, *42*(24), 10773–10781. https://doi.org/10.1002/2015GL066235

Paerl, H. W., & Huisman, J. (2008). Blooms like it hot. *Science*, *320*(5872), 57–58. https://doi.org/10.1126/science.1155398

Pi, X., Feng, L., Li, W., Zhao, D., Kuang, X., & Li, J. (2020). Water clarity changes in 64 large alpine lakes on the Tibetan Plateau and the potential responses to lake expansion. *ISPRS Journal of Photogrammetry and Remote Sensing*, *170*, 192–204. https://doi.org/10.1016/j.isprsjprs.2020.10.014

Read, J. S., & Rose, K. C. (2013). Physical responses of small temperate lakes to variation in dissolved organic carbon concentrations. *Limnology and Oceanography*, *58*(3), 921–931. https://doi.org/10.4319/lo.2013.58.3.0921

Sharma, S., Legendre, P., Boisclair, D., & Gauthier, S. (2012). Effects of spatial scale and choice of statistical model (linear versus tree-based) on determining species–habitat relationships. *Canadian Journal of Fisheries and Aquatic Sciences*, *69*(12), 2095–2111. https://doi.org/10.1139/cjfas-2011-0505

Shen, X. (2024). Global lake surface color dataset (1984-2021) [Dataset]. Retrieved from https://data.tpdc.ac.cn/en/disallow/9d8f8716-3ffd-4e20-8bbf-fe07a372fb9c

Sikder, M. S., Wang, J., Allen, G. H., Sheng, Y., Yamazaki, D., Song, C., et al. (2023). Lake-TopoCat: A global lake drainage topology and catchment database. *Earth System Science Data*, *15*(8), 3483–3511. https://doi.org/10.5194/essd-15-3483-2023

Toming, K., Kotta, J., Uuemaa, E., Sobek, S., Kutser, T., & Tranvik, L. J. (2020). Predicting lake dissolved organic carbon at a global scale. *Scientific Reports*, *10*(1), 8471. https://doi.org/10.1038/s41598-020-65010-3

Topp, S. N., Pavelsky, T. M., Dugan, H. A., Yang, X., Gardner, J., & Ross, M. R. (2021). Shifting patterns of summer lake color phenology in over 26,000 US lakes. *Water Resources Research*, *57*(5), e2020WR029123. https://doi.org/10.1029/2020WR029123

Topp, S. N., Pavelsky, T. M., Stanley, E. H., Yang, X., Griffin, C. G., & Ross, M. R. (2021). Multi-decadal improvement in US lake water clarity. *Environmental Research Letters*, *16*(5), 055025. https://doi.org/10.1088/1748-9326/abf002

Topp, S. N., Pavelsky, T. M., Yang, X., Gardner, J., & Ross, M. R. (2020). LimnoSat-US: A remote sensing dataset for U.S. Lakes from 1984-2020 [Dataset]. https://doi.org/10.5281/zenodo.4139695

USGS. (2019). *Landsat collection 1 level 1 product definition*. Department of the Interior U.S. Geological Survey.

Wang, S., Li, J., Shen, Q., Zhang, B., Zhang, F., & Lu, Z. (2015). MODIS-based radiometric color extraction and classification of inland water with the Forel-Ule scale: A case study of Lake Taihu. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *8*(2), 907–918. https://doi.org/10.1109/JSTARS.2014.2360564

Wernand, M., & Van der Woerd, H. (2010). Spectral analysis of the Forel-Ule Ocean colour comparator scale. *Journal of the European Optical Society-Rapid Publications*, *5*, 10014s. https://doi.org/10.2971/jeos.2010.10014s

Wetzel, R. G., Likens, G. E., Wetzel, R. G., & Likens, G. E. (2000). Light and temperature. *Limnological Analyses*, 15–32. https://doi.org/10.1007/978-1-4757-3250-4

Weyhenmeyer, G. A., Müller, R. A., Norman, M., & Tranvik, L. J. (2016). Sensitivity of freshwaters to browning in response to future climate change. *Climatic Change*, *134*(1–2), 225–239. https://doi.org/10.1007/s10584-015-1514-z

Williamson, C. E., Overholt, E. P., Pilla, R. M., Leach, T. H., Brentrup, J. A., Knoll, L. B., et al. (2015). Ecological consequences of long-term browning in lakes. *Scientific Reports*, *5*(1), 18666. https://doi.org/10.1038/srep18666

Woerd, H. J., & Wernand, M. R. (2015). True colour classification of natural waters with medium-spectral resolution satellites: SeaWiFS, MODIS, MERIS and OLCI. *Sensors*, *15*(10), 25663–25680. https://doi.org/10.3390/s151025663

Woolway, R. I., Sharma, S., Weyhenmeyer, G. A., Debolskiy, A., Golub, M., Mercado-Bettín, D., et al. (2021). Phenological shifts in lake stratification under climate change. *Nature Communications*, *12*(1), 2318. https://doi.org/10.1038/s41467-021-22657-4

Yang, X., O'Reilly, C. M., Gardner, J. R., Ross, M. R. V., Topp, S. N., Wang, J., & Pavelsky, T. M. (2022). The color of Earth's lakes. *Geophysical Research Letters*, *49*(18), e2022GL098925. https://doi.org/10.1029/2022GL098925

Yao, F., Livneh, B., Rajagopalan, B., Wang, J., Crétaux, J.-F., Wada, Y., & Berge-Nguyen, M. (2023). Satellites reveal widespread decline in global lake water storage. *Science*, *380*(6646), 743–749. https://doi.org/10.1126/science.abo2812

Zhu, Z., Wang, S., & Woodcock, C. E. (2015). Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images. *Remote Sensing of Environment*, *159*, 269–277. https://doi.org/10.1016/j.rse.2014.12.014