ARTICLE

https://doi.org/10.1038/s41586-019-1648-7

Widespread global increase in intense lake phytoplankton blooms since the 1980s

J. C. Ho[1,2]*, A. M. Michalak[1]* & N. Pahlevan[3,4]

Freshwater phytoplankton blooms affect public health and ecosystem services globally[1,2], with harmful impacts resulting either from a bloom's high intensity or the presence of toxin-producing phytoplankton species. Freshwater blooms result in economic losses of over US$4 billion annually in the United States alone, primarily from harm to aquatic food production, recreation and tourism, and drinking-water supplies[3]. Studies documenting bloom conditions in lakes have either focused only on individual or regional subsets of lakes[4–6], or have been limited by lack of long-term observations[7–9]. Here, we use three decades of high-resolution Landsat 5 satellite imagery to investigate long-term trends in intense summertime near-surface phytoplankton blooms for dozens of large lakes globally. We find that peak summertime bloom intensity has increased in a majority (68 per cent) of the lakes studied, revealing a global exacerbation of bloom conditions. Lakes that have experienced a significant ($P < 0.1$) decrease in bloom intensity are rare (8 per cent). The reason behind the increase in phytoplankton bloom intensity remains unclear, however, as temporal trends do not track consistently with temperature, precipitation, fertilizer-use trends, or other previously hypothesized drivers. We do find that lakes with a decrease in bloom intensity warmed less compared to other lakes, suggesting that lake warming may already be counteracting management efforts to ameliorate eutrophication[10,11]. Our findings support calls for water-quality management efforts to account better for the interactions between climate change and local hydrologic conditions[12,13].

The reported incidence of toxic phytoplankton blooms has risen dramatically over the past half-century[14]. While it is generally understood that nutrient loading drives phytoplankton blooms[15], the degree to which bloom conditions are changing globally and the factors driving change among multiple interacting stressors[16] are still uncertain[17]. An understanding of global patterns, trends, and drivers is necessary, however, for designing effective management and remediation strategies[18]. While past studies synthesizing information on long-term trends in lake phytoplankton blooms have been limited by data availability, recent advances in cloud-based parallel computing have made it possible to leverage high-resolution (e.g., 30m) freely accessible satellite imagery over large areas, enabling the study of long-term environmental trends on a global scale[19,20].

Here, we take advantage of these advances to generate long-term record of intense, near-surface phytoplankton blooms for dozens of large lakes across the globe. We use data from the Landsat 5 satellite to generate time series of peak summer bloom intensity from 1984 to 2012 for 71 lakes in 33 countries across six continents (Fig. 1). In total, the data span 30,922 scenes and 72.6 billion lake pixels. The study lakes span a broad range of physical characteristics and degree of anthropogenic impacts (Supplementary Table 1) (see Methods for a full description of the implemented approach). Seasonal peak bloom intensity for a given lake and year is defined based on the maximum observed lake-wide near-infrared signal magnitude, with a first-order correction of atmospheric interference using the shortwave-infrared signal[21]. Remotely-sensed observations within the near-infrared part of the electromagnetic spectrum are sensitive to intense, near-surface algal blooms (see Methods). An initial superset of 154 lakes was selected based on their inclusion in earlier studies leveraging satellite remote sensing[22,23], thereby reducing the likelihood of persistent cloudiness obscuring images. These lakes all have surface areas over 100km^2; globally, lakes within this size range contain approximately 95% of all water stored in lakes[24]. Lakes for which little signal was observed throughout the study period as well as those for which the signal was far outside the range over which the original algorithm was designed[21] were eliminated. A smaller number of additional lakes were eliminated due to previously documented evidence of a lack of phytoplankton blooms. Of the final selected lakes, 38 have a documented presence of harmful cyanobacterial species, while the rest have evidence of other phytoplankton species (10 lakes) or no reported evidence whatsoever (23 lakes). Given the heterogeneity in lake characteristics, each lake's interannual bloom intensity time series is normalized by its own long-term mean and standard deviation to assess relative change over time. This approach eliminates the need to compare absolute magnitudes across lakes, which has been an important barrier to past syntheses across lakes[25].

We find that the implemented algorithm is able to successfully capture previously documented spatial gradients in phytoplankton bloom severity within individual lakes and temporal trends in phytoplankton bloom intensity for specific lakes (see Methods, Extended Data Figure 1). Using atmospheric radiative transfer simulations, we also find that the algorithm is insensitive to reported variations in Landsat 5 orbit or image radiometric quality, primarily owing to the strong signal arising from intense, near-surface blooms identified in study lakes (see Supplementary Information). These results suggest that a single algorithm can indeed identify major phytoplankton blooms despite the large differences in optical properties across lakes[26], as long as the focus is on interannual rather than inter-lake variability. This lends support for the approach implemented in this study for tracking long-term trends globally. In the following, we use all 71 study lakes to assess global trends in summertime peak phytoplankton bloom intensity. We also use a subset of 49 lakes with at least 14 years of data to explore more detailed historical temporal patterns in phytoplankton bloom conditions, with the 14-year threshold selected based on studies of global lake temperatures[22,27].

[1]Department of Global Ecology, Carnegie Institution for Science, Stanford, CA, USA. [2]Department of Civil and Environmental Engineering, Stanford University, Stanford, CA, USA. [3]NASA Goddard Space Flight Center, Greenbelt, MD, USA. [4]Science Systems and Applications Inc, Lanham, MD, USA. *e-mail: jeffho@stanford.edu; michalak@stanford.edu

We find that peak summertime phytoplankton bloom intensity has increased in over two-thirds of study lakes since the 1980s (48 of 71 lakes) (Fig. 1). Increases in bloom intensity are statistically significant for close to a third of all lakes ($p<0.1$ for 22 of 71 lakes), whereas only six lakes exhibited a statistically significant decrease in intensity ($p<0.1$). A similar proportion of lakes has increasing intensity among those with a documented presence of cyanobacteria (24 of 38 lakes) as those without (24 of 33 lakes), and the proportion of lakes with increases in bloom intensity is also consistent across lakes of different areas, volumes, mean and maximum depths, and latitudes (see Supplementary Table 1, Supplementary Information). These results suggest that the observed trends are widespread globally and across lake types, in contrast to earlier hypotheses of differential impacts as a function of latitude[28] or morphometry[29]. This finding provides a global perspective that is consistent with surveys of sedimentary records across temperate-subarctic lakes[6] showing sharp increases in cyanobacterial pigment concentration after 1985. This finding also corroborates putative trends of increasing harmful cyanobacterial blooms globally[17], and counters the hypothesis that increased reporting of toxic blooms is instead a byproduct of increased scientific attention[30].

We find that lake phytoplankton bloom histories follow one of four prototypical pathways, termed here "sustained improvement," "improvement then deterioration," "deterioration," and "no significant trend" (Fig. 2a–d; see Methods). The two pathways that include deteriorating conditions reveal that increases in peak bloom intensity occurred predominantly in the latter half of the study period (Fig. 2b, c). For example, three quarters of study lakes (51 of 68) with sufficient data over the second half of the study period (1998-2012) exhibited an increase in bloom intensity during this period, while only a third (22 of 66) experienced an increase during the first half (1984-1997). The reason behind the temporal coherence of changes in phytoplankton bloom intensity remains unclear, as temporal trends do not track consistently with temperature, precipitation, fertilizer use trends, satellite data availability, or geomorphological characteristics for individual lakes (Extended Data Figures 2-5; see also Supplementary Information), nor are there widespread trends in the seasonal timing of peak bloom intensity (see Supplementary Information).

We find that although lakes that exhibited "sustained improvement" were rare ($n=6$), they experienced less warming (or more cooling) relative to those that exhibited "improvement then deterioration" ($p=0.09$) (Fig. 3; Extended Data Figure 6), suggesting that lake warming may have counteracted management efforts in the latter group. This finding suggests that nutrient reduction targets based on historical relationships between bloom severity and nutrient loading may have to be revisited in the context of climate change, as has been hypothesized[11]. Generalizing the impact of warming across a wide range of lakes is inadvisable, however, as trends across the full lake ensemble showed little direct correlation with temperature (Fig. 4; Extended Data Figures 2, 3; see also Supplementary Information). Rather, these findings suggest that the effects of global lake warming differ depending on lake-specific characteristics[31], and highlight the importance of assessing the role of lake attributes in modulating the impact of temperature on nutrient-phytoplankton relationships[32].

Overall, this study provides global view of trends in intense lacustrine near-surface phytoplankton blooms over the last three decades. We examine bloom histories for lakes with widely differing characteristics and geographic locations, and demonstrate the promise of long-term satellite observations for tracking intense bloom conditions across a heterogeneous set of systems to augment geographically- and temporally-limited *in situ* monitoring efforts. Our results corroborate the putative reported increase in bloom occurrence and intensity globally, and highlight that lakes that have exhibited a long-term decrease in bloom intensity are rare. Results further show that sustained decreases in bloom intensity are more likely to have occurred in lakes with little or no warming, suggesting that rising lake temperatures may hamper environmental recovery, and illustrating the importance of identifying factors that make some lakes more susceptible to the effects of warming.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1648-7.

Received: 11 January 2018; Accepted: 6 August 2019;
Published online 14 October 2019.

1. Pick, F. R. Blooming algae: a Canadian perspective on the rise of toxic cyanobacteria. *Can. J. Fish. Aquat. Sci.* **73**, 1149–1158 (2016).
2. Ndlela, L. L., Oberholster, P. J., Van Wyk, J. H. & Cheng, P. H. An overview of cyanobacterial bloom occurrences and research in Africa over the last decade. *Harmful Algae* **60**, 11–26 (2016).
3. Kudela, R. M. et al. *Harmful Algal Blooms. A Scientific Summary For Policy Makers.* (IOC/UNESCO, 2015).
4. Hampton, S. E. et al. Sixty years of environmental change in the world's largest freshwater lake - Lake Baikal, Siberia. *Glob. Chang. Biol.* **14**, 1947–1958 (2008).
5. Duan, H. et al. Two-Decade Reconstruction of Algal Blooms in China's Lake Taihu. *Environ. Sci. Technol.* **43**, 3522–3528 (2009).
6. Taranu, Z. E. et al. Acceleration of cyanobacterial dominance in north temperate-subarctic lakes during the Anthropocene. *Ecol. Lett.* **18**, 375–384 (2015).
7. Carvalho, L. et al. Sustaining recreational quality of European lakes: Minimizing the health risks from algal blooms through phosphorus control. *J. Appl. Ecol.* **50**, 315–323 (2013).
8. Beaulieu, M., Pick, F. & Gregory-Eaves, I. Nutrients and water temperature are significant predictors of cyanobacterial biomass in a 1147 lakes data set. *Limnol. Oceanogr.* **58**, 1736–1746 (2013).
9. Kosten, S. et al. Warmer climates boost cyanobacterial dominance in shallow lakes. *Glob. Chang. Biol.* **18**, 118–126 (2012).
10. Posch, T., Köster, O., Salcher, M. M. & Pernthaler, J. Harmful filamentous cyanobacteria favoured by reduced water turnover with lake warming. *Nat. Clim. Chang.* **2**, 809–813 (2012).
11. Paerl, H. W. et al. Mitigating cyanobacterial harmful algal blooms in aquatic ecosystems impacted by climate change and anthropogenic nutrients. *Harmful Algae* **54**, 213–222 (2016).
12. Winder, M. Lake warming mimics fertilization. *Nat. Clim. Chang.* **2**, 771–772 (2012).
13. Paerl, H. W. & Huisman, J. Blooms like it hot. *Science.* **320**, 57–58 (2008).
14. Carmichael, W. in *Cyanobacterial Harmful Algal Blooms: State of the Science and Research Needs* (ed. Hudnell, H. K.) 105–125 (Springer-Verlag New York, 2008).
15. Schindler, D. W., Carpenter, S. R., Chapra, S. C., Hecky, R. E. & Orihel, D. M. Reducing Phosphorus to Curb Lake Eutrophication is a Success. *Environ. Sci. Technol.* **50**, 8923–8929 (2016).
16. Winter, J. G., Young, J. D., Landre, A., Stainsby, E. & Jarjanazi, H. Changes in phytoplankton community composition of Lake Simcoe from 1980 to 2007 and relationships with multiple stressors. *J. Great Lakes Res.* **37**, 63–71 (2011).
17. Huisman, J. et al. Cyanobacterial blooms. *Nat. Rev. Microbiol.* **16**, 471–483 (2018).
18. McCrackin, M. L., Jones, H. P., Jones, P. C. & Moreno-Mateos, D. Recovery of lakes and coastal marine ecosystems from eutrophication: A global meta-analysis. *Limnol. Oceanogr.* **62**, 507–518 (2017).
19. Gorelick, N. et al. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **202**, 18–27 (2017).
20. Zhu, Z. et al. Benefits of the free and open Landsat data policy. *Remote Sens. Environ.* **224**, 382–385 (2019).
21. Ho, J. C., Stumpf, R. P., Bridgeman, T. B. & Michalak, A. M. Using Landsat to extend the historical record of lacustrine phytoplankton blooms: A Lake Erie case study. *Remote Sens. Environ.* **191**, 273–285 (2017).
22. Schneider, P. & Hook, S. J. Space observations of inland water bodies show rapid surface warming since 1985. *Geophys. Res. Lett.* **37**, (2010).
23. Sharma, S. et al. A global database of lake surface temperatures collected by in situ and satellite methods from 1985–2009. *Sci. Data* **2**, 150008 (2015).
24. Messager, M. L., Lehner, B., Grill, G., Nedeva, I. & Schmitt, O. Estimating the volume and age of water stored in global lakes using a geo-statistical approach. *Nat. Commun.* **7**, 1–11 (2016).
25. Hampton, S. Understanding Lakes Near and Far. *Science.* **342**, 815–816 (2013).
26. Spyrakos, E. et al. Optical types of inland and coastal waters. *Limnol. Oceanogr.* **63**, 846-870 (2017).
27. O'Reilly, C. M. et al. Rapid and highly variable warming of lake surface waters around the globe. *Geophys. Res. Lett.* **42**, 10, 773-10,781 (2015).
28. Downing, J. in *Global Environmental Change* (ed. Freedman, B.) 221–229 (Springer Netherlands, 2014).
29. Kraemer, B. M. et al. Morphometry and average temperature affect lake stratification responses to climate change. *Geophys. Res. Lett.* **42**, 4981–4988 (2015).

30. Fristachi, A. et al. in *Cyanobacterial Harmful Algal Blooms: State of the Science and Research Needs* (ed. Hudnell, H. K.) 45–103 (Springer-Verlag New York, 2008).
31. Adrian, R. et al. Lakes as sentinels of climate change. *Limnol. Oceanogr*. **54**, 2283–2297 (2009).
32. Rigosi, A., Carey, C. C., Ibelings, B. W. & Brookes, J. D. The interaction between climate warming and eutrophication to promote cyanobacteria is dependent on trophic state and varies among taxa. *Limnol. Oceanogr*. **59**, 99–114 (2014).
33. Wessel, P., Smith, W. H. F., Scharroo, R., Luis, J. & Wobbe, F. Generic Mapping Tools: Improved Version Released. *Eos Trans. AGU* **94**, 409–410 (2013).

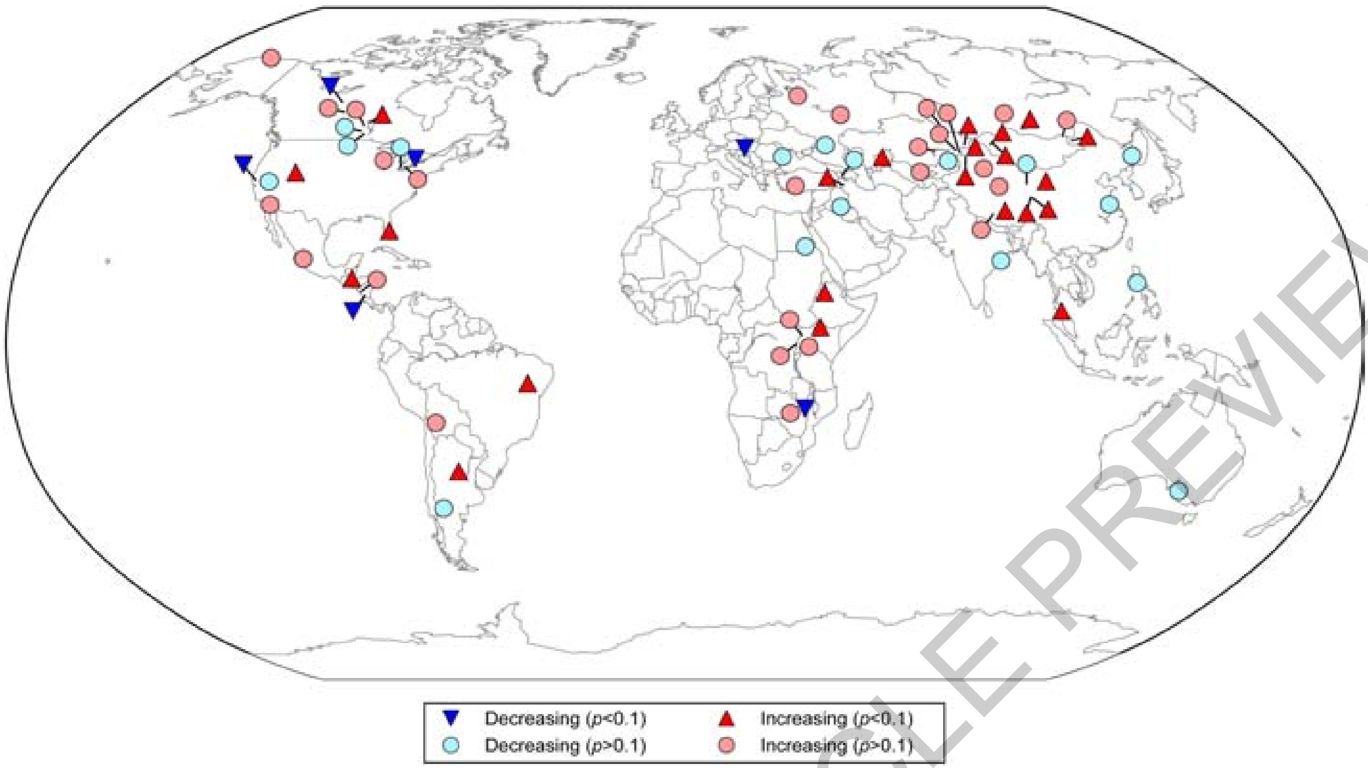


Fig. 1 | Global distribution of lake bloom intensity trends shows peak summertime bloom intensity has increased since the 1980s. The map shows bloom intensity trends for all 71 study lakes over 1984-2012 (Supplementary Table 1). Colors and symbols indicate whether bloom intensity decreased (blue) or increased (red), and whether the trend is statistically significant (triangles for $p<0.1$; circles for $p>0.1$). Basemap generated using Generic Mapping Tools[33].

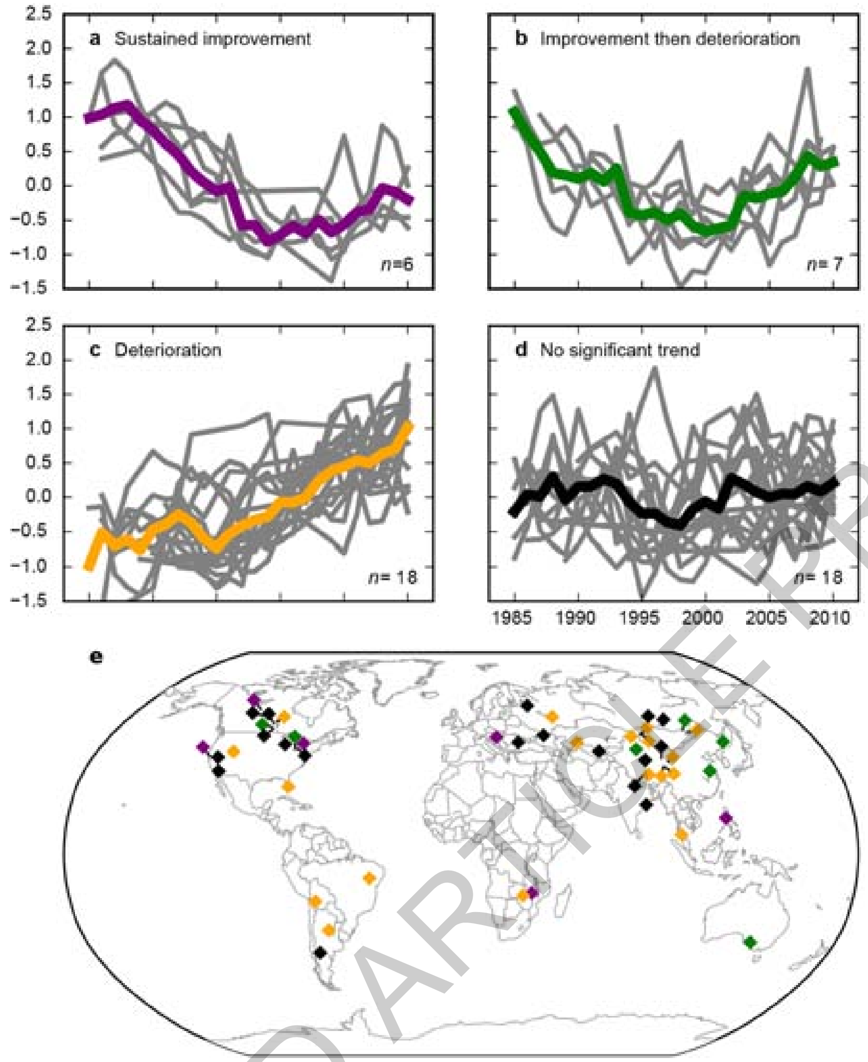


Fig. 2 | Lake bloom histories follow one of four prototypical pathways. (**a**,**b**,**c**,**d**) Time series for lakes with at least 14 years of data (n=49) categorized by historical pathway. Grey lines show 5-year moving average of normalized bloom intensity, with colored lines showing pathway averages across lakes. Each lake's bloom intensity z-score time series is calculated using its own historical mean and standard deviation. (**e**) Global distribution of lake pathways. Basemap generated using Generic Mapping Tools[33].

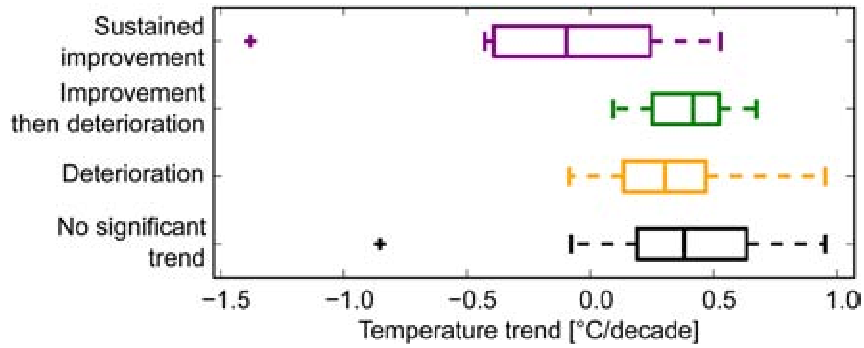


Fig. 3 | Lakes that experienced improvements in bloom conditions tend to have experienced little to no warming. Box plots of water temperature trend (1985-2012) binned by lake historical pathway.

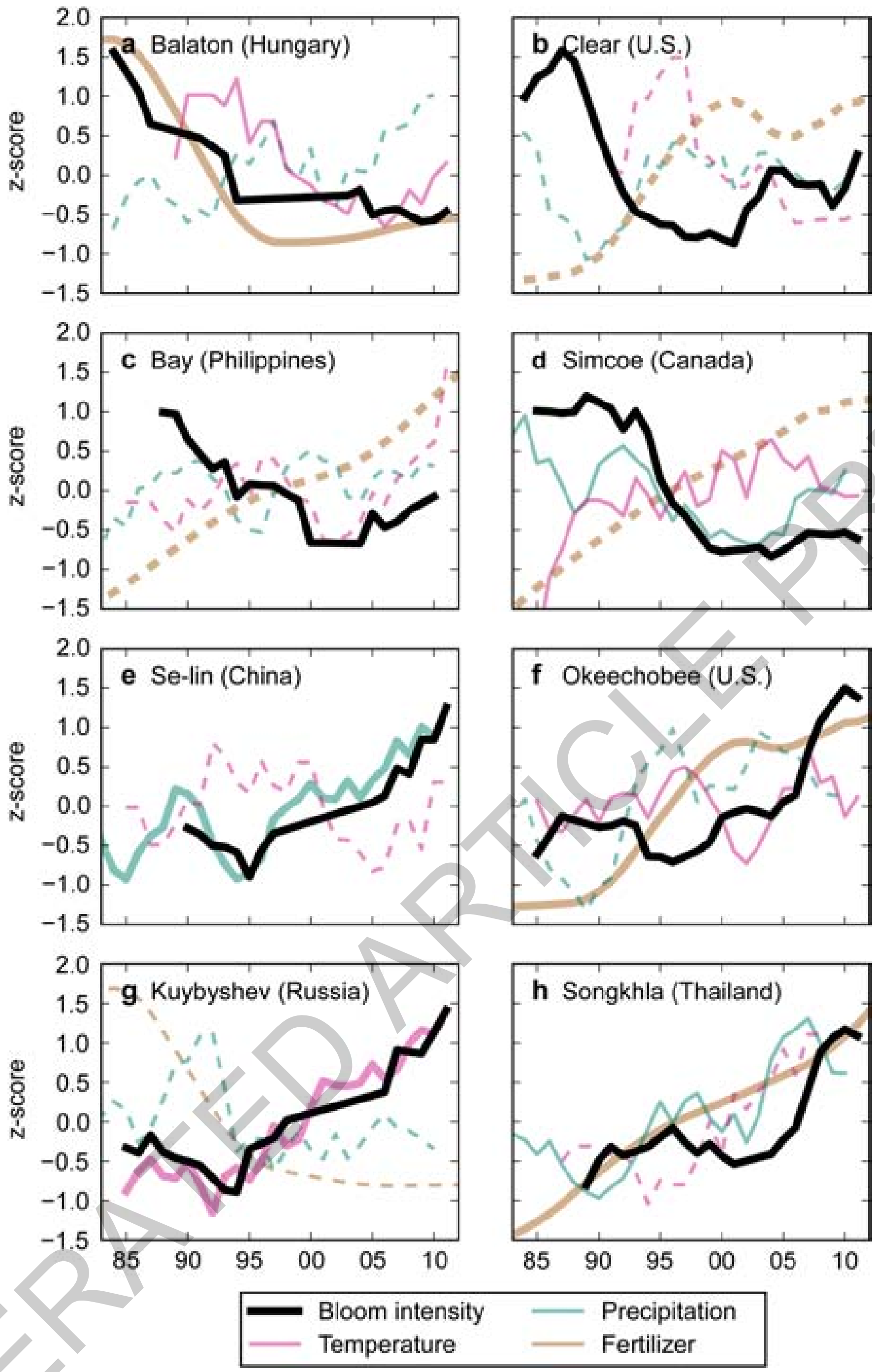


Fig. 4 | Lake bloom histories show no consistent correspondence with temperature, precipitation, and fertilizer use. For eight prototypical lakes, panels show five-year moving average of normalized near-surface bloom intensity, summer lake temperatures, as well as total precipitation and fertilizer application rate over each lake's watershed. Lakes in panels (**a-d**) follow the "sustained improvement" pathway while lakes in panels (**e-h**) follow the "deterioration" pathway. Thicker temperature, precipitation, and fertilizer lines indicate that the Pearson correlation coefficient with bloom intensity is significant (p<0.1). Dashed lines indicate anti-correlations.

METHODS

Satellite source data and implementation of bloom detection algorithm. We used all Landsat 5 Thematic Mapper (L5 TM) images over study lakes (1984-2012) covering the five months encompassing summer (June to October or December to April, depending on lake latitude, similar to a global study of satellite-estimated lake temperatures[22]). Given its long-term archive, L5 TM can be harnessed to assess bloom severity in the upper-layer of the water column[21,34]. Our analysis used images from the L5 TM top-of-atmosphere (TOA) reflectance image collection in Google Earth Engine[35], collected originally from the U.S. Geological Survey[36]. The images in this collection represent unitless planetary TOA reflectance (ρ_λ)[37]:

$$\rho_\lambda = \frac{\pi L_\lambda d^2}{ESUN_\lambda \cos(\theta_s)} \tag{1}$$

where L_λ is the spectral radiance at the sensor's aperture [W/(m^2 sr μm)], d is the Earth-sun distance [astronomical units], $ESUN_\lambda$ is the mean exoatmospheric solar irradiance [W/(m^2 μm)], and θ_s is the solar zenith angle [degrees]. L5 TM Surface Reflectance products[38] were not available worldwide for the full period at the time of algorithm development[21].

We implemented a compositing technique to create images covering the whole surface area of each lake. Based on Landsat's 16-day revisit period, we created composites for each lake by stitching together all L5 TM scenes overlapping with the lake during 16-day time intervals. For each year, intervals started on the first day of the first month analyzed (Either June 1 or December 1) and ran until 144 days later, or 9 intervals total. Overlapping pixels from multiple scenes were averaged in each final composite. Artifacts of this process were observed for some lakes (e.g., horizontal or vertical stripes at scene boundaries), but were not found to substantially impact subsequent analyses.

To mask land within each 16-day composite image, we modified lake polygons from the Global Lakes and Wetlands Database[39]. The original polygons were adjusted manually to fully cover the lake surface area based on images with maximum lake extent over the study period. Within these polygons, we used the Fmask algorithm[40,41] as implemented in Google Earth Engine (Erickson, T.A., USGS Landsat 5 TOA Reflectance (Orthorectified) with Fmask, https://code.earthengine.google.com/dataset/LANDSAT/LT5_L1T_TOA_FMASK, 2016) to exclude cloud and cloud shadow pixels, and to identify the water surface area in each composite. We tested the use of other cloud detection algorithms (e.g., the Landsat Automatic Cloud Cover Assessment procedure[42]), but Fmask resulted in fewer misclassification errors due to turbid water and haze. We also tested static land cover maps to mask land pixels, but the Fmask water layer better accounted for dynamic changes in lake shorelines over time.

We then applied a bloom detection algorithm based on the near-infrared (NIR) band, using the shortwave-infrared (SWIR) band to minimize effects of atmospheric interference and a "greenness" filter to distinguish suspended sediment[21]. The algorithm subtracts the pixel value in the SWIR band (L5 TM Band 5, 1.55-1.75 μm), weighted with an empirical parameter, from the value in the NIR band (L5 TM Band 4, 0.76-0.90 μm) as a measure of near-surface bloom intensity:

$$B = F_G(\rho_{B4} - 1.03\rho_{B5}) \tag{2}$$

where B is the bloom intensity, ranging from 0 to 0.1 where 0.1 generally represents intense near-surface phytoplankton blooms, ρ_{B4} is L5 TM TOA Band 4, ρ_{B5} is L5 TM TOA Band 5, and F_G represents a "greenness" filter that masks out pixels below a certain hue (H) threshold based on L5 TM TOA in Bands 1, 2, and 3:

$$F_G = \begin{cases} 1 \ if \ H < 1.6 \\ 0 \ if \ H > 1.6 \end{cases} \tag{3}$$

$$H = \begin{cases} \dfrac{\rho_{B2} - \rho_{B1}}{\rho_{B3} + \rho_{B2} - 2\rho_{B1}} & if \ \rho_{B1} = \min(\rho_{B1}, \rho_{B2}, \rho_{B3}) \\ \dfrac{\rho_{B3} - \rho_{B2}}{\rho_{B3} + \rho_{B1} - 2\rho_{B2}} + 2 & if \ \rho_{B2} = \min(\rho_{B1}, \rho_{B2}, \rho_{B3}) \\ \dfrac{\rho_{B1} - \rho_{B3}}{\rho_{B2} + \rho_{B1} - 2\rho_{B3}} + 1 & if \ \rho_{B3} = \min(\rho_{B1}, \rho_{B2}, \rho_{B3}) \end{cases} \tag{4}$$

Under intense, near-surface algal bloom conditions, backscattering due to phytoplankton abundance dominates the water-leaving radiance in the NIR region that is otherwise dampened by pure water absorption[43,44] (see Supplementary Information for additional discussion of algorithm sensitivity to *in situ* bloom conditions, also Extended Data Figures 7 and 8). While this approach has proven effective in identifying the extent of near-surface intense phytoplankton blooms, we emphasize that retrievals of concentrations of specific bloom severity metrics (e.g., chlorophyll-a) are beyond the scope of this study.

Selection of initial 76 study lakes. We first selected 154 lakes from those included in a study of global lake temperatures[23] with temperature data collected via satellite. The rationale for choosing these lakes was twofold: 1) there was a lower likelihood of persistent cloudiness obscuring images because these lakes had previously been successfully explored using satellite remote sensing, and 2) the *in situ* temperature data for other lakes were not collected in a consistent way that would be representative of the whole lake (e.g., some data points were collected at point locations on specific shores of the lake).

After an initial exploration of three randomly selected years of composite images for each lake, we sub-selected 95 lakes for analysis based on the criteria that lake pixels have nonzero bloom intensity values below a threshold of 0.1 in a substantial portion of images. We did this to identify lakes that had ranges of bloom intensity values most similar to Lake Erie, where the algorithm was originally validated[21]. Compared to the 59 lakes that were not selected, these 95 lakes in general were shallower and had smaller lake volume, suggesting that the approach may have been more applicable for detecting blooms in shallower lakes.

We then performed a literature search to explore whether the observed bloom intensity signal in lakes was likely to be indicative of real phytoplankton blooms (of any type) or a false positive. We used ISI Web of Science, Google Scholar, and Google Search for the lake name, lake name + "algal bloom", and lake name + "eutrophic". Based on the results of this search, we determined that 78 lakes had either some evidence *for* phytoplankton blooms (51 of 78) or no evidence *against* phytoplankton blooms (27 of 78), while the remaining 17 lakes had strong evidence against the signal representing phytoplankton blooms (e.g., in one lake a high "bloom intensity" signal was erroneously caused by high ice reflectance). Two lakes were also removed from the dataset at this stage based on a lack of L5 TM data during the study period (i.e., data available in only one or two years).

The remaining 76 lakes were selected for further analysis. In these lakes, 54% had support from literature for evidence of cyanobacterial presence or blooms, with 29% specifically dominated by *Microcystis sp*. In total, 10,892 composites were compiled from 30,922 L5 TM scenes during the selected months over study lakes, with a median of 139 composites and 283 scenes per lake. The total number of composites per lake ranged from 24 for Lake Edward (43 scenes total) to 250 for Lake Winnipeg (1771 scenes total). Lakes with a greater number of scenes tended to be in North America, as expected based on historical L5 TM coverage[45]. The number of years with available data per lake ranged from 5 to 28 years with a mean of 21 years (Supplementary Table 1); among lakes with at least 14 years of data, the mean was 22 years per lake.

Validation of well-known spatial gradients of bloom intensity. We further evaluated the proposed approach by comparing geographic regions within lakes with known spatial gradients of bloom intensity. We searched the literature for descriptions of spatial gradients for the 76 study lakes, for example on the basis of chlorophyll-a or phytoplankton biomass observations, and then examined algorithm output values in regions that would be expected to show the largest differences. Using this approach, we identified 48 pairs of regions across 22 lakes (Supplementary Table 2). For instance, in Lake Balaton we identified three regions based on documented chlorophyll-a and biomass gradients from the southwest to the northeast[46–48]. Because the southwest basins are eutrophic to hypertrophic, the northwest basin is mesotrophic, and the middle basin is in between[46], the expectation is therefore that bloom intensities would be higher in a region in the southwest basin relative to a region in the middle basin, which would then also be higher relative to a region in the northeast basin.

We documented the expected bloom intensity in each region qualitatively (i.e., high, medium, or low) as well as the strength of the evidence supporting the expected direction of the gradient between regions (i.e., strong, medium, weak) (e.g., some are based on extensive *in situ* sampling over many years, while others are based on more qualitative inferences). For each of the 48 pairs of regions, we computed the difference in intensity between regions by comparing the mean intensity across all pixels within each region of the lake over the full study period.

The gradient between regions inferred using the implemented algorithm was in the correct direction for over three quarters of region pairs (37 of 48 pairs). For the Lake Balaton example described earlier, all three comparisons between mean pixel values in the three regions were consistent with the expected sign of the difference in bloom intensity (high versus medium, high versus low, medium versus low).

The prior evidence was not strong for 10 out of the 11 pairs of regions for which the gradient was in the direction opposite from that expected. For example, two pairs were for Songkhla Lake, where the evidence on the spatial gradient came from one publication that was not peer-reviewed and another that was based purely on computational modeling rather than *in situ* observations[49,50]. Because the vast majority of the discrepancies in identifying spatial gradients were for region pairs with weaker support in the literature, results indicate that the implemented algorithm is indeed able to identify well-established spatial gradients in bloom intensity across a variety of lakes.

We note that the qualitative approach used here represents a first step toward understanding how global patterns in bloom intensity may be measured using satellite remote sensing. While the algorithm used here was validated quantitatively for data on Lake Erie only, the results of our qualitative approach, in concert with atmospheric radiative transfer simulations (see Supplementary Information), indicate that the algorithm provides a useful signal of bloom intensity for the lakes in this study. Additional validation may be needed for other applications.

Generation of bloom intensity time series and trends. To generate long-term time series of summertime maximum bloom intensities for each lake, we summed the algorithm output values over the whole lake for each composite, as a measure of bloom intensity. This approach assumes that the algorithm output value is correlated with measures such as near-surface chlorophyll-a or phytoplankton biomass, which is supported by validation based on well-known spatial gradients, as described above. We took the largest composite bloom intensity each year as an estimate of summertime maximum bloom intensity, similar to others[21,51]. This approach focuses on the relative spatial intensity of phytoplankton blooms over time, and therefore minimizes the impact of noise in L5 TM images over bodies of water[52].

We removed estimates from each time series when the observed lake surface area was less than 80% of maximum, or less than the mean minus one standard deviation of the whole time series lake surface area, whichever was lower. These guidelines were determined heuristically, based on removing composites that visually had large portions of the lake surface area unavailable due either to missing L5 TM scenes or high cloud cover. We also removed estimates for years where there were fewer than 3 composites due to missing data, as these were expected to be less representative of the summertime maximum. Main findings were not found to be sensitive to minor variations in these thresholds.

We further adjusted for variations in observed lake surface due to clouds or missing L5 TM scenes by dividing the annual bloom intensity estimates by the observed surface area of each lake in each year. This had the benefit of correctly adjusting observed bloom severity trends for lakes where the water surface area changed substantially over time. For example, for a subset of lakes that have dried up during the study period (e.g., Lake Urmia, or the Aral Sea), a decline in water surface area would otherwise be incorrectly observed as a decline in bloom intensity. Normalizing by water surface area more accurately reflects the true bloom conditions in those lakes over time. Although in principle this could also make blooms of constant severity in an otherwise shrinking lake look like an increasing trend, this scenario was not found among study lakes with declining water surface area. Given the highly varying local conditions for study lakes with respect to bloom intensity and water surface area trends with time, normalizing by water surface area provided the best approach overall for accounting for variations due to clouds and missing L5 TM scenes.

Finally, in order to compare data across different lakes, we normalized each lake's annual peak bloom intensity time series by its own long-term mean and standard deviation values, creating bloom intensity z-scores. This is similar to other studies that have treated historical bloom data from remote sensing[53], tracked long-term trends in cyanobacteria[6], or estimated long-term trends of other parameters in lakes[27].

Evaluation of bloom intensity time series and trends. To further evaluate the implemented approach, we compared the temporal evolution of peak bloom intensity in well-studied lakes to those described in existing literature, and also compared bloom intensity trends overall to trends in the SWIR TOA reflectance.

From each time series of normalized annual peak bloom intensity, we tested for the presence of monotonic time trends using the S statistic from the Mann-Kendall trend test[54], and estimated the magnitude of temporal trends using Thiel-Sen's slope[55]. These are both non-parametric procedures known to be more robust to outliers and more accurate for skewed or heteroskedastic data[56], and have been used widely for assessing temporal trends in limnologic studies and those evaluating trends in phytoplankton blooms specifically[6,31,57,58]. Trend analyses over the whole study period were performed for all lakes (n=76).

For lakes with large changes in bloom intensity documented over the study period, the data developed here accurately matched both the direction and timing of changes described in literature. For example, substantial improvements in bloom conditions have been reported for Lake Balaton[59,60], Clear Lake[61,62] and Lake Simcoe[16,63], and all showed statistically significant ($p<0.1$) decreases in peak bloom intensity in the data developed here over the same timeframes as described in earlier studies. Lake Balaton's water quality improvement in the 1990s, coincident with sewage controls and a decline in agriculture[60], was reproduced correctly in the time series (Fig. 4a). Clear Lake experienced a similar decline in bloom intensity in the 1990s, likely due to trophic cascade stemming from drought[62], that was also correctly reproduced (Fig. 4b). Furthermore, for Lake Simcoe, an improvement in water quality occurring soon after 1995, coincident with a widespread invasion of zebra mussels[16,64,65], was also successfully reproduced (Fig. 4d). The developed data similarly reproduced histories for lakes with documented increases in bloom intensity over the study period, such as Lake Winnipeg[66] and Lake Baikal[57,67], and captured documented temporal patterns of decreasing and increasing bloom intensity in ecosystems as diverse as Lake Erie[68] and Tsimlyansk Reservoir[69] (Extended Data Figure 9).

We further evaluated the performance of the approach for detecting false positives (i.e., high derived bloom intensity for low bloom intensity spectra) and false negatives (i.e., low derived bloom intensity for high bloom intensity spectra) using the atmospheric radiative transfer simulations described in the Supplementary Information. The analyses demonstrated the robustness of the algorithm in limiting false positives, i.e., correctly identifying instances where no blooms were present. However, the analyses also identified that aerosol optical thickness (AOT) has an impact on the incidence of false negatives whereby higher AOT (e.g., hazy conditions) resulted in an increased likelihood of missing high bloom intensity events.

Because SWIR can be used as a proxy for AOT, we then evaluated whether derived bloom intensity trends in study lakes could erroneously be due to trends in aerosol conditions affecting the likelihood of false negatives. This was accomplished by comparing peak bloom intensity trends to trends in the SWIR TOA reflectance. For five lakes with statistically significant trends ($p<0.1$) in both peak bloom intensity and SWIR TOA reflectance, we found that observed bloom intensity trends coincided with trends in SWIR TOA reflectance, increasing the risk that apparent trends in bloom intensity could be due to a change in the likelihood of false negatives. For these lakes, peak bloom intensity and SWIR trends were in opposite directions (i.e., increasing SWIR resulting in increased likelihood of false negatives consistent with a decreased bloom intensity trend, and vice versa for decreasing SWIR). Decreasing bloom intensity trends were observed in four of the five lakes (Chao, Gaoyou, Taihu, Sarykamshskoye), with an increasing bloom intensity trend observed in the other (Kremenshugskoye). Increasing SWIR trends in the four former lakes were consistent with trends in aerosols in Eastern China and central Asia over the study period, while a reduction in aerosols in Eastern Europe[70] was consistent with the latter. Because we could not say whether or not the observed bloom intensity trends were in fact attributable to trends in SWIR, to be conservative we removed these five lakes from subsequent analysis. This resulted in a final set of n=71 lakes for further analysis. For trends estimated over shorter periods (1984-1997 and 1998-2012), slightly fewer lakes were used (n=66 and n=68, respectively) because at least two years of observations per period were required to compute trends. A subset of lakes with at least 14 years of data (n=49) was also used to explore temporal patterns in peak bloom intensity.

Beyond this analysis of SWIR trends, we found no other evidence that any potential misclassification of bloom intensity trends occurred. To assess whether our approach could have been incorrectly measuring trends in other environmental variables, we explored historical patterns of potential confounders documented in the literature. Trends in Secchi depth did not match the observed bloom intensity trends (e.g., in Lake Simcoe[71], in Great Salt Lake[72], or in Lake Okeechobee[73]), indicating a lack of evidence to suggest that the implemented algorithm was potentially measuring other constituents of water quality. Nor did we find evidence that global changes in atmospheric constituents, such as aerosols, dust, and water vapor, could explain the overall geographic pattern of bloom observations (except in the five aforementioned lakes) because such constituents have a spatial coherence at large regional scales[70,74–76] whereas lake trends observed here were highly spatially heterogeneous (Fig. 1). For individual lakes, bloom intensity trends also did not track well with observed trends in submerged aquatic vegetation (e.g., Lake Okeechobee[73]), gypsum (e.g., Salton Sea[77]) or cloud cover (e.g., Lake Nicaragua[78]). Taken together, this suggested that our findings about the global proportion of lakes with increasing bloom intensity trends were likely to be robust.

Characteristics of final 71 study lakes. The 71 study lakes (Supplementary Table 1) spanned a wide range of surface areas (158 to 67,052 km^2) and maximum depth (2 to 1637 m) comparable to previous global studies of lakes[29]. Of the 49 lakes with at least 14 years of bloom data, a large majority warmed over the study period (88%) (Extended Data Figure 6), with the temperature trend ranging from -1.40 °C/decade (i.e., cooling) to 0.93 °C/decade. Most of these 49 lakes also experienced an increase in annual precipitation (61%, ranging from -49 to 173mm/decade) while close to half experienced an increase in fertilizer application rate (49%, ranging from -1.47 to 2.41 Mg-N/km^2/decade) (Supplementary Table 1).

Categorization of lakes by prototypical historical pathway. In order to bin the lakes by prototypical historical pathway, we fit a linear model with time for each lake time series using ordinary least squares regression:

$$y = \beta_1 t + \beta_0 \quad (5)$$

where y represents normalized maximum summertime bloom intensity, t represents the year of the observation, and β_1 and β_0 are the fitted model parameters. Bloom intensity values from individual lake pixels (B from Eq. 2) were summed for each image composite, and the maximum summed bloom intensity for each year was used to create the time series y after correcting for missing data, subtracting the long-term mean and dividing by the long-term standard deviation. Lakes for which the linear term was statistically significant ($p<0.1$) were categorized as "sustained improvement" if the peak bloom intensity trend was decreasing with time ($\beta_1<0$) and "deterioration" if the peak bloom intensity trend was increasing with time ($\beta_1>0$).

For the remaining lakes, we fit a quadratic model to each lake time series:

$$y = \beta_2 t^2 + \beta_1 t + \beta_0 \quad (6)$$

where a third term is added indicating a change in peak bloom intensity with t^2. Lakes for which the quadratic term was statistically significant ($p<0.1$) and $\beta_2>0$ were categorized as "improvement then deterioration." The remaining lakes were categorized as "no significant trend."

This approach utilized both simple monotonic trends with time as well as an assessment of the degree of improvement and deterioration to categorize bloom intensity trends. Categorization of lakes ("Eutrophication", "Restoration", "No Change") based on simple changes in lake parameters (increasing, decreasing, no consistent change, respectively) has been used previously to understand long-term trends[6], as have comparisons of multiple measurements in time to assess the balance between historical deterioration and improvement[18].

Data availability

The Landsat 5 Thematic Mapper imagery used in this study is available from the U.S. Geological Survey http://earthexplorer.usgs.gov and through Google Earth Engine https://earthengine.google.com. The bloom intensity trend estimates, historical pathway categories and environmental driver variables generated for each lake and analyzed in this study are provided in Supplementary Table 1. The temperature, precipitation, fertilizer use and lake geomorphological data supporting the findings of this study are publicly available[23,79,80] (see section on 'Environmental driver, watershed, and geomorphological characteristic datasets' in Supplementary Information).

Code availability

Google Earth Engine's web interface allows the bloom detection algorithm[21] to be applied on any Landsat 5 Thematic Mapper images. Access will be provided upon request.

34. Tebbs, E. J., Remedios, J. J. & Harper, D. M. Remote sensing of chlorophyll-a as a measure of cyanobacterial biomass in Lake Bogoria, a hypertrophic, saline-alkaline, flamingo lake, using Landsat ETM+. *Remote Sens. Environ*. **135**, 92–106 (2013).
35. Google. Earth Engine. (2016). at <https://earthengine.google.com/>
36. USGS. Landsat Missions. (2016). at <http://landsat.usgs.gov/>
37. 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*. **113**, 893–903 (2009).
38. USGS. Landsat Surface Reflectance Level-2 Data Products. (2017). at <https://landsat.usgs.gov/landsat-surface-reflectance-data-products>
39. Lehner, B. & Döll, P. Development and validation of a global database of lakes, reservoirs and wetlands. *J. Hydrol*. **296**, 1–22 (2004).
40. Zhu, Z. & Woodcock, C. E. Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sens. Environ*. **118**, 83–94 (2012).
41. Zhu, Z., Wang, S. & Woodcock, C. E. Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4-7, 8, and Sentinel 2 images. *Remote Sens. Environ*. **159**, 269–277 (2015).
42. Irish, R. R. Landsat 7 automatic cloud cover assessment. *AeroSense 2000* **4049**, 348–355 (2000).
43. Moore, T. S. et al. Bio-optical Properties of Cyanobacteria Blooms in Western Lake Erie. *Front. Mar. Sci*. **4**, 1–20 (2017).
44. Gower, J., King, S., Borstad, G. & Brown, L. Use of the 709 nm band of meris to detect intense plankton blooms and other conditions in coastal waters. *Eur. Sp. Agency, (Special Publ. ESA SP* **1161**, 365–368 (2005).
45. Goward, S. et al. Historical record of landsat global coverage: Mission operations, NSLRSDA, and international cooperator stations. *Photogramm. Eng. Remote Sensing* **72**, 1155–1169 (2006).
46. Palmer, S. C. J. et al. Validation of Envisat MERIS algorithms for chlorophyll retrieval in a large, turbid and optically-complex shallow lake. *Remote Sens. Environ*. **157**, 158–169 (2015).
47. Pálffy, K., Présing, M. & Vörös, L. Diversity patterns of trait-based phytoplankton functional groups in two basins of a large, shallow lake (Lake Balaton, Hungary) with different trophic state. *Aquat. Ecol*. **47**, 195–210 (2013).
48. Hajnal, É. & Padisák, J. Analysis of long-term ecological status of Lake Balaton based on the ALMOBAL phytoplankton database. *Hydrobiologia* **599**, 227–237 (2008).
49. Chesoh, S., Lim, A. & Tongkumchum, P. Trend of Water Quality and Model for Forecasting Eutrophication Occurrence in Songkhla Lake, Thailand. in *Proceedings of Taal2007: The 12th World Lake Conference* 834–839 (2008).
50. Suwanidcharoen, S. & Liengcharernsit, W. Development of phytoplankton model with application to Songkhla Lake, Thailand. *Lowl. Technol. Int*. **14**, 50–59 (2012).
51. Stumpf, R. P., Wynne, T. T., Baker, D. B. & Fahnenstiel, G. L. Interannual variability of cyanobacterial blooms in Lake Erie. *PLoS One* **7**, e42444 (2012).
52. Pahlevan, N., Balasubramanian, S. V., Sarkar, S. & Franz, B. A. Toward long-term aquatic science products from heritage Landsat missions. *Remote Sens*. **10**, 1–24 (2018).
53. Palmer, S. C. J. et al. Satellite remote sensing of phytoplankton phenology in Lake Balaton using 10years of MERIS observations. *Remote Sens. Environ*. **158**, 441–452 (2015).
54. Mann, H. B. Nonparametric Tests Against Trend. *Econometrica* **13**, 245–259 (1945).
55. Sen, P. K. Estimates of the Regression Coefficient Based on Kendall's Tau. *J. Am. Stat. Assoc*. **63**, 1379–1389 (1968).
56. Rousseeuw, P. J. & Leroy, A. M. *Robust Regression and Outlier Detection*. (John Wiley & Sons, Inc., 1987).
57. Izmest'eva, L. R. et al. Lake-wide physical and biological trends associated with warming in Lake Baikal. *J. Great Lakes Res*. **42**, 6–17 (2016).
58. Gobler, C. J. et al. Ocean warming since 1982 has expanded the niche of toxic algal blooms in the North Atlantic and North Pacific oceans. *Proc. Natl. Acad. Sci*. **114**, 4975–4980 (2017).
59. Padisák, J. & Koncsos, L. Trend and noise: long-term changes of phytoplankton in the Keszthely Basin of Lake Balaton, Hungary. *Int. Assoc. Theor. Appl. Limnol*. **28**, 194–203 (2002).
60. Tátrai, I., Istvánovics, V., Tóth, L.-G. & Kóbor, I. Management measures and long-term water quality changes in Lake Balaton (Hungary). *Fundam. Appl. Limnol*. **172**, 1–11 (2008).
61. Mioni, C., Kudela, R., Baxa, D. & Sullivan, M. *Harmful cyanobacteria blooms and their toxins in Clear Lake and the Sacramento-San Joaquin Delta (California)*. (2011).
62. Winder, M., Reuter, J. & Schladow, G. *Clear Lake Historical Data Analysis*. (University of California, Davis, 2010).
63. North, R. L. et al. The state of Lake Simcoe (Ontario, Canada): The effects of multiple stressors on phosphorus and oxygen dynamics. *Inl. Waters* **3**, 51–74 (2013).
64. Evans, D. O., Skinner, A. J., Allen, R. & McMurtry, M. J. Invasion of zebra mussel, Dreissena polymorpha, in Lake Simcoe. *J. Great Lakes Res*. **37**, 36–45 (2011).
65. Baranowska, K. A., North, R. L., Winter, J. G. & Dillon, P. J. Long-term seasonal effects of dreissenid mussels on phytoplankton in Lake Simcoe, Ontario, Canada. *Inl. Waters* **3**, 285–296 (2013).
66. Schindler, D. W., Hecky, R. E. & McCullough, G. K. The rapid eutrophication of Lake Winnipeg: Greening under global change. *J. Great Lakes Res*. **38**, 6–13 (2012).
67. Hampton, S. E. et al. Sixty years of environmental change in the world's largest freshwater lake - Lake Baikal, Siberia. *Glob. Chang. Biol*. **14**, 1947–1958 (2008).
68. Allinger, L. & Reavie, E. The ecological history of Lake Erie as recorded by the phytoplankton community. *J. Great Lakes Res*. **39**, 365–382 (2013).
69. Nilson, E. *Investigating potential agricultural-related causes of eutrophication in the Tsimlyansk Reservoir through GIS and remote sensing*. (Central European University, 2014).
70. Pozzer, A. et al. AOD trends during 2001-2010 from observations and model simulations. *Atmos. Chem. Phys*. **15**, 5521–5535 (2015).
71. Guan, X., Li, J. & Booty, W. G. Monitoring Lake Simcoe Water Clarity Using Landsat-5 TM Images. *Water Resour. Manag*. **25**, 2015–2033 (2011).
72. Belovsky, G. E. et al. The Great Salt Lake Ecosystem (Utah, USA): long term data and a structural equation approach. *Ecosphere* **2**, art33 (2011).
73. Havens, K. et al. Extreme weather events and climate variability provide a lens to how shallow lakes may respond to climate change. *Water* **8**, 1–18 (2016).
74. Chin, M. et al. Multi-decadal aerosol variations from 1980 to 2009: A perspective from observations and a global model. *Atmos. Chem. Phys*. **14**, 3657–3690 (2014).
75. Ginoux, P., Prospero, J. M., Gill, T. E., Hsu, N. C. & Zhao, M. Global-Scale Attribution of Anthropogenic and Natural Dust Sources and Their Emission Rates Based on Modis Deep Blue Aerosol Products. *Rev. Geophys*. **50**, 1–36 (2012).
76. Wang, J., Dai, A. & Mears, C. Global water vapor trend from 1988 to 2011 and its diurnal asymmetry based on GPS, radiosonde, and microwave satellite measurements. *J. Clim*. **29**, 5205–5222 (2016).
77. Tiffany, M. A., Ustin, S. L. & Hurlbert, S. H. Sulfide irruptions and gypsum blooms in the Salton Sea as detected by satellite imagery, 1979-2006. *Lake Reserv. Manag*. **23**, 637–652 (2007).
78. Chang, N. Bin, Bai, K. & Chen, C. F. Smart Information Reconstruction via Time-Space-Spectrum Continuum for Cloud Removal in Satellite Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens*. **8**, 1898–1912 (2015).
79. Wei, Y. et al. NACP MsTMIP: Global and North American Driver Data for Multi-Model Intercomparison. (2014). https://doi.org/10.3334/ORNLDAAC/1220
80. Center for International Earth Science Information Network - CIESIN - Columbia University. Gridded Population of the World, Version 4 (GPWv4): Population Density. (2016). https://doi.org/10.7927/H4NP22DQ.

Acknowledgements We thank T. Ballard and D. Del Giudice for discussions and helpful input, Y. Fang and E. Sinha for help with obtaining and processing environmental driver data sets, as well as N. Gorelick and T.A. Erickson for guidance with Google Earth Engine. This research was supported by the National Science Foundation (NSF) under Grant 1313897. Additional support was provided by the Natural Sciences and Engineering Research Council of Canada (NSERC) under a Postgraduate Scholarship-Doctoral award (PGSD3- 438855-2013), by a 2015 Google Earth Engine Research Award, by NASA ROSES Grant NNX16AI16G, and by USGS Landsat Science Team Award 140G0118C0011. CRU-NCEP precipitation was provided by the Multi-scale synthesis and Terrestrial Model Intercomparison Project (MsTMIP; http://nacp.ornl.gov/MsTMIP.shtml). Funding for the MsTMIP activity was provided through NASA Grant NNX10AG01A. Data management support for preparing, documenting, and distributing model driver data was performed by the Modeling and Synthesis Thematic Data Center at Oak Ridge National Laboratory (ORNL; http://nacp.ornl.gov), with funding through NASA Grant NNH10AN681.

Author contributions J.C.H. and A.M.M. designed the research and analysed the results. J.C.H. and A.M.M. wrote the manuscript with input from N.P. J.C.H. performed the majority of the computations with input from A.M.M. N.P. performed the MODTRAN simulations, analysed the MODTRAN results, and wrote the corresponding sections of the methods.

Competing interests The authors declare no competing interests.

Additional information

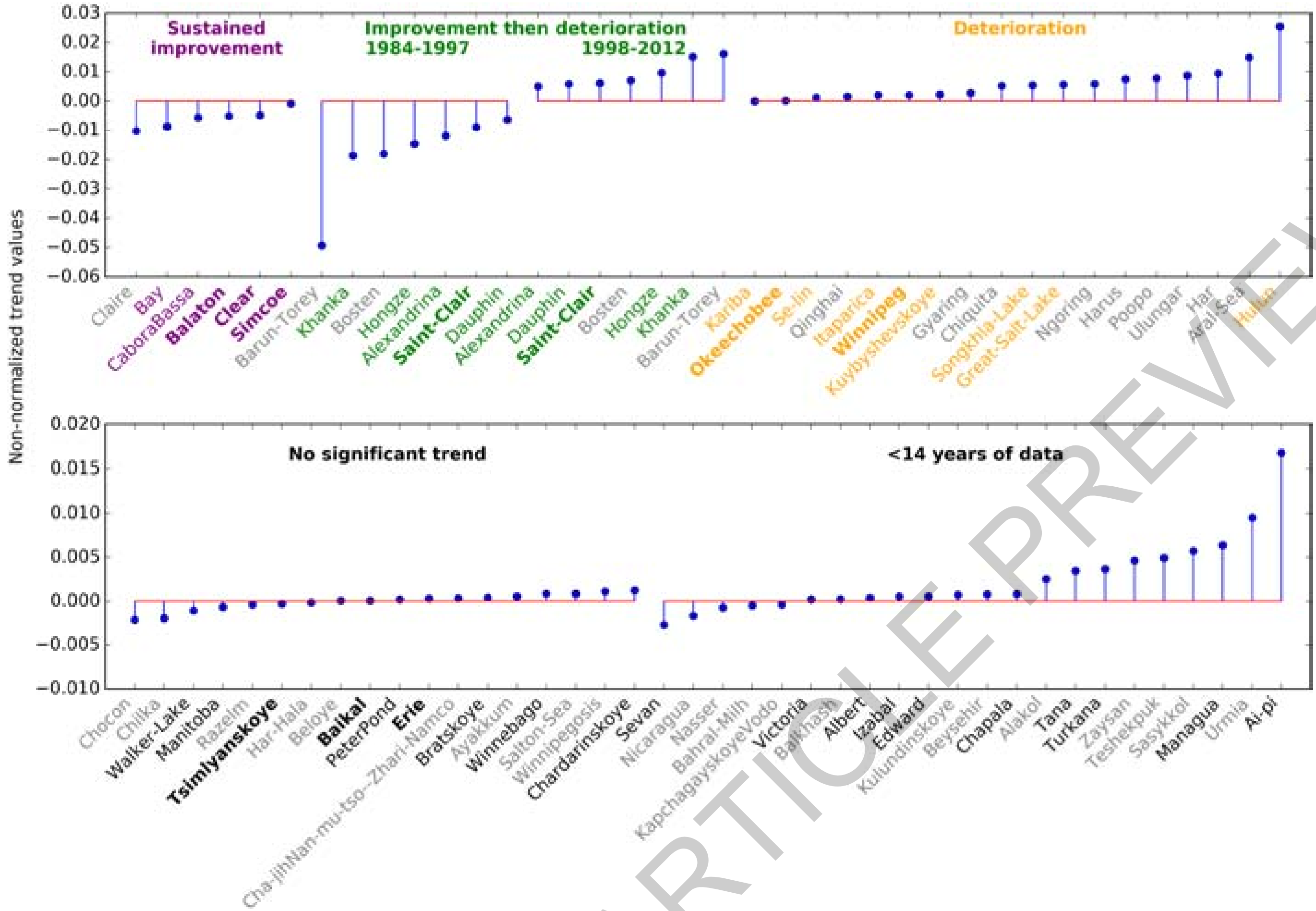
Extended data is available for this paper at https://doi.org/10.1038/s41586-019-1648-7.

Supplementary information is available for this paper at https://doi.org/10.1038/s41586-019-1648-7.

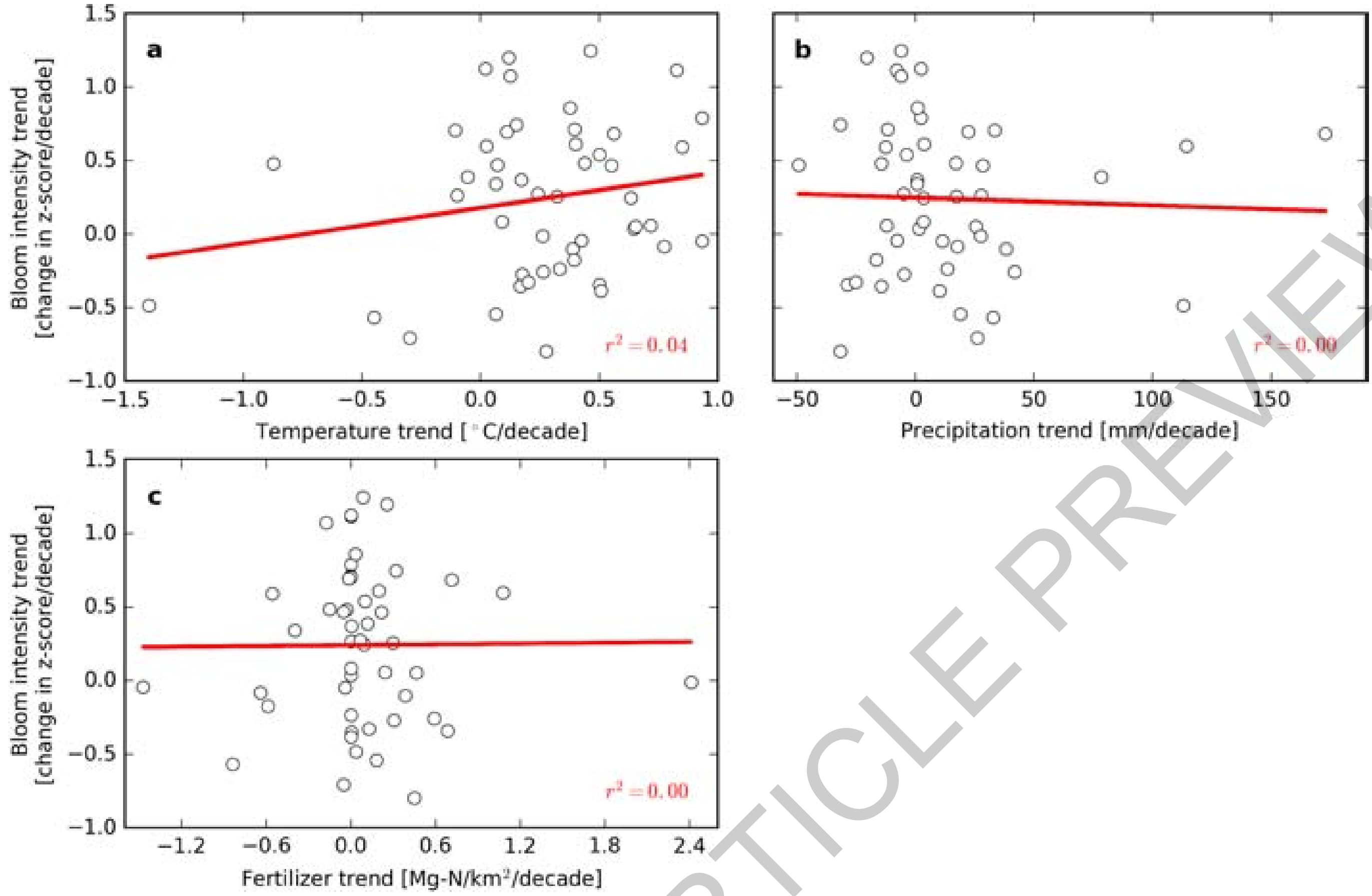
Correspondence and requests for materials should be addressed to J.C.H. or A.M.M.

Peer review information *Nature* thanks Xi Chen and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

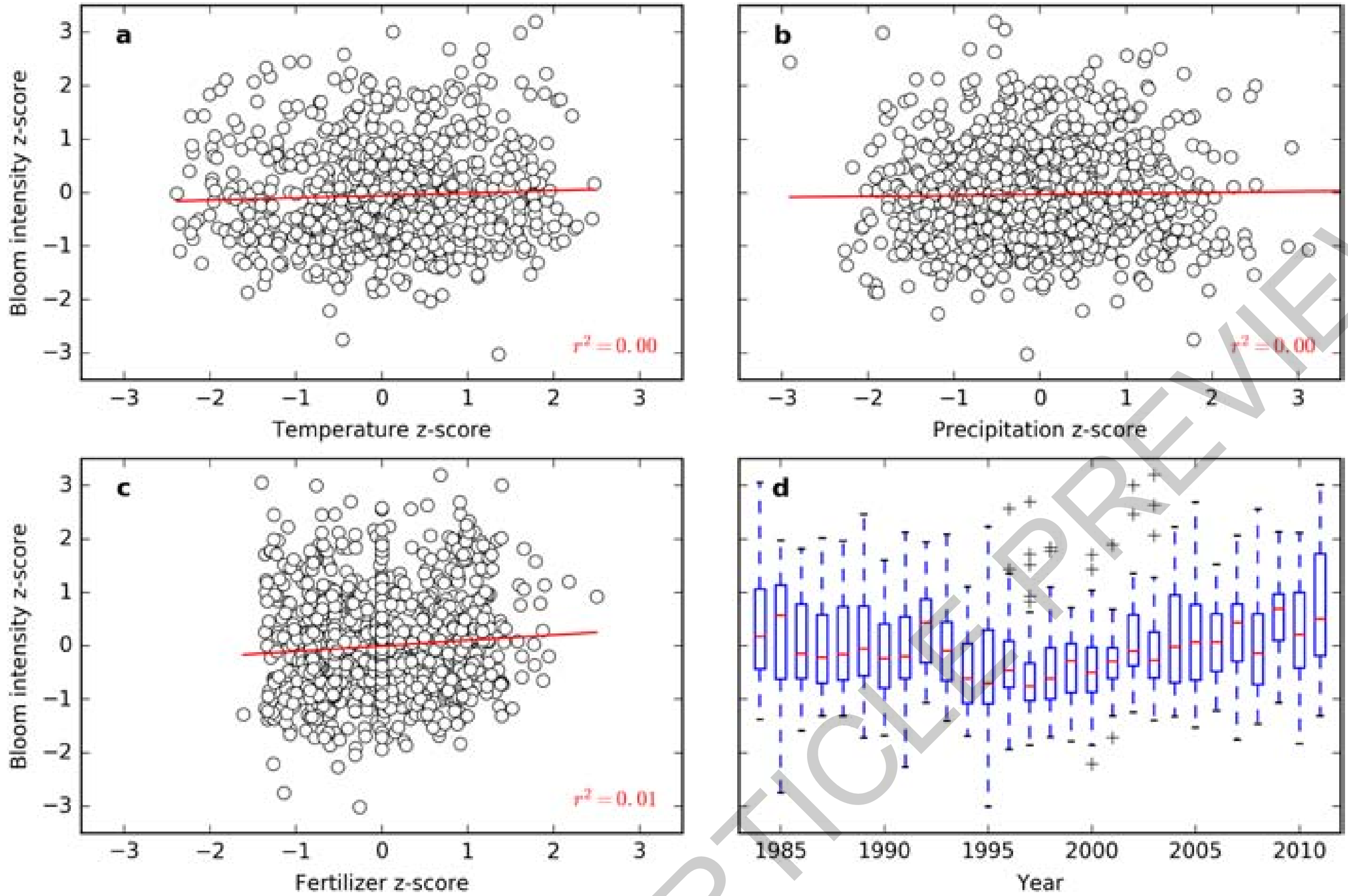
Reprints and permissions information is available at http://www.nature.com/reprints.



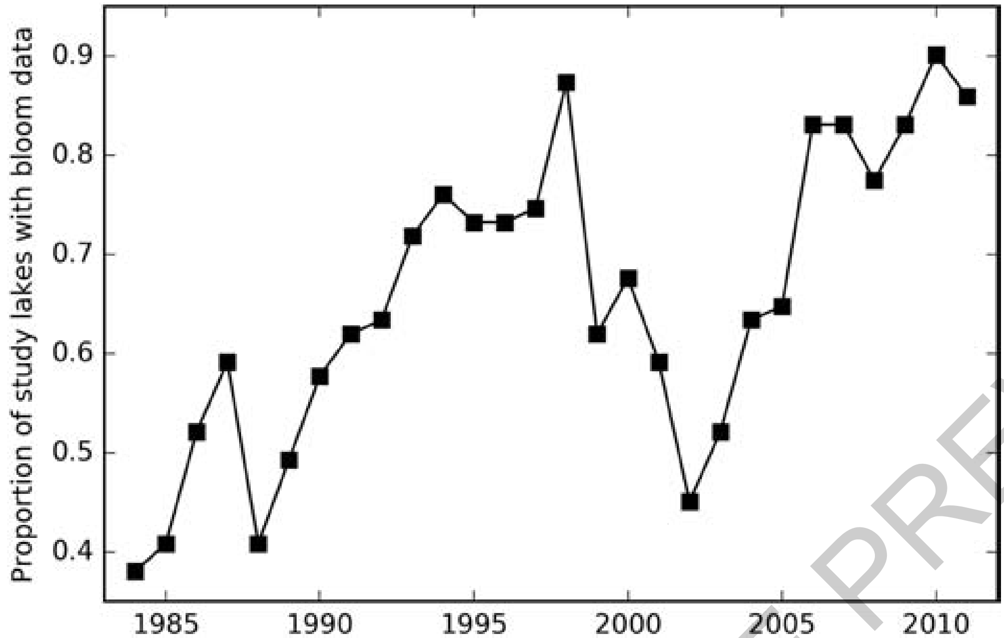
Extended Data Fig. 1 | Lakes with evidence of cyanobacteria (names in color) and with well-documented evidence of major ecological change (bolded names) have trends of a similar magnitude to those in other lakes. Vertical axis shows temporal trends in peak bloom intensity before normalization for all 71 study lakes, categorized by historical pathway. These temporal trends are the Thiel-Sen's slope values calculated using the maximum summertime lake-wide bloom intensity time series for each lake. Trends for lakes following the "improvement then deterioration" pathway are separated into trends over 1984-1997 and 1998-2012 to show trend values in each sub-period separately.



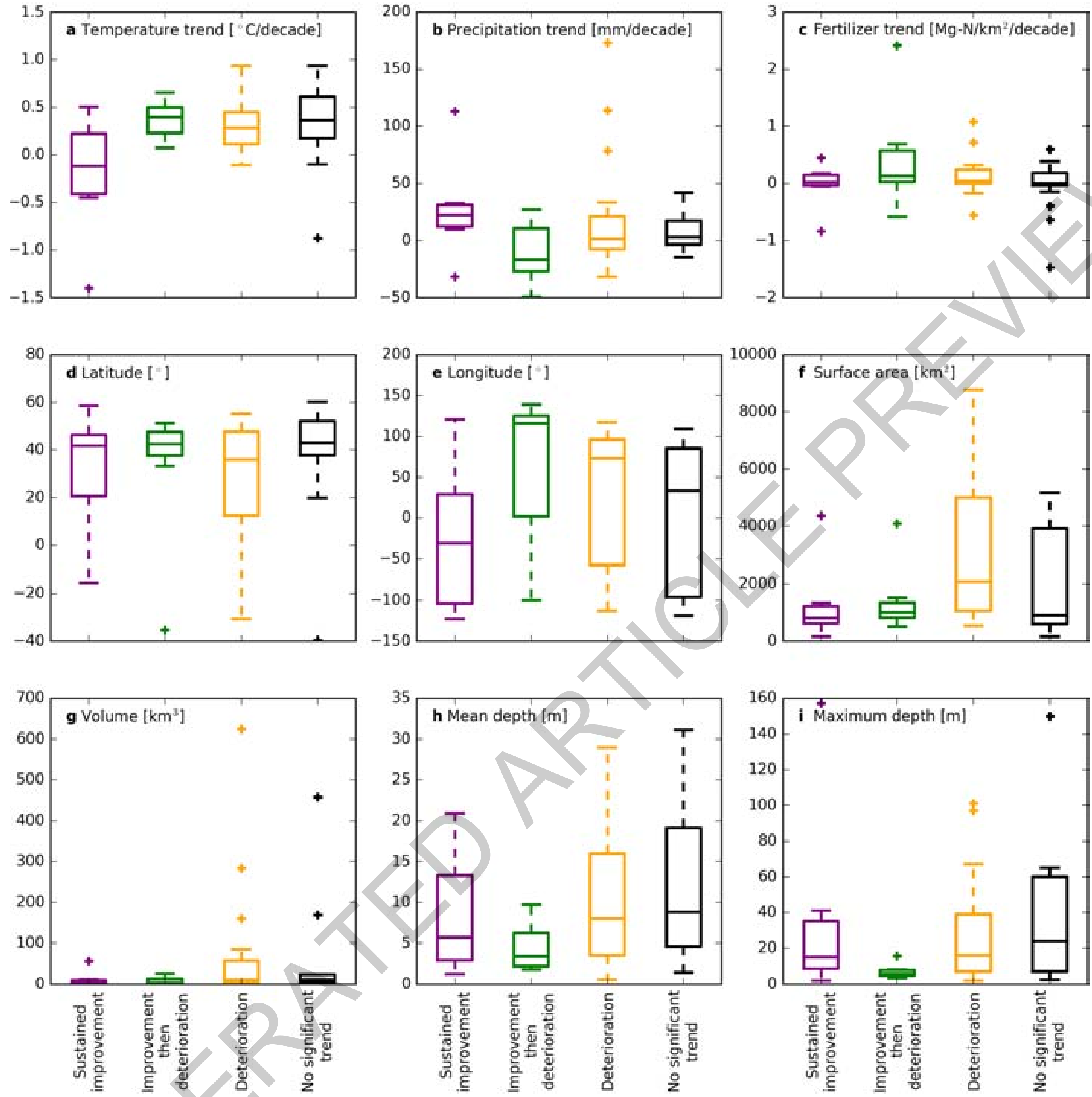
Extended Data Fig. 2 | Low correlations between trends in bloom intensity and environmental factors. Scatter plots of bloom intensity trend versus (**a**) temperature trend, (**b**) total precipitation trend, and (**c**) fertilizer application trend for study lakes with at least 14 years of data (n=49) where each circle represents one lake. Red lines indicate linear fit of white circles.



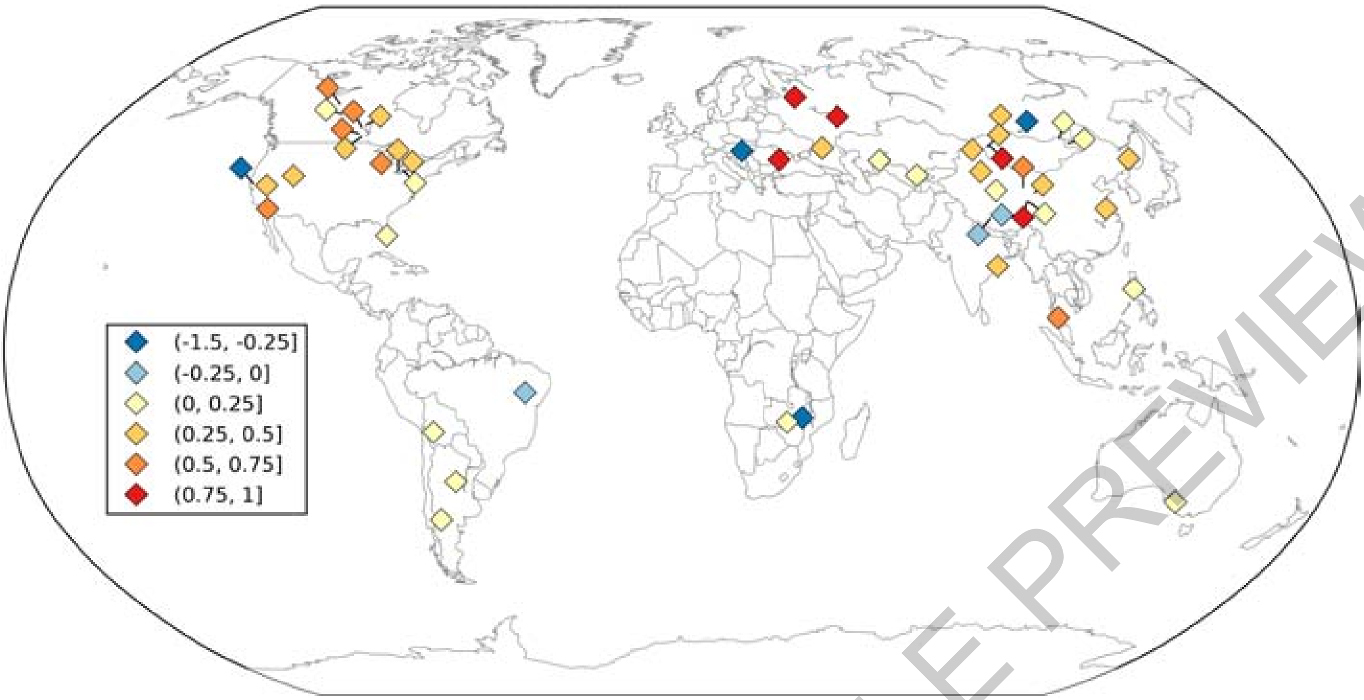
Extended Data Fig. 3 | No relationship observed between bloom intensity and environmental factors aggregated from all lakes. (**a**,**b**,**c**) Scatter plots of bloom intensity *z*-score versus (**a**) temperature (n=784), (**b**) precipitation (n=936), and (**c**) fertilizer (n=980) *z*-scores, where each circle represents one year for one lake. Each lake variable *z*-score is calculated using the mean and standard deviation of its own time series. Red lines indicate linear fit of white circles. (**d**) Box plots of bloom intensity z-score (n=980 total), where each box plot shows the distribution of *z*-scores for all lakes with available data each year.



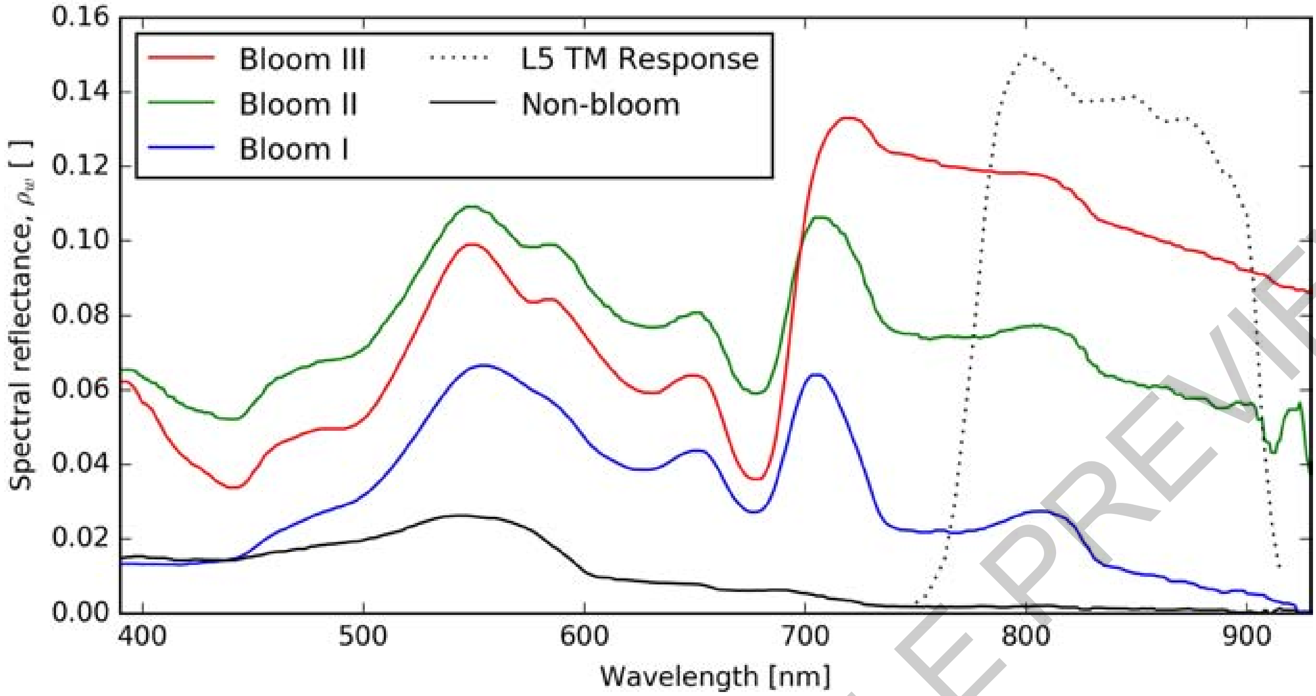
Extended Data Fig. 4 | Availability of bloom intensity data during study period. Number of lakes with a bloom intensity observations after correction for clouds and number of composites (see Methods) divided by the total number of study lakes ($n = 71$) for each year.



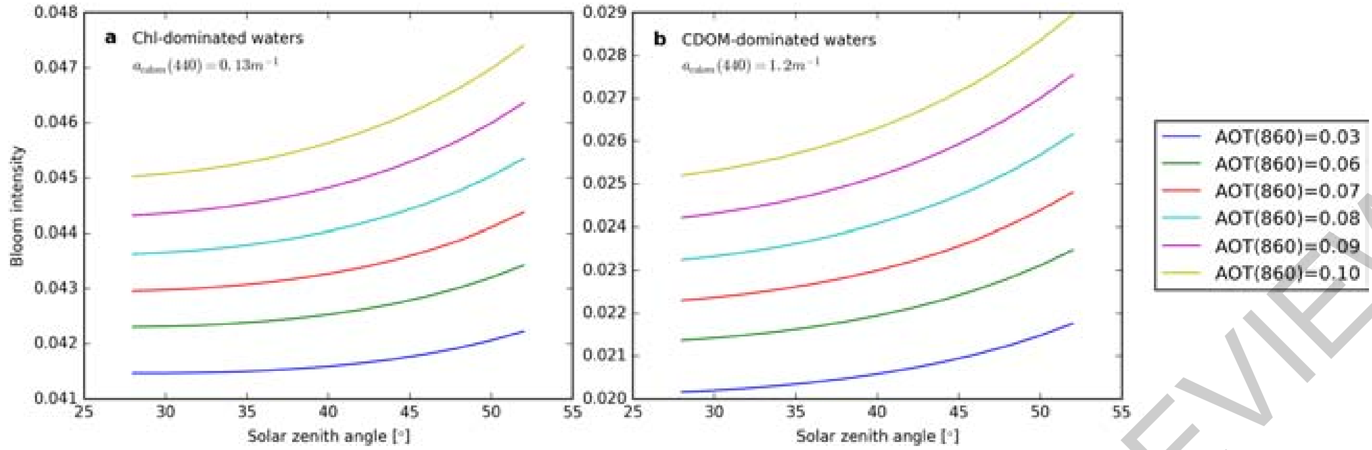
Extended Data Fig. 5 | Box plots showing distributions of lake variables by historical pathway. (**a-c**) Environmental drivers, and (**d-i**) geomorphological factors. Panel (**a**) is equivalent to Fig. 3.



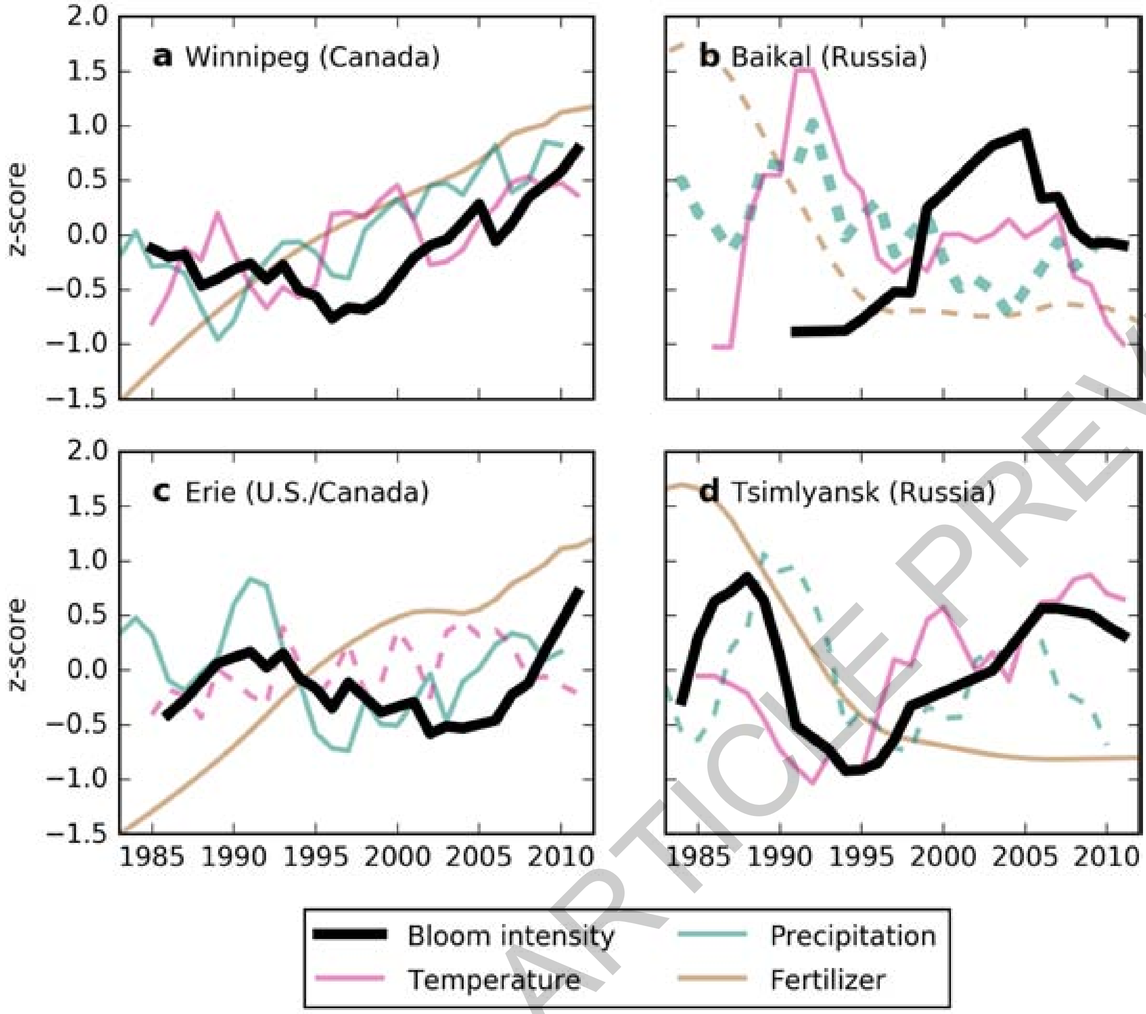
Extended Data Fig. 6 | Global distribution of trends in lake temperature. For the lakes with at least 14 years of bloom data (n=49), the maps show the temporal trend in lake surface water temperature [°C/decade]. Basemap generated using Generic Mapping Tools[33].



Extended Data Fig. 7 | Spectral reflectance curves used by simulations to test algorithm sensitivity. Spectral reflectance curves (ρ_w [])associated with three phytoplankton bloom concentrations and one non-bloom water condition measured in Lake Erie are shown[43]. Bloom I, II, and III correspond to near-surface chlorophyll-a concentrations of 100.1 mg m^{-3}, 143.7 mg m^{-3}, and 106.3 mg m^{-3}, respectively, and total suspended solid concentrations of 30.1 g m^{-3}, 20.0 g m^{-3}, and 22.7 g m^{-3}, respectively. The non-bloom curve corresponds to chlorophyll-a and total suspended solid concentrations of 5.8 mg m^{-3} and 1.8 g m^{-3}, respectively. The normalized spectral response of the NIR channel of L5 TM is also shown. The spectra used in the sensitivity analyses demonstrate the robustness of the bloom intensity measure used in this study (Eq. 2).



Extended Data Fig. 8 | Observed bloom intensity shows minimal sensitivity to solar zenith angle changes that would result from a change in Landsat 5 orbit. Sensitivity of derived bloom intensity varies on the order of 0.001 for waters dominated by both (**a**) chlorophyll (Chl) and (**b**) colored dissolved organic matter (CDOM) for changes in solar zenith angle that would be expected due to a change in satellite orbit. The simulated variation due to solar zenith angle is even smaller for coarse aerosol types (i.e., smaller values of aerosol optical thickness, AOT). The environments in panel (**a**) and (**b**) correspond to "Bloom III" and "Bloom I", respectively" in Extended Data Figure 7.



Extended Data Fig. 9 | Similar to Fig. 4, but showing historical bloom intensity patterns for four additional lakes with well-documented temporal trends. Panels show five-year moving average of normalized bloom intensity, summer lake temperatures, as well as total precipitation and fertilizer application rate over each lake's watershed. Thicker temperature, precipitation, and fertilizer lines indicate that the Pearson correlation coefficient with bloom intensity is significant ($p<0.1$). Dashed lines indicate anti-correlations.