Phytoplankton and cyanobacteria abundances in mid-21st century lakes depend strongly on future land use and climate projections

Karan Kakouei1 | Benjamin M. Kraemer1 | Orlane Anneville2 | Laurence Carvalho3 | Heidrun Feuchtmayr4 | Jennifer L. Graham5 | Scott Higgins6 | Francesco Pomati7,8 | Lars G. Rudstam9 | Jason D. Stockwell10 | Stephen J. Thackeray4 | Michael J. Vanni11 | Rita Adrian1,12

1Department of Ecosystem Research, Leibniz-Institute of Freshwater Ecology and Inland Fisheries (IGB), Berlin, Germany
2Food and Environment (INRAE), UMR CARTEL, University of Savoie Mont-Blanc, French National Research Institute for Agriculture, Thonon-les-Bains, France
3UK Centre for Ecology & Hydrology, Lancaster University, Lancaster, UK
4Lake Ecosystems Group, UK Centre for Ecology & Hydrology, Bailrigg, Lancaster, UK
5United States Geological Survey, Troy, New York, USA
6IISD Experimental Lakes Area, Winnipeg, Manitoba, Canada
7Eawag, Department of Aquatic Ecology, Swiss Federal Institute of Water Science and Technology, Dübendorf, Switzerland
8Department of Integrative Biology (IBZ), ETH-Zürich, Zürich, Switzerland
9Department of Natural Resources and the Environment, Cornell University, Ithaca, New York, USA
10Rubenstein Ecosystem Science Laboratory, University of Vermont, Burlington, Vermont, USA
11Department of Biology, Miami University, Oxford, Ohio, USA
12Department of Biology, Freie Universität Berlin, Berlin, Germany

Abstract

Land use and climate change are anticipated to affect phytoplankton of lakes worldwide. The effects will depend on the magnitude of projected land use and climate changes and lake sensitivity to these factors. We used random forests fit with long-term (1971–2016) phytoplankton and cyanobacteria abundance time series, climate observations (1971–2016), and upstream catchment land use (global Clumondo models for the year 2000) data from 14 European and 15 North American lakes basins. We projected future phytoplankton and cyanobacteria abundance in the 29 focal lake basins and 1567 lakes across focal regions based on three land use (sustainability, middle of the road, and regional rivalry) and two climate (RCP 2.6 and 8.5) scenarios to mid-21st century. On average, lakes are expected to have higher phytoplankton and cyanobacteria due to increases in both urban land use and temperature, and decreases in forest habitat. However, the relative importance of land use and climate effects varied substantially among regions and lakes. Accounting for land use and
climate changes in a combined way based on extensive data allowed us to identify urbanization as the major driver of phytoplankton development in lakes located in urban areas, and climate as major driver in lakes located in remote areas where past and future land use changes were minimal. For approximately one-third of the studied lakes, both drivers were relatively important. The results of this large scale study support the best approaches for mitigating the effects of human activity on lake phytoplankton and cyanobacteria will depend strongly on lake sensitivity to long-term change and the magnitude of projected land use and climate changes at a given location. Our quantitative analyses suggest local management measures should focus on retaining nutrients in urban landscapes to prevent nutrient pollution from exacerbating ongoing changes to lake ecosystems from climate change.

**KEYWORDS**
climate change, forecast, freshwater lakes, land use change, machine learning, phytoplankton, cyanobacteria

1 | INTRODUCTION

Lakes across the globe are threatened by multiple stressors (Vörösmarty et al., 2010). For example, land use and climate change alter the flow of nutrients and energy through lentic ecosystems (De Senerpont Domis et al., 2013; Woolway & Merchant, 2019). Catchment land use (agriculture and urban areas) is also a major influence on lake algal biomass (Allan, 2004; Carpenter et al., 1998; Jones et al., 2020; Soranno et al., 2015) and serves as a proxy for lakes’ altered nutrient concentrations (Foley et al., 2005; Vanni et al., 2011). Constructed forest areas may amplify eutrophication by increasing soil nutrient contents (Liu et al., 2008). However, high proportions of upstream forest and grassland areas prevent direct inflow of nutrients into freshwater lakes and can sustain water quality where anthropogenic activities are minor (Wiley et al., 2010). Higher water temperature and stable thermal conditions associated with climate change may influence algal biomass, particularly favoring cyanobacteria, through direct (e.g., promoting algal growth) and indirect (e.g., nutrient release from sediments during anoxia) mechanisms (Kraemer et al., 2015). Nutrient poor deep lakes may experience declining algal biomass because warming can limit nutrient transport from deep waters into the euphotic zone by enhancing thermal stratification and reducing the extent of winter mixing (Kraemer, Chandra, et al., 2017; Salmaso et al., 2018; Tierney et al., 2010; Winder & Sommer, 2012). Oligotrophic lakes might also experience increasing algal biomass due to alterations in nutrient dynamics and warming temperatures (Carey et al., 2012; Sterner et al., 2020). Climate change may also alter algal biomass by affecting the flow of nutrients coming into lakes, which promotes algal growth (Carey et al., 2012; Catalán et al., 2016; Creed et al., 2018; Sterner et al., 2020), and colored dissolved organic matter, which suppresses algal biomass (Weyhenmeyer et al., 2016).

Previous studies underlined the interplay of climate changes, nutrient loading, and trophic state of lakes (Elliott, 2012 and references therein; Richardson et al., 2018). Although the potential impacts of land use change (Sinha & Michalak, 2016; Vanni et al., 2011) and climate warming (Chapra et al., 2017; Kraemer, Chandra, et al., 2017; O’Reilly et al., 2003; Urrutia-Cordero et al., 2017) on lake phytoplankton have been addressed separately as single stressors, the lack of simulated land use and climate data at sufficient spatiotemporal resolution have hampered projections of their potential combined impacts on lake phytoplankton. Recent developments including the Clumondo global land use models (Van Asselen & Verburg, 2013) and ISIMIP2b global time-series climate model data (Frieler et al., 2017; https://www.isimip.org/protocol/#isimip2b) now allow such projections worldwide (e.g., Woolway & Merchant, 2019). These new model output data provide opportunities to disentangle and explore the combined effects of projected land use and climate changes on lake ecosystems.

To test lake ecosystem responses to increasingly severe environmental change, we used long-term observations from 14 European and 15 North American lake basins to project mid-21st century phytoplankton (including cyanobacteria) and in particular cyanobacteria abundance based on a range of land use and climate scenarios. We then applied our projections of phytoplankton and cyanobacterial responses to more than 1500 lakes, each within 160 km of one of the 29 focal lake basins and with similar morphometric characteristics, to compare the responses across different combinations of land use and climate scenarios. We expected an increase in phytoplankton and cyanobacteria abundances under both the best- and worst-case climate and land use scenarios, mainly because temperature and urban areas are projected to increase in the future. We also expected that the relative importance of climate versus land use change impacts on algae would depend on the trophic state and geographic location of lakes. Our assessment enabled a comprehensive estimate
of the scope of lake phytoplankton changes under projected future environmental conditions with implications for lake ecosystem services and management priorities.

2 | MATERIALS AND METHODS

2.1 | Overview

Random forest (RF) models fit with observed phytoplankton and cyanobacteria data from 1971 to 2016 were used to predict phytoplankton (including cyanobacteria) and cyanobacteria responses to projected land use and climate scenarios in 29 lake basins across Europe and North America for the time period 2044 to 2056 (Table S1). We used two scenarios (RCP 2.6 and RCP 8.5) and four earth system models per scenario (GFDL-ESM2M, HadGEM2-ES, IPSL-CM5A-LR, and MIROC5) for climate, and three scenarios for potential future land use changes (sustainability, middle of the road, and regional rivalry) (O’Neill et al., 2017; Van Asselen & Verburg, 2013). Together, the land use and climate data used for model fitting and projections offer rich coverage of the major environmental variables that influence algae in lakes, including agricultural and urban land use (as a proxy for nutrient loading), water temperatures, and light conditions.

2.2 | Study sites

We gathered a suite of long-term climate and biological sample data, and calculated the proportion of upstream watershed land use from 13 European and 11 North American freshwater lakes for which unique physical and geographical characteristics were available (Table S1). Lake Champlain and Lake Zürich were divided into five and two lake basins, respectively, and each basin was analyzed separately, resulting in a total of 29 lake basins. These lake basins vary in their geographic and morphological characteristics, thus allowing coverage of wide environmental gradients.

2.3 | Biological sample data

Phytoplankton and cyanobacteria abundance data (cells L\(^{-1}\)) were collected from 29 focal lake basins at weekly to monthly intervals between 1971 and 2016 (see Table S1 for more information on lakes and total sample size per lake basin). The long-term phytoplankton samples have been consistently taken from the same sampling location at each lake except for Lake Lucerne, where the sampling site was changed in 1998. In Lake Geneva, the sampling depth has changed from the upper 10 m of water column to the upper 18 meters in 2001 because the strata in which the maximum phytoplankton growth occurs have deepened over time (Anneville et al., 2019). In Cheney Reservoir, sample collection was changed from near-surface to the integrated photic zone in 2004 (Graham et al., 2017). Samples were integrated across depth using integrated samplers or by combining separate samples from discrete depths across the epilimnion and hypolimnion. The samples for each lake have been counted under a microscope using consistent methods over the entire sampling duration and have been counted by the same taxonomist whenever possible. Exceptions are Windermere with various sample processors in early years, the Oneida Lake, for which the sample processor changed in 1996 (Idrisi et al., 2016), and lakes Geneva, Annecy, and Bourget, for which sample processors changed in 2007. The data include measures of phytoplankton and cyanobacteria abundances, which are representative for large-scale changes. To allow comparisons across the focal lake basins, abundance data of individual phytoplankton and cyanobacteria taxa from each sampling date were converted to total abundance (cells L\(^{-1}\)). For more detail about the sampling methods and sample processing for the different lakes see the metadata description in the Forest Ecosystem Monitoring Cooperative (FEMC) repository that contains the Global Evaluation of the Impacts of Storms on freshwater Habitat and Structure of phytoplankton Assemblages (GEISHA) database (https://www.uvm.edu/femc/data/archive/project/geisha-stormblitz.fr) and the references listed in Table S1.

2.4 | Land use predictors

We used the Clumondo land use model for Europe and North America, which consists of image-based land use data from the year 2000, and three demand-based future scenarios of (1) sustainability, (2) middle of the road, and (3) regional rivalry for the year 2050 (see O’Neill et al., 2017 for more details on the implementation of scenario storylines in Clumondo). The three scenarios constrain or promote land system conversion as influenced by local factors (Van Asselen & Verburg, 2013). Based on the intensification theory, the Clumondo model focuses on simulating the intensification of agricultural management upon a combination of increasing demands and decreasing land available for expansion of cropland area (Eitelberg et al., 2015; Turner & Ali, 1996). The land use scenarios were specifically adjusted according to the scenario narratives, which describe the demands according to the future evolution of key aspects of society such as human development, economy, lifestyle, policies, and natural resources (O’Neill et al., 2017). For the sustainability scenario, demands were included for tons of crop production, livestock units (goats and sheep), built-up area, and tons of carbon; middle of the road and regional rivalry scenarios included demands for tons of crop production, livestock units (goats and sheep), and built-up area (O’Neill et al., 2017). These three modeled scenarios assume an increasing demand for cropland production resulting from a combination of cropland expansion and cropland intensification (Eitelberg et al., 2015). The Clumondo data are classified at a spatial resolution of 5 arc minutes (~9.25 km), where pixels represent the dominant land use types. For a comprehensive coverage of our study area, we merged the 24 original sub-categories into five distinct categories of cropland, urban, forest, grassland, and bare. For example, all
sub-categories that described urban land use (i.e., peri-urban or villages and urban areas) were considered to form the category "urban" (see Table S2 for details on sub-categories). For each of the 29 focal basins, we calculated the proportion of each land use category in its upstream watershed area. The land use proportions were calculated for Clumondo land use data from the year 2000 (i.e., to be used for the sample data and the baseline period) and the three land use scenarios (i.e., to be used for the future period).

2.5 Climate predictors

All climate data were derived from the hind and forecasted models from the Inter-sectoral Impact Model Intercomparison Project (ISIMIP2b), which provides data according to the impacts of climate change across sectors and spatial scales (Frieler et al., 2017). We downloaded 10 bias-corrected climate variables of mean water column temperature and surface water temperature, near-surface minimum, mean, and maximum air temperature, short- and long-wave radiation, surface wind speed, surface air pressure, and surface relative humidity, which were available at a daily time step and at a grid resolution of 0.5°. Based on current ecological understanding, all these variables are expected to drive total algae and cyanobacteria in lake ecosystems (Kakouei et al., 2020; Litchman & Klausmeier, 2008). Modeled time-series data were extracted for the grid point situated closest to the center of each lake. To account for the uncertainty stemming from climate scenarios or models, we used four climate model projections—GFDL-ESM2M, HadGEM2-ES, IPSL-CM5A-LR, and MIROC5 from the CLM45 impact model for the historic period (1971–2016) and two defined periods of temporal baseline (2006–2018) and horizon 2050 (2044–2056) under two climate scenarios of RCP 2.6 and RCP 8.5. The RCP 8.5 is the most aggressive scenario that might be very unlikely; however, this scenario is reported to be consistent with the historical total cumulative CO2 emissions, thus serving a useful tool for quantifying climate risks over the midterm time horizons (Schwalm et al., 2020). These two scenarios (RCP 2.6 and 8.5) and all four available model projections (GFDL-ESM2M, HadGEM2-ES, IPSL-CM5A-LR, and MIROC5) span a range of potential future global radiative forcing from anthropogenic impacts. Furthermore, using the least severe (RCP 2.6) and the most severe (RCP 8.5) scenarios allowed us to cover the whole range of possible variability in lake phytoplankton and cyanobacteria responses to climate change.

To fit the statistical relationships between the biological sample data and climate data, we needed to add the climate data from the observation period (i.e., 1971–2016) to the biological data set. We extracted the time-series daily mean data for the 10 climate variables according to the historic period (1971–2005), and averaged the values across the four model projections (GFDL-ESM2M, HadGEM2-ES, IPSL-CM5A-LR, and MIROC5). The time-series daily mean data were added to the biological data set according to the day of sampling. As the biological sample data were available until 2016, we filled the gap 2006–2016 using the average of the four model projections from the RCP 6.0, which is reported to align best with the climate conditions over the last decades (Schwalm et al., 2020). These data, which were from the grid point situated closest to the center of each lake, were merged with the biological sample data of the respective lake basin to extract the daily mean value of each climate variable at the sampling dates.

2.6 Lake attributes

Lake attributes are unique values that describe morphological characteristics and geographic location for each of the 29 focal lake basins (gray bars in Figure 1; Table S2). We used the measured values for our focal lake basins. We added these attributes as predictors to our models to enhance their predictive power by discriminating between different phytoplankton abundances and consequently different phytoplankton responses among lakes that might be located in the same area.

2.7 Machine learning approach

We used RF, a non-parametric machine learning (ML) tool, to assess the relative importance of predictor variables and to predict future changes in phytoplankton and cyanobacteria abundances. RF averages the results from an ensemble of decision trees (i.e., bootstrapping), each of which selects a random subset of data and predictor variables for partitioning (Breiman, 2001). The bootstrapping approach in RF minimizes potential issues associated with outliers or collinearity among predictor variables, by using one-third of the variables each time, resulting in a different set of variables for each tree in the RF models (Breiman, 2001). Averaging the results across all trees enables RF to handle complex interactions between the response and predictor variables (Breiman, 2001). In our study, each forest comprised 2000 trees and each tree used one-third of the variables to predict unknown data. The RF node size was set to “5,” which is appropriate for regression models. We used the R-package randomForestSRC (version 3.6.2) for training the model and performing predictions (R Development Core Team, 2016).

The statistical relationships were set up by linking the abundance data to the predictor variables (Table S2). The predictor variables consist of 10 climate predictors, five land use categories, and nine fixed lake attributes variables. To give equal weighting to all predictor variables and prevent strong relative influences for predictors with significantly large values, we normalized the predictor variables without creating a distortion in the final results. For this, we used the scale function from the stats package in R (R Development Core Team, 2016), which standardizes the values of each predictor in the data set by normalizing each predictor column to a mean of zero and a variance of one. Furthermore, the abundance of phytoplankton and cyanobacteria was Log10-transformed prior to fitting the model, and the predictions for the baseline and future periods were back-transformed for the subsequent percent-change calculations.
2.8 | Model cross-validation

To account for potential uncertainties stemming from the ecological models, we did a 100-fold cross-validation by dividing the phytoplankton or cyanobacteria data sets of each lake 100 times into a randomly selected 75% training set and 25% test set. For each of the 100-fold, the training sets of all 29 ake basins were joined into one single data set, resulting in 100 training sets. We trained a model using each of these 100 training sets, and predicted the phytoplankton and cyanobacteria abundance for the test set of each lake separately. We calculated adjusted $R^2$ values for the ordinary least squares regression fit comparing predicted and observed values in the test data set for each of the 100-fold and each lake. Furthermore, we then calculated the Pearson correlation coefficients between predicted and observed values in the test data set (Figure S1).

2.9 | Projected phytoplankton and cyanobacteria changes in 29 focal lake basins

We defined two 13-year periods, one as baseline (2006–2018) and one as future projection (2044–2056) to assess the effects of land use and climate changes on the abundance of phytoplankton and cyanobacteria. The baseline period was defined to account for the natural annual climate fluctuations (i.e., differences between warm and cold years, or wet and dry years) and to guarantee a consistent comparison between the future period and the baseline. Each of the baseline and future periods consisted of lake attributes, proportion of upstream land use data, and climate data for the 29 focal lake basins. The climate data consisted of monthly mean values calculated by averaging the daily values from summer months (June through September) of each year. This data set includes 24 possible combinations consisting of the two climate scenarios (RCP2.6 and RCP8.5), four climate models (GFDL-ESM2M, HadGEM2-ES, IPSL-CM5A-LR, and MIROC5 from the CLM45), and the three land use scenarios (sustainability, middle of the road, and regional rivalry) for each period. The 24 combinations allowed explicit consideration of the potential uncertainties stemming from land use and climate models or scenarios. Considering the 100-fold cross-validation, we predicted 2400 single phytoplankton or cyanobacteria abundance values for each month during the baseline and projected periods. Furthermore, predicting abundances for each month over the 13-year baseline and projected periods incorporated natural interannual fluctuations, such as differences between warm and cold or wet and dry years (Kakouei, 2018). The final 2400 percent-change values for each focal lake basin were calculated by subtracting the value of each of the 2400 combinations averaged over the future time period from the value of the same combination averaged over the baseline period.

In addition to predicting the combined effects of land use and climate change on phytoplankton and cyanobacteria abundances of the focal lake basins, we predicted the single effects of land use and climate change separately. More specifically, we predicted the effects of climate change once land use was kept unchanged for the future period, and vice versa. This approach was performed to assess the relative importance of land use and climate in driving future phytoplankton or cyanobacteria abundances.

2.10 | Projected phytoplankton and cyanobacteria changes in 1567 lakes from focal regions

To assess the regional effects of land use and climate changes on phytoplankton and cyanobacteria, we selected 3942 freshwater lakes within a 160 km buffer around the 29 focal lake basins. The 160 km distance is within the expected dispersal distance for phytoplankton taxa; thus, the 29 lake basins and the 3942 lakes across the focal regions should have comparable phytoplankton communities (Kristiansen, 1996). We excluded 2375 of the 3942 lakes that were within the 160 km buffer because they fell outside of the observed range of physical attributes (elevation, watershed area, lake area, total volume, mean discharge, and mean depth) of the 29 focal lake basins. We calculated the proportion of upstream land use data and gathered climate data and lake attributes (from the Hydrolakes attribute table; Messager et al., 2016) for the remaining 1567 lakes. Using these data, we extrapolated our predictions to these 1567 lakes. We predicted phytoplankton and cyanobacteria abundances for all the 2400 combinations (two climate scenarios, four climate models, three land use models, and 100 ecological models) in the baseline and future time periods. Then, we calculated the differences between the two periods by subtracting each of the 2400 values in the future time period from the value from the same combination in the baseline period. Finally, we calculated the mean of these 2400 percent-change values for each of the 1567 lakes.

3 | RESULTS

3.1 | Relative influences of predictors

Phytoplankton and cyanobacteria abundances from the 29 lake basins differed in their predictability in cross-validation, with respective adjusted $R^2$ of $0.32 \pm 0.23$ (mean $\pm$ SD) and $0.33 \pm 0.21$ when comparing log-transformed predicted to observed values from the test set (Figure S1). Lake temperature and light from climate predictors; cropland, forest, and urban areas from land use predictors; and seasonality (day of year) had the highest relative influence on both phytoplankton and cyanobacteria abundances over the observed 29 time series (Figure 1). Temperature was the top ranked variable for both phytoplankton and cyanobacteria, but the relative influence was substantially higher for cyanobacteria. Moreover, total lake volume and light intensity were in the top five variables for phytoplankton, while cropland and forest were in the top five variables that had relatively high influence on cyanobacteria (Figure 1).
3.2 Projected land use and climate change scenario output

On average, cropland areas in the upstream watersheds of all 1596 studied lake basins were projected to decrease slightly in the sustainability scenario, but to increase slightly in the middle of the road and regional rivalry scenarios (Figure 2a). On average, the upstream forest areas were projected to decrease for all land use scenarios by at least 11% (Figure 2b), whereas the urban areas were projected to increase for all land use scenarios by at least 22% (Figure 2c). Thus, in lake basins where land use is affecting phytoplankton and cyanobacteria, the overall future land use effect predominantly reflects forest and urban cover changes because cropland changes are projected to be minor overall (Figure 2a).

The surface water temperature was projected to increase by $0.92°C (±0.14, SD)$ and $1.66°C (±0.22, SD)$ on average across lakes according to the RCP 2.6 and RCP 8.5 scenarios, respectively (Figure 2d). Water temperature, near-surface air temperature, light, and surface air pressure were projected to increase for the 29 lake basins in the future; whereas wind speed and relative humidity were projected to decrease in the future (Figure S2).

![Figure 1](image1.png) **Figure 1** Relative importance of the influence of land use, climate, seasonality, and lakes attributes on phytoplankton (a) and cyanobacteria abundances (b) from the 29 focal lake basins. The random forests (RFs) were fit using long-term (1971–2016) empirical data from 29 focal lake basins in the Northern Hemisphere.

![Figure 2](image2.png) **Figure 2** The projected (2044–2056) changes in percent upstream cropland (a), forest (b), and urban (c) areas relative to the baseline period (2006–2018) according to the three future land use scenarios of sustainability (blue), middle of the road (orange) and regional rivalry (red) in the watersheds of the 29 focal lake basins. (d) Projected (2044–2056) changes in lake surface water temperatures (SWTs) according to the RCP 2.6 (blue) and RCP 8.5 (red) scenario relative to the baseline period (2006–2018). The horizontal lines within the violin plots depict the 25%, 50%, and 75% quartiles.
3.3 | Relative importance of land use versus climate change

Despite overall increasing trends in average projected phytoplankton and cyanobacteria abundances, the projected responses compared to baseline conditions (2006–2018) varied substantially across the 29 focal lake basins (green circles in Figure 3a,b). The combined effect of land use and climate change was projected to increase phytoplankton abundance in 18 of the 29 lake basins (green circles in Figure 3a), while cyanobacteria abundance was projected to increase in 14 lake basins (green circles in Figure 3b).

The relative importance of land use and climate change driving lake phytoplankton varied strongly across lakes. Land use changes were the primary driving factor for future phytoplankton and cyanobacteria in 15 and 17 of the 29 focal lake basins (pie charts Figure 3a,b). The absolute ratio between land use- and climate-induced changes in phytoplankton was >1 for 14 of the 29 lake basins (t test, \( p < .05 \), pie charts in Figure 3). The magnitude of land use induced changes in cyanobacteria abundances was greater than the magnitude of climate-change induced changes (Figure 3, pie charts) except for lakes where land use changes were minimal such as in the Experimental Lakes Area in Canada (Lakes 114, 224, 239, 373, and 442). For cyanobacteria, the ratio between land use- and climate-induced changes in abundances was >1 for 18 of the 29 focal lake basins. Furthermore, meso- to oligotrophic lakes were projected to be more prone to small changes in watershed land use in their upstream areas compared to eutrophic lakes (Figure S3).

3.4 | Combined effects

The combined effect of all land use and climate scenarios was projected to increase phytoplankton on average by 111% (±24%, SD) for the 29 lake basins and 13% (±19%, SD) for the 1567 lakes. Cyanobacteria abundance was projected to increase on average by 215% (±88%, SD) on average for the 29 focal lake basins and by 23% (±63%, SD) for the remaining 1567 lakes. The phytoplankton were projected to increase in 1475 (94%) of these lake basins, while cyanobacteria were projected to increase in 1371 (87%) of the 1567 lake basins. The magnitude of overall changes was smaller for the 1567 lakes as compared to the 29 focal lake basins, because the mean change for the 1567 lakes was down-weighted by the large number of lakes (40%, 618 of 1567) in the remote Experimental Lakes Area in Canada, which was projected to be experiencing minimal changes in land use or climate.

The magnitude of projected changes in phytoplankton and cyanobacteria abundances was related to the average phytoplankton and cyanobacteria abundances during the 2006–2018 baseline period, that is, related to the trophic state of the lakes (Figure 4a,b). Among the 1567 lakes, the phytoplankton tended to decrease in phytoplankton-poor lakes (Figure 4c), whereas the cyanobacteria abundance tended to increase in cyanobacteria-poor lakes (Figure 4d).

The proportion of upstream urban and cropland areas was positively correlated with phytoplankton and cyanobacteria abundances, whereas forest showed a negative correlation with their abundances. These partial effects were relatively stronger for cyanobacteria as compared to phytoplankton (Figure S4).

Under all six possible combinations of land use and climate scenarios, the average phytoplankton and cyanobacteria abundances from the 1596 studied lakes (29 focal lake basins and the 1567 lakes) were projected to increase by 2050 (Figure 5). Differences between the worst-case land use and climate scenario effects and the least-case scenario were minor (Figure 5). However, the projected changes across scenarios varied less for phytoplankton abundance (−10–16%) as compared to cyanobacterial abundance (−18–30%) (Figure 5). The greatest changes in lake phytoplankton and cyanobacteria were projected by the land use regional rivalry scenario in combination with the two climate scenarios (Figure 5a,b).

4 | DISCUSSION

We projected phytoplankton and cyanobacteria abundances in 1567 lakes by using empirical relationships between weekly to monthly long-term biological sample data and the historical and future upstream land use and climate data. By the 21st mid-century, phytoplankton and cyanobacteria abundance is projected to increase for these 1567 lakes on average primarily due to increases in urban land use and temperature, and decreases in forest land use. Accounting for land use and climate changes in a combined way based on the extensive data allowed us to identify urbanization as the major driver of phytoplankton development in lakes located in urban areas, and climate as major driver in lakes located in remote areas where past and future land use changes were minimal. For approximately one-third of the studied lake basins, the combined effects of land use and climate change were important. Lake surface water temperature, urban cover, and forest cover were the most important drivers of algal abundance. However, lakes varied in the direction of their projected changes in phytoplankton.

On average, lakes are expected to have higher phytoplankton abundance due to increasing temperatures and decreasing forest areas replaced by increasing urban land use. Both historical changes and future phytoplankton projections suggest that the increased nutrient loading associated with the already known impacts of urbanization or deforestation as major drivers of eutrophication of lakes (Carpenter et al., 2011; Correll, 1998; Wiley et al., 2010). Increasing water temperature and altered mixing regimes directly and indirectly increase algal biomass (Adrian et al., 2009, 2016; De Senerpont Domis et al., 2013; Kraemer, Chandra, et al., 2017). The projected changes in land use and climate are thus expected to further increase lakes’ phytoplankton and cyanobacterial abundances that may lead to development of nuisance algal blooms, particularly increasing blooms of cyanobacteria (Carpenter et al., 2008; Schindler, 2012).
FIGURE 3  Average percent changes of (a) phytoplankton and (b) cyanobacteria abundances as an average of the projected (2044–2056) changes across all land use and climate scenarios relative to the baseline period (2006–2018) (green and blue circles). The size of circles represents the standard deviation to the mean percent changes (coefficient of variation, CV) in phytoplankton and cyanobacteria abundances projected for 2050. The size of the circles gets larger while the variability in phytoplankton and cyanobacteria responses across all land use and climate scenarios get higher. The pie charts depict the projected proportion of land use (gold) or climate (purple) change effects in altering future (a) phytoplankton and (b) cyanobacteria abundances. Please note that about one-third of the lake basins are predominantly affected by changes in land use (primarily lakes located in urban and agricultural areas), one-third predominantly affected by climate change (primarily lakes in remote areas), and for one-third of the lakes both changes in land use and climate are important. Full lake names and their abbreviations are given in Table S1.
Projected (2044–2056) percent changes of (a) phytoplankton and (b) cyanobacteria abundances in relation to their respective abundance during the baseline period (2006–2018) according to all land use and climate scenarios for all 1567 studied lake basins. Note that the strongest changes are projected for lakes with low phytoplankton and cyanobacteria abundances. Panels (c) and (d) show the direction of change (decreasing in blue, increasing in green) in the abundance of (c) phytoplankton and (d) cyanobacteria in relation to their respective abundance during the baseline period for the 1567 lake basins. Note that phytoplankton abundance decreases in algae-poor lakes and increases in phytoplankton-rich lakes. However, cyanobacteria increase in lakes with lower cyanobacteria abundances in the projected period (2044–2056).
While evidence for an increase in algae in a warming world with a growing human footprint is strong, we found substantial variation among lakes which we attribute to the magnitude of projected changes in land use and climate within each lake’s watershed (Figure 3). However, the influence of a specific variable on the projected change in phytoplankton and cyanobacterial abundance reflects both the magnitude of the variable’s projected changes and its importance in the model. For a predictor to have a strong influence on variation among lakes, the magnitude of projected change in that predictor must itself vary among lakes. Furthermore, that predictor must also have a relatively high importance in the model. For example, both cropland and urban areas were among the most important predictor variables in the RF model (Figure 1). However, only projected changes in urban areas varied from lake to lake and could therefore explain lake-to-lake variation. Alternatively, light is among the most important predictor variables in the RF model, but the magnitude of its projected changes and variation among lakes were minimal—thus it did not have a strong influence on lake-to-lake variation in projected phytoplankton and cyanobacteria abundance (see Figure S2c,d). Consequently, the projected phytoplankton and cyanobacterial conditions in lakes will depend heavily on the realized land use and climate changes, which will vary dramatically across geographical areas based on regional characteristics and the demand for available natural resources (McCullough et al., 2019; Watson et al., 1998).

Furthermore, phytoplankton and cyanobacteria responses to land use and climate changes vary according to lake trophic state with meso- to oligotrophic lakes being more prone to small changes in watershed land use in their upstream areas compared to eutrophic lakes (Figure S3; Carpenter et al., 2008; Rusak et al., 2018; Schindler, 2012). In meso- to oligotrophic lakes nutrients are limiting, thus any changes in nutrients loading may cause strong responses in phytoplankton and cyanobacteria. In contrast, nutrients may not be a limiting factor in eutrophic lakes, thus any changes in nutrient loading may have fewer consequences for phytoplankton and cyanobacteria. Climate change can strengthen these effects particularly in deep meso- to oligotrophic lakes through positive effects of warming on algae production—but counteract the effect as prolonged thermal stratification prevents mixing of nutrients from the hypolimnion into the photic zone over longer time periods (Kraemer et al., 2015). In shallow productive lakes, which turn anoxic during thermal stratification, warming acts like an additional eutrophication effect via an increase in internal nutrient loading released from anoxic sediments (Adrian et al., 2009; De Senerpont Domis et al., 2013; Jane et al., 2021; Wilhelm & Adrian, 2008). Wind and precipitation extremes causing an increase in external nutrient loading via run-off, resuspension of lake sediment, altered water column mixing may add to the effects of nutrient enrichment through land use changes particularly in meso- to oligotrophic lakes (Carpenter et al., 2008; Rusak et al., 2018; Schindler, 2012; Stockwell et al., 2020). Heat extremes may directly promote cyanobacteria, which are well adapted to high water temperatures and indirectly through an increase in internal nutrient loading during prolonged thermal stratification (Huber et al., 2012; Huismann et al., 2018). Yet another reason for the variation in responses across the studied lakes might be attributed to geographic location (e.g., latitude and elevation) and lake-specific characteristics and morphometry (e.g., lake area and depth) within and across regions (e.g., Richardson et al., 2017). For example, surface water temperature of deep lakes at high elevations (e.g., Alpine lakes) might change very slowly, but these lakes experience prolonged periods of thermal stratification (Kraemer et al., 2015) with repercussions for internal nutrient mixing.

The overall results for the 1567 lakes showed that combinations of land use and climate scenarios lead to larger increases in cyanobacteria abundance compared to phytoplankton. This can be attributed to the relatively large influence of land use (nutrient loading) and climate (temperatures) variables on cyanobacteria abundances. Another reason for the strong cyanobacteria responses might be described by the partial effects of upstream land use (urban, forest, and cropland areas), which were much stronger for cyanobacteria compared to phytoplankton (Figure S4). The differences across scenario combinations are also substantially greater for cyanobacteria.
than phytoplankton, consistent with previous findings (Bucak et al., 2018). The stronger cyanobacteria responses, compared to phytoplankton, may be explained by changes in nutrient loading that offset the effects of temperature on phytoplankton growth. The phytoplankton in this study included cyanobacteria taxa to be comparable with Chl-a in lakes, thus the increase in phytoplankton and cyanobacteria abundances can also be described by a distinct shift in phytoplankton taxa such as nitrogen-fixing cyanobacteria at the expense of diatoms (Markensten et al., 2010). Furthermore, the larger increase in temperature according to climate scenario RCP 8.5 compared to RCP 2.6 leads to a relatively long growing season for phytoplankton and cyanobacteria, thus leading to a stronger effect of RCP 8.5 on both groups.

Overall, the magnitude of responses induced by land use change was usually greater than those induced by climate change. Whereas lakes in remote areas (where changes in land use were minimal and predicted to remain low) were strongly affected by climate warming, lakes in watersheds with some urbanization were more affected by land use. Increasing external nutrient inputs due to strong increases in urban use and decreases in forest use increased projected phytoplankton and cyanobacteria abundances in our study. In other studies, however, significant reductions in negative impacts on water quality and biological communities associated with increase in forest use resulted in decreased algal productivity (Wiley et al., 2010). The magnitude of these responses reflects the potency of land use impacts on lake ecosystems. Land use change induced nutrient increase can drastically facilitate algal growth rates even under slight temperature increase (Litchman & Klausmeier, 2008). Furthermore, despite the controlling influence of climate variables such as water temperatures on lake organisms (e.g., Sommer & Lengfellner, 2008), community-based metrics such as cyanobacteria or phytoplankton abundances are more sensitive to land use change than climate change (Wiley et al., 2010). The higher sensitivity of community metrics to land use may stem from the enhanced nutrient loads via external sources (land use change) in comparison to internal sources (climate change; release from sediments) in, for example, eutrophic lakes. Thus, climate change impacts also act via an increase (increase in internal nutrient loads during prolonged thermal stratification in eutrophic lakes) or a decrease in nutrient availability (reduced upward mixing of nutrients in deep lakes during prolonged stratification periods).

Climate-change effects on phytoplankton and cyanobacteria abundances depended on lake average cyanobacteria or phytoplankton abundances during the baseline period (Figure 4). Warmer water temperatures were reported to cause phytoplankton-poor lakes to get poorer and phytoplankton-rich lakes to get richer (Kraemer et al., 2017)—the reason might be attributed to differences in the temperature effects on trophic interactions and resource availability across the lake phytoplankton abundance gradient (Kraemer, Mehner, et al., 2017). Land use and climate change, respectively, in conjunction with high nutrient inputs and warming temperatures, strongly favor the growth of phytoplankton which are less efficiently consumed by grazers (mostly cyanobacteria) (Kosten et al., 2012; Rigosi et al., 2014). Moreover, increasing grazing activity and phytoplankton consumption may reduce phytoplankton in phytoplankton-poor lakes (Kraemer, Chandra, et al., 2017; Kratina et al., 2012). However, the cyanobacteria were projected to increase in cyanobacteria-poor lakes, which might be explained by the fact that even low levels of increasing nutrient inputs facilitate cyanobacteria growth rates under slight temperature increase (Litchman & Klausmeier, 2008) and may hint at an initiation of cyanobacterial expansion in historically oligotrophic lakes (Sterner et al., 2020).

We examined the implications of various land use and climate change scenarios for lake algae. Comparison of the overall responses across different land use and climate change scenario combinations clearly illustrate (1) land use and climate changes are important factors that interactively promote or suppress lake algae, and (2) land use change often predominates over climate change in affecting lake algae and cyanobacteria. Therefore, land use management will continue to play a very important role in developing climate adaptation strategies (as suggested by Wiley et al., 2010). Furthermore, the strongly varying lake-to-lake responses of phytoplankton and cyanobacteria to land use and climate changes clearly points to the need for local management strategies. To guarantee sustainable development on land to prevent further eutrophication of lakes, local management measures that focus on retaining nutrients in agricultural and urban landscapes are likely to be most effective in preventing nutrient pollution from exacerbating ongoing changes to lake ecosystems from climate change.

ACKNOWLEDGMENTS
This study was funded through the “LimnoSCenES” project (Research on social-ecological scenarios for biodiversity and ecosystem service changes in north temperate freshwater ecosystems over the next half century; AD 91/22-1) supported by the network programming and funding research on biodiversity and ecosystem services across European countries and territories (BiodivERsA) and the Belmont Forum. RA was funded by the MANTEL-ITN project within the European Union’s Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie grant agreement no. 722518. Data collection for Acton Lake was supported by US National Science Foundation grants 9318452, 9726877, 0235755, 0743192, and 1255159. Windermere monitoring is currently supported by Natural Environment Research Council award number NE/R016429/1 as part of the UK-SCAPE programme delivering National Capability. Oneida Lake monitoring is supported by Cornell University and New York State Department of Environmental Conservation. Data for Lakes Annecy, Bourget and Geneva were collected as part of the long-term monitoring programs conducted by the Observatory of alpine Lakes (OLA, © SOERE OLA-IS, AnaEE-France, INRAE of Thonon-les-Bains) and co-funded by the INRAE and SILA, CISALB, CIPEL, respectively. Cheney Reservoir data collection was funded by the U.S. Geological Survey Cooperative Matching Funds Program and the City of Wichita, Kansas. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government. Compilation
of data sets was provided in part by the Global Evaluation of the Impacts of Storms on freshwater Habitat and structure of phytoplankton Assemblages (GEISHA) Working Group, supported by the U.S. Geological Survey (G17AC00044), the Centre de Synthèse et d’Analyse sur la Biodiversité (CESAB) of the Fondation pour la recherche sur la Biodiversité (FRB, www.fondationbiodiversite.fr), and the Vermont Water Resources and Lake Studies Center (Award G16AP00087). Cheney Reservoir data collection was funded by the U.S. Geological Survey Cooperative Matching Funds Program and the City of Wichita, Kansas. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government. Lake Müggelsee data collection was funded by the IGB long-term ecological research programme.

CONFLICT OF INTEREST
The authors have no conflict of interest to declare.

DATA AVAILABILITY STATEMENT
Data that support the findings of this study have been deposited in https://zenodo.org/record/5501851#.YT3xTy0ethF. USGS water-quality data are available through the National Water Information System at: http://doi.org/10.5066/F7P5SKJN. USGS phytoplankton data are available as a USGS data release at: https://doi.org/10.5066/F7ZG6QFX.

ORCID
Karan Kakouei https://orcid.org/0000-0001-8665-6841
Benjamin M. Kraemer https://orcid.org/0000-0002-3390-9005
Orlane Anneville https://orcid.org/0000-0003-1075-5661
Laurence Carvalho https://orcid.org/0000-0002-9768-9902
Heidrun Feuchtmayr https://orcid.org/0000-0002-2028-4843
Jennifer L. Graham https://orcid.org/0000-0002-6420-9335
Scott Higgins https://orcid.org/0000-0001-9427-7024
Francesco Pomati https://orcid.org/0000-0002-9746-3735
Lars G. Rudstam https://orcid.org/0000-0002-3732-6368
Jason D. Stockwell https://orcid.org/0000-0003-3393-6799
Stephen J. Thackeray https://orcid.org/0000-0003-3274-2706
Michael J. Vanni https://orcid.org/0000-0001-8690-9124
Rita Adrian https://orcid.org/0000-0002-6318-7189

REFERENCES


**SUPPORTING INFORMATION**

Additional supporting information may be found in the online version of the article at the publisher’s website.