

Geographic redistribution of microcystin hotspots in response to climate warming

Received: 19 November 2022

Accepted: 4 September 2023

Published online: 12 October 2023

 Check for updates

Julian Merder¹✉, Ted Harris², Gang Zhao^{1,4}, Dimitrios M. Stasinopoulos³, Robert A. Rigby³ & Anna M. Michalak¹✉

High concentrations of cyanobacterial toxins such as microcystin represent a global challenge to water quality in lakes, threatening health, economies and ecosystem stability. Lakes are sentinels of climate change but how warming will affect microcystin concentrations is still unclear. Here we examine how warming impacts the probability of exceeding microcystin water quality thresholds across 2,804 lakes in the United States and show how future warming will alter these probabilities. We find that higher temperatures consistently increase the likelihood of microcystin occurrence but that the probability of microcystin concentrations above water quality thresholds is highest for water temperatures between 20 and 25 °C. Regions with temperatures that promote microcystin will shift to higher latitudes in the coming decades, leading to relative changes in exceedance probabilities of more than 50% in many basins of the United States. High nitrogen concentrations amplify the impact of rising temperatures, calling for increased awareness of a substantial hazard to ecosystems and human health under global warming.

Harmful algal blooms and their associated toxins pose a severe threat to health and water quality^{1,2}. Microcystins, produced by cyanobacteria commonly known as blue-green algae, are one of the most frequently detected classes of such toxins globally³, with several hundred microcystin congeners currently known⁴. As a potent hepatotoxin, microcystin affects liver function by inhibiting protein phosphatases⁵ and acute concentrations can lead to fatalities in wild and domestic animals and, in rare cases, humans^{6,7}. Microcystin can also impact animal and human health through chronic exposure^{8,9}, including being classified as a potential carcinogen¹⁰. As an example of real-world impacts, around 500,000 people living near Lake Erie were instructed in 2014 to not drink tap water because microcystin concentrations exceeded safe limits in finished drinking water¹.

Climate change is one of the greatest challenges to water quality and aquatic ecosystems^{11,12}. Lakes, which hold 87% of Earth's liquid surface freshwater¹³, are sentinels of climate change¹⁴ with lake summer surface temperatures expected to increase by 0.34 °C per decade on average¹⁵. This trend will have severe consequences for lakes, with

impacts on critical abiotic and biotic processes such as mixing regimes, evaporation, lake ice phenology and the growth rates and composition of freshwater taxa^{11,12,14}.

How temperature affects total microcystin concentrations within lakes, however, is still unclear. Generally, cyanobacteria have been identified as big winners of climate warming and are likely to increase globally in abundance and dominance, outcompeting other species due to boosted growth rates and intensified water stratification^{16–20}. Larger blooms, however, do not necessarily contain higher microcystin concentrations, as not all taxa can produce this class of toxins²¹ and high genetic diversity leads to varying toxin-producing capabilities even within a species²². Warming has also been found to promote toxic strains over non-toxic strains²³ but toxin cell quotas have been reported to decrease with increasing temperatures^{24,25}. While laboratory studies have reported temperature optima for elevated microcystin concentrations²⁶, field studies widely disagree on the direction of the temperature effect on microcystin concentrations, reporting negative, positive or negligible^{27–29} relationships in lakes. This complicates water quality

¹Department of Global Ecology, Carnegie Institution for Science, Stanford, CA, USA. ²Kansas Biological Survey and Center of Ecological Research, University of Kansas, Lawrence, KS, USA. ³School of Computing and Mathematical Sciences, University of Greenwich, Greenwich, UK. ⁴Present address: Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China. ✉e-mail: jmerder@carnegiescience.edu; michalak@carnegiescience.edu

management because rising temperatures may counteract the effect of nutrient management strategies³⁰.

Here, we use 3,027 measurements from 2,804 lakes across the United States that were sampled in 2007, 2012 and 2017 as part of the National Lakes Assessment³¹ (NLA) to determine how temperature impacts microcystin occurrence (defined here as concentrations above the 0.1 $\mu\text{g l}^{-1}$ detection limit of enzyme-linked immunosorbent assays (ELISA)^{32,33} used by the NLA) and concentrations across large geographic regions. We use a logistic model to represent occurrence and a log-normal model to represent concentrations above the analytical detection limit while controlling for pH, Secchi depth, the natural-log transformations of each of total nitrogen (TN), total phosphorus (TP), chlorophyll *a* (Chl-*a*), dissolved organic carbon (DOC) and lake depth (*D*) and area (*A*), all of which have been found to impact or correlate with microcystin concentrations in previous studies^{22,27,28,34,35}. The two models are then combined into a zero-adjusted model²⁷ that represents microcystin both below and above the detection limit (Methods).

The models are fitted in the framework of generalized additive models for location, scale and shape³⁶ (GAMLSS) which allows us to identify potential nonlinear relationships and to model the full probability distribution of microcystin concentrations, making it possible to quantify the probability of exceeding any defined water quality threshold under specific environmental conditions (Methods).

We then use the combined model to assess the impact of future warming on the geographic distribution of areas with elevated probabilities of microcystin concentrations above critical water quality thresholds across the continental United States. To do so, we use projections from 15 generalized circulations models (GCMs) (Extended Data Table 1) participating in the coupled model intercomparison project phase 6 (CMIP6) (ref. 37) under the 'middle-of-the-road' shared socioeconomic pathway scenario SSP 2-4.5 (ref. 38) (Methods).

Exceedance probabilities of water quality thresholds

The World Health Organization (WHO) provisional guideline values for drinking water³⁹ list concentrations of 0.3 and 1 $\mu\text{g l}^{-1}$ for children and adults, respectively. The US Environmental Protection Agency (EPA) has set a microcystin threshold of 8 $\mu\text{g l}^{-1}$ as the water quality criterion for recreational waters protective of human health while either swimming or taking part in recreational activities on the water⁴⁰. During the three sampling years of the NLA, 18.0% and 9.9% of lakes investigated in this study had concentrations above the drinking water guideline for children and adults, respectively, and 1.3% had concentrations above the recreational water quality criterion (Fig. 1). The model developed here (Methods) was applied under the observed environmental conditions at each lake at the time of sampling and predicts 16.8–20.0% (99% interval; Methods) of lakes would have concentrations above the children's drinking water guideline. The predicted intervals are 7.6–10.1% for the adult drinking water guideline and 1.1–2.3% for the recreational water criterion. The model is therefore highly effective at representing the probability of exceeding critical water quality standards under the range of conditions observed during the NLA.

For individual lakes, the predicted probabilities range from 0 (Waldo Lake, Oregon) for exceeding the recreational water criterion to 0.91 for exceeding the children's drinking water guideline (Roundup Lake, Nebraska) under the environmental conditions observed at the time of sampling. In regions such as the upper Midwest corn belt, the probability of exceeding the 0.3 $\mu\text{g l}^{-1}$ water quality guideline for drinking water for children can be >0.50 on average across entire basins (Fig. 2a; five basins), >0.25 for exceeding the 1.0 $\mu\text{g l}^{-1}$ water quality guideline for adults (Fig. 2c; 14 basins) and >0.10 for the 8 $\mu\text{g l}^{-1}$ recreational water criterion (Fig. 2e; two basins). These results indicate that exceeding water quality thresholds in those areas is already common rather than being the exception. The probabilities show a

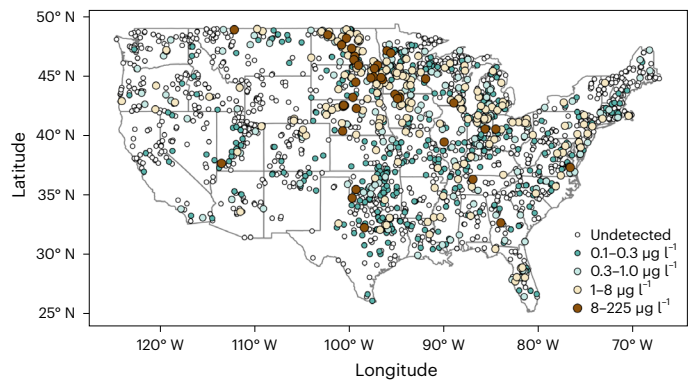


Fig. 1 | Microcystin concentrations across the United States. Observed microcystin concentrations across 3,027 sampling points (2,804 lakes) over three survey years (2007, 2012 and 2017).

strong geographic signature (Fig. 2) based on temperature, degree of eutrophication and other environmental conditions (Extended Data Fig. 1), which we discuss next.

Drivers of microcystin occurrence and exceedance probabilities

We find that temperature (*T*), TN, Chl-*a* and pH help explain both the occurrence (logistic model) of microcystin and median concentrations above the detection limit (log-normal model). In addition, lake area (*A*) and DOC have an effect on microcystin occurrence, thus indirectly also impacting microcystin concentration in the combined model, while lake depth (*D*) affects concentration above the detection limit. The partial effects of these covariates are shown in Fig. 3a–f for occurrence, in Fig. 3g–l for the median and spread (Methods) of the concentrations above the detection limit and in Fig. 4 for the probability of microcystin concentrations exceeding a specific threshold. Neither log TP, nor its interaction with log TN, showed a significant relationship to microcystin. Also, Secchi depth did not provide significant additional explanatory power.

Temperature

We find that while the probability of microcystin occurrence increases monotonically with temperature (Fig. 3a), the median concentration above the detection limit peaks at 22 °C (Fig. 3g). Consequently, the probability of exceeding specific concentration thresholds is also highest for temperatures ranging from 20 to 25 °C (Fig. 4). This contrasts with previous studies based on a more limited set of observations that had found a negative or negligible effect of temperature^{27,28}. Laboratory or single lake studies, however, have also shown an optimum temperature in the range of 20–25 °C for elevated microcystin concentrations^{26,41}, supporting our findings.

The more flexible nonlinear approach applied here thus made it possible to observe the presence of a temperature optimum at the field scale across thousands of lakes. Because both the logistic and log-normal models include proxies for biomass (Chl-*a*), the higher probability of microcystin occurrence at higher temperatures is probably attributable to an increase in the dominance and relative abundance of toxic strains of cyanobacteria²³, resulting from increased growth rates and water column stratification at higher temperatures.

The decreasing exceedance probabilities above 22 °C are consistent with studies reporting a reduction in toxin quota at higher temperatures and a decoupling between optimal growth rates and highest toxin concentration^{24,26}. Physiological reasons for this pattern are not fully understood^{24,25}. Another plausible reason for decreasing concentrations when temperature exceeds 22 °C is an increase in microcystin removal by bacteria that are capable of degrading microcystin⁴².

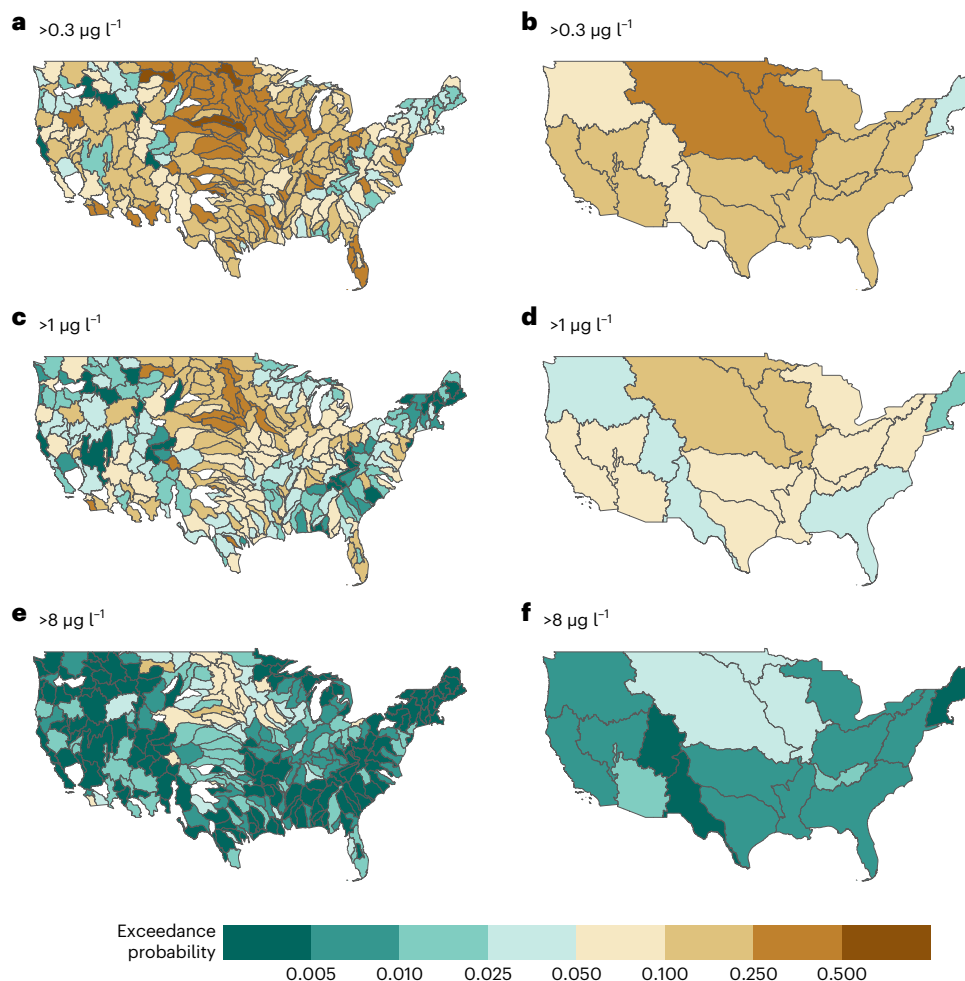


Fig. 2 | The upper Great Plains and the upper Midwest corn belt show the highest probabilities of exceeding microcystin thresholds. The exceedance probabilities for each lake are calculated on the basis of the environmental conditions measured at the time of microcystin sampling. **a–f**, Exceedance

probabilities are then averaged across lakes on the HUC6 basin scale (**a,c,e**) and HUC2 regional scale (**b,d,f**) for the thresholds $0.3 \mu\text{g l}^{-1}$ (**a,b**), $1 \mu\text{g l}^{-1}$ (**c,d**) and $8 \mu\text{g l}^{-1}$ (**e,f**).

Other environmental drivers

As expected, the model developed here confirms that in-lake nitrogen has a clear positive effect on microcystin occurrence and exceedance probabilities (Figs. 3 and 4). This is in line with previously observed results indicating that TN is a dominant driver of microcystin^{27,28,34}. This pattern may reflect that cells deal with excess N by shunting it out into N-rich metabolites^{26,43} such as microcystin. In addition, however, we also identified a direct impact of TN on the variability in microcystin concentrations above the detection limit (Fig. 3I), suggesting that the uncertainty in microcystin concentration increases with TN, that is, that the likelihood of microcystin concentration extremes scales directly with in-lake TN.

Consistent with earlier field and laboratory studies^{27,41}, the probability of exceeding microcystin thresholds grows with increasing Chl-a and pH. Intense blooms can lead to depleted carbon dioxide concentrations, thereby increasing pH. Studies also suggest that depleted carbon dioxide concentrations favour cyanobacteria over other taxa^{25,44}. For DOC, earlier studies report both negative and positive effects on microcystin occurrence or concentration^{27,34,45}. Here, we find that DOC has a positive effect on microcystin occurrence, which increases the probability of microcystin concentrations exceeding water quality thresholds. The effect of DOC on microcystin may be the result of light limitation⁴⁶ or the production of reactive oxygen species

formed under ultraviolet radiation, which in turn has been shown to spur toxin production³⁵. We also find an effect of lake area and depth on the probability of exceeding microcystin thresholds. Both lake area and lake depth increase the likelihood and strength of stratification in summer months⁴⁷, which has been observed to promote cyanobacterial dominance over eukaryotic competitors⁴⁸. Another way in which lake morphology may impact microcystin concentrations is through the coupling between lake water and lake sediments, where microcystin degradation has been found to take place⁴⁹.

Compound effects in the presence of warming

We find that the impact of rising temperatures is greater under high TN concentrations (Fig. 5a) and the sensitivity to changes in TN concentrations is highest for temperatures between 20 and 25 °C (Fig. 5b). This suggests that the impact of warming can be supercharged by eutrophication. At the time of the NLA, regions with high TN concentrations (Extended Data Fig. 1a) did not coincide with temperature regions associated with the highest risk for elevated microcystin concentrations (Extended Data Fig. 1g) but this could change under future conditions. That the whole may be greater than the sum of its parts is in line with the ‘allied attack’ concept⁵⁰ suggesting a potential reinforcing effect between eutrophication and rising temperatures and was previously observed to boost microcystin in a pond water experiment¹⁶.

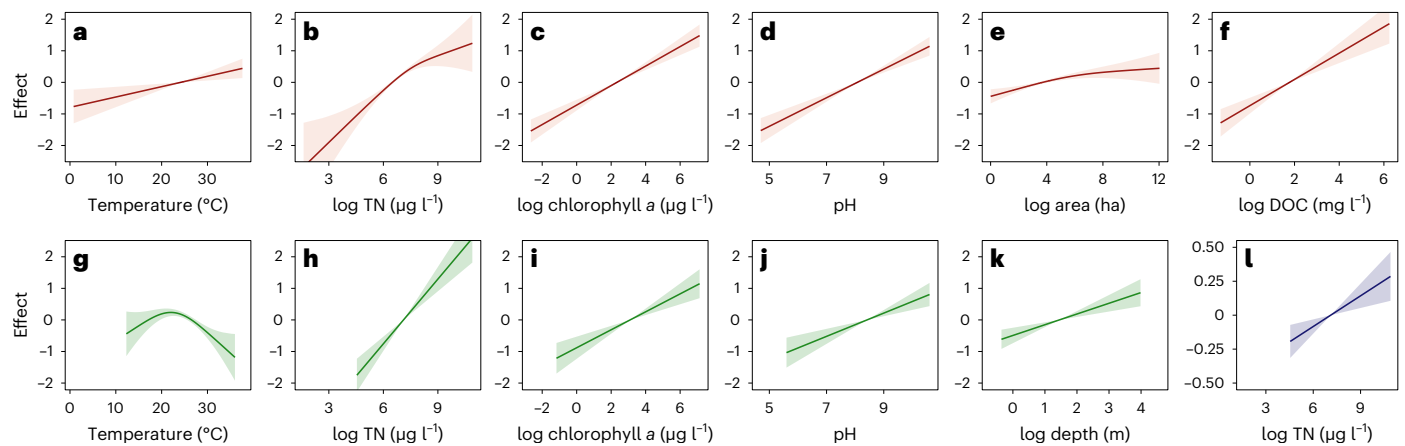


Fig. 3 | Partial effects of environmental parameters. **a–f**, Partial effects of the environmental parameters on the log odds of microcystin being detected (that is, microcystin occurrence) for the logistic model. **g–l**, Partial effects of the environmental parameters on concentrations above the detection limit for the log-normal model; **g–k**, show the partial effect on $\log(\mu)$, where μ is the predicted median of the microcystin concentration distribution, while **l** shows the effect on $\log(\sigma)$, where σ is the predicted standard deviation of the log microcystin

concentration above the detection limit; Methods. The ranges on the horizontal axes represent the ranges observed in the sampled data. The range for the log-normal model is smaller than for the logistic model because it only includes the range of observed values for those samples where the microcystin concentration was above the detection limit. The shaded areas represent the 95% confidence intervals.

Here, we observe and quantify this effect across thousands of lakes (Fig. 5). Interestingly, such a compounding effect was not observed for cyanobacterial biovolume¹⁹.

Geographic redistribution under rising temperatures

We find that the probability of exceeding water quality thresholds will increase under SSP 2-4.5, applied to 15 climate models participating in the CMIP6 intercomparison (Methods). Taking the water quality guideline for children as an example, the relative increase in exceedance probability will be >50% for 14 basins by the mid-century and 28 basins by the late-century relative to historical conditions (Fig. 6a,c,e). We project that about a quarter of basins will see increases of >25% for the mid- (66 basins) and late-century (78 basins). Even at the Hydrologic Unit Code (HUC) regional scale HUC2 (Fig. 6b,d,f), three regions will see increases of >25% by mid-century and two regions will see increases of >50% by late-century.

Geographically, the regions with the greatest relative increase in exceedance probabilities are located in the north of the United States, where summer temperatures are projected to be close to the optimum for high microcystin concentrations (Fig. 7). For some regions, such as the Great Lakes, the exceedance probabilities are already high (Fig. 2) based on relatively high TN concentrations. But even for northern regions that currently have relatively low exceedance probabilities (Fig. 2), a relative change in exceedance probabilities of up to 50% means that those regions are much more likely to experience hazardous microcystin concentrations in the future. Interestingly, even at present the relative increase in exceedance probabilities is >10% higher than under historical conditions for more than a quarter of basins (Fig. 6a), indicating that the impact of regional shifts in temperature is already being felt.

Conversely, only one basin is expected to see a relative reduction in exceedance probabilities of >25% by the late-century, while less than a quarter of basins will experience a reduction of >10% by the mid-century and a third by the late-century (Fig. 6d,e,f). These areas coincide with regions where temperatures are expected to rise above the range associated with the highest probability of exceeding microcystin thresholds.

Similar patterns emerge for other microcystin thresholds, that is, 1 and 8 $\mu\text{g l}^{-1}$ (Extended Data Figs. 2 and 3). Overall, the divergent responses in the north and south result from the northward migration of the region with summertime temperatures closest to the 22 °C optimum (Fig. 7). However, the divergent response is not symmetric

and there are more regions experiencing increases than decreases (Extended Data Fig. 4). Under increasing temperatures, microcystin levels are also expected to reach detectable concentrations more often (Extended Data Fig. 5) with average relative increases in occurrence of 9.5% by the late-century across HUC2 regions.

Given that microcystin occurrence is defined here on the basis of the detection limit of 0.1 $\mu\text{g l}^{-1}$, it is likely that smaller concentrations will become even more frequent. What impact an increasingly chronic microcystin exposure even under low concentrations will have on health and ecosystem functions is not well understood. In a recent study, it was shown that chronic exposure with an estimated daily intake of even only 0.15–0.27 μg (MC-LReq per day) led to detectable microcystin concentrations in human blood sera and was associated with signs of renal impairments⁹. Chronic exposure is especially dangerous for people with existing liver disorders such as non-alcoholic fatty liver disease⁸. Microcystin has also been reported to be transferred to higher trophic levels through the food chain⁵¹, to lead to growth inhibition of edible plants in bioassays and to be detected on plants grown for human consumption after crop spray irrigation^{52,53}.

Compound events⁵⁴, for example, events of synchronously high in-lake TN concentrations together with high-risk temperatures, will further increase the likelihood of microcystin extreme concentrations in the future (Fig. 5). High concentrations may occur under changing land use patterns or from agricultural runoff after weather events such as heavy rainfall. Hotter spring days or extended periods of temperatures around 22 °C in autumn may also increase the frequency and duration of high-concentration events. Although more data will be needed to assess seasonal differences in microcystin occurrence and concentration, summer-based monitoring programmes will need to extend to the spring and autumn months.

Conversely, preventing excess nutrients from reaching water bodies and reducing eutrophication is key to counteracting the effects of rising temperatures. Efforts to reduce inland water eutrophication over the past decades have focused mainly on phosphorus, due to the assumption that atmospheric N_2 -fixing cyanobacteria will buffer nitrogen levels in the water column⁵⁵. However, N_2 -fixing is not sufficient to offset nitrogen loss due to processes such as denitrification⁵⁶ and microcystin-producing taxa such as the genus *Microcystis* require external sources of nitrogen⁵⁷ from runoff or atmospheric deposition. In some systems, nitrogen loading has been shown to selectively promote

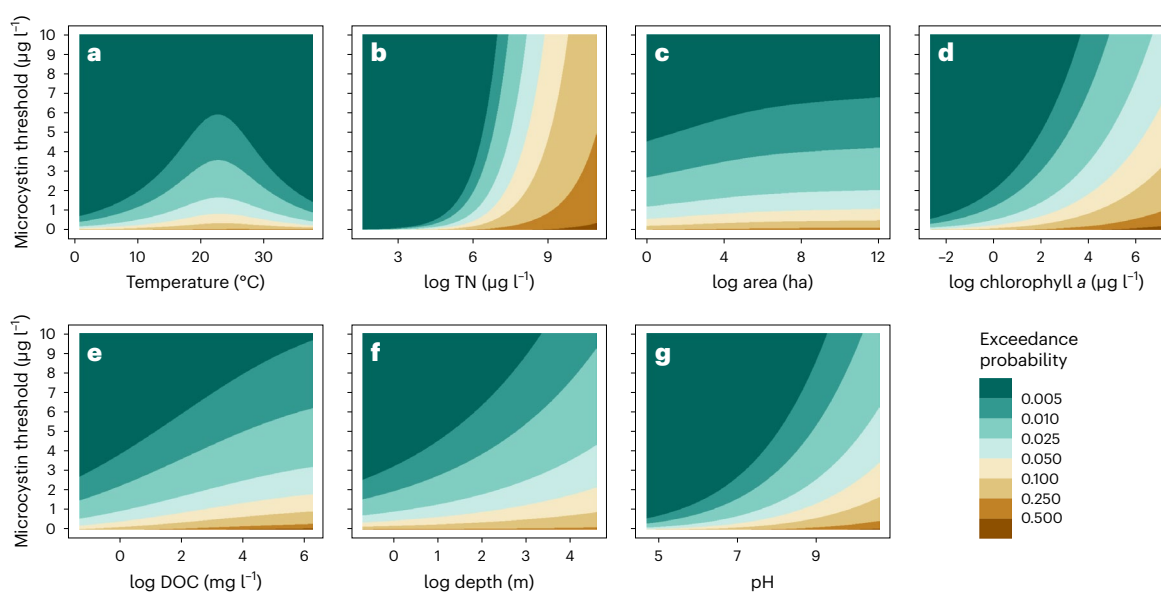


Fig. 4 | Exceedance probabilities under varying temperature, log TN, log area, log chlorophyll *a*, log DOC, log depth and pH. For easier visual interpretation the plots are grouped into selected exceedance probability bins, although exceedance probabilities are continuous. **a–g.** For a given microcystin threshold on the vertical axis the colours represent the exceedance probability as

a function of temperature (**a**), TN (**b**), area (**c**), chlorophyll *a* (**d**), DOC (**e**), depth (**f**) and pH (**g**). In each panel the other (log-transformed) environmental variables are held at their median (pH of 8.27, log chlorophyll *a* of 2.23, log TN of 6.53, temperature of 24.4 °C, log DOC of 1.77, log depth of 1.57 and log area of 3.55).

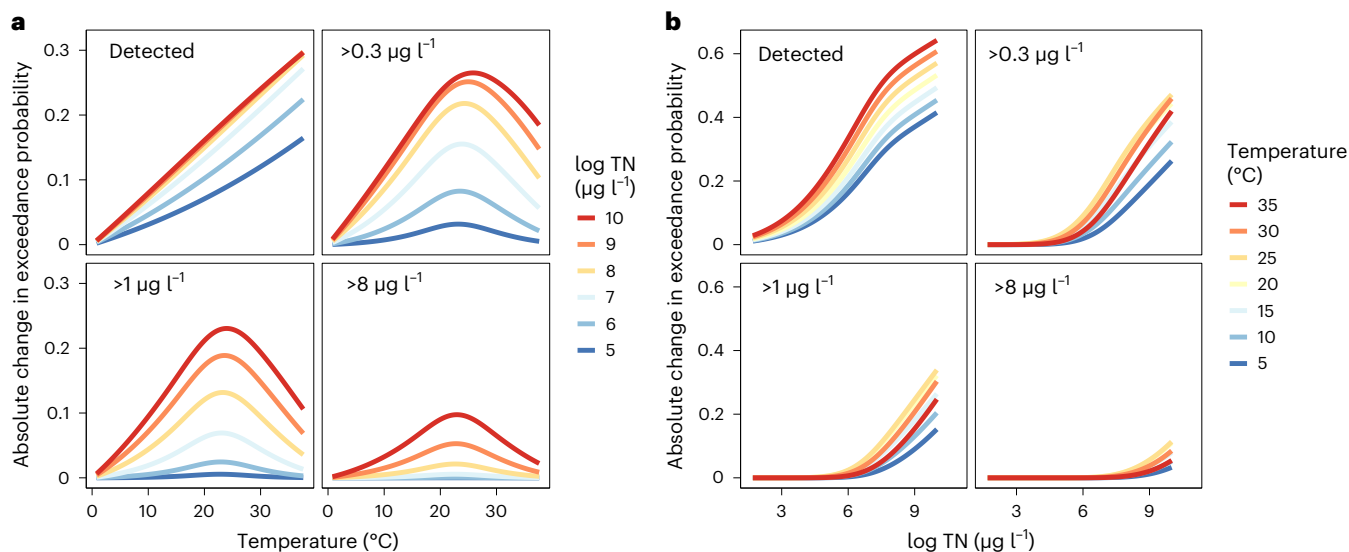


Fig. 5 | High in-lake TN concentrations amplify the effect of warming on the probability of exceeding microcystin water quality guidelines. **a.** The absolute change in exceedance probabilities for microcystin being above the detection limit (occurrence) and other selected thresholds when changing temperatures under different in-lake nitrogen concentrations. All variables

(other than temperature and log TN) are kept at their respective medians. **b.** The absolute change in exceedance probabilities when changing TN concentrations under different temperature conditions. All variables (other than temperature and TN concentration) are kept at their respective medians.

the abundance of toxic *Microcystis* strains, while orthophosphate has not⁵⁸. Therefore, the results presented here add to the increasing evidence that an exclusive focus on phosphorus reductions alone will not be sufficient to mitigate dangers from nitrogen-rich cyanotoxins²⁶.

Additional complexities and broader considerations

The ELISA used here (Methods) represent a congener-independent, robust and highly sensitive assessment of total microcystin concentrations^{59,60}. Its main advantage relative to methods such as liquid

chromatography-mass spectrometry is that it quantifies the total concentration across a very high fraction (>80%) of microcystin congeners⁶⁰. A loss of microcystins during preparation and storage can in principle result in an underestimation of the true microcystin concentration⁶¹ with ELISA, while microcystin conjugates and byproducts could lead to overestimations of the toxicity of a sample^{62,63}. Also, variable cross-reactivity of congeners can lead to either under- or overestimations of concentrations⁶³. The uncertainty resulting from these factors, however, would be highly unlikely to change the primary conclusions of the current study (Methods).

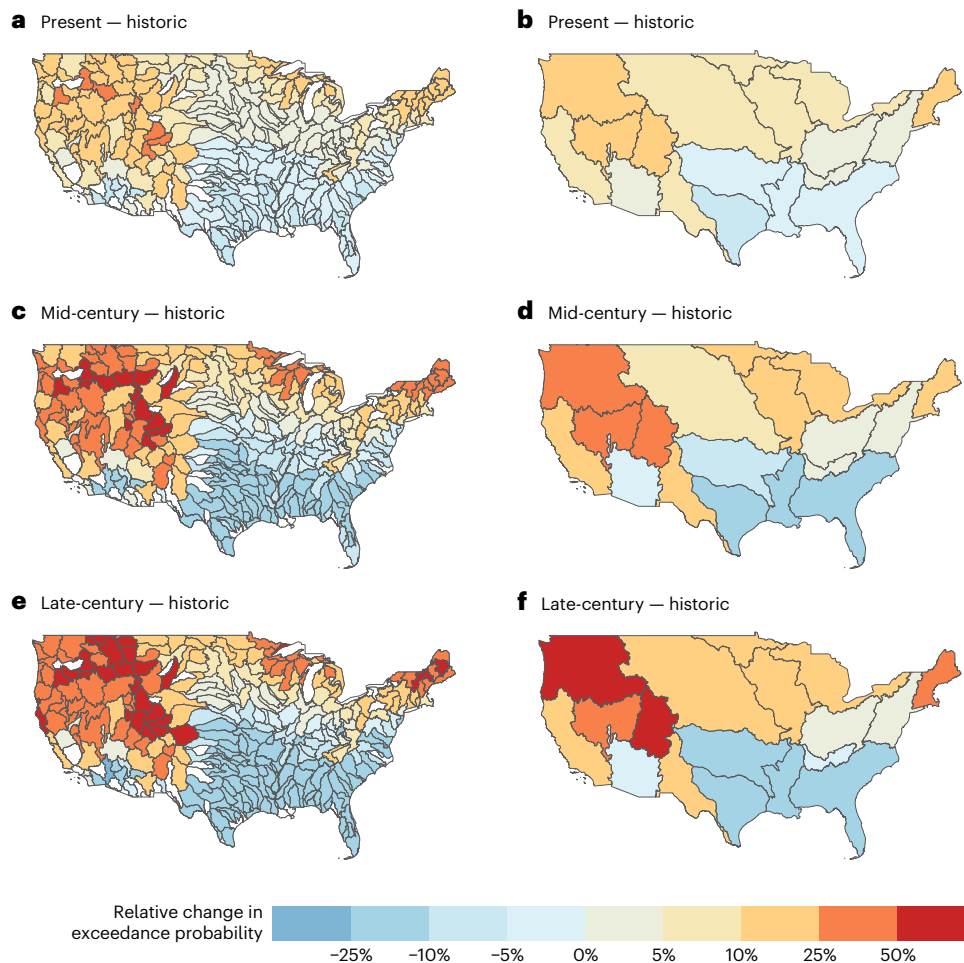


Fig. 6 | Warming will change the exceedance probability of the children's drinking water guideline ($0.3 \mu\text{g l}^{-1}$), with relative increases exceeding 50% in some areas. a–f, The relative increase in risk due to warming relative to the historic period (1950–1979), averaged across lakes within individual HUC6

basins (a,c,e) and HUC2 regions (b,d,f), for present (1990–2019) (a,b), projected mid-century (2030–2059) (c,d) and projected late-century (2070–2099) (e,f) summertime temperatures. All other variables are held constant at their values at the time of sampling. White areas represent basins with no sampled lakes.

Future large-scale monitoring using congener-specific approaches such as liquid chromatography-mass spectrometry could be used to answer additional questions about the impact of warming on specific strains and congeners, such as a shift in their diversity and relative abundance⁶⁴. Routine monitoring of all known congeners is almost impossible, however, because new microcystin congeners continue to be discovered. Low levels of several congeners that individually fall under the detection limit could also lead to an underestimation of the total microcystin concentration. Therefore, monitoring using congener-specific methods could supplement but not replace assessments based on ELISA.

Additional factors such as the local cyanobacterial strain composition and genetic diversity, grazing by zooplankton, viral lysis or microbial interactions can also influence microcystin concentrations^{65,66} but were not considered here because such data are currently non-existent for the United States across large scales. The identified effects of temperature therefore represent the net effect of temperature across various biotic and physiological responses. While strain evolution could in principle impact the future trends explored here, an earlier study found that increased eutrophication and temperature did not favour strain evolution towards more toxin-producing strains in one of the most important microcystin producers *Microcystis aeruginosa*⁶⁷. Future work exploring biotic factors could further deepen understanding of the processes that control microcystin and bloom dynamics.

In addition, future changes in temperature are also likely to affect other drivers included in the model and the future projections presented here do not account for correlations among changes to those variables (Extended Data Fig. 6). Given that temperature has been found to increase cyanobacterial biomass either directly or indirectly^{17,68}, increases in temperature may also increase other predictors in the model, such as Chl-a. As such, microcystin concentrations would be expected to increase even more than presented here, making our conclusions a conservative estimate of future impact. The long-term coupling of environmental variables, however, may impact cyanobacterial taxa differently^{19,69}, which makes future microcystin projection much more complex under climate change and highlights the need for more research focusing on the interplay of environmental variables.

Lastly, although this study focuses on lakes in the United States, microcystin is a global hazard⁶⁵ and the results shown here indicate that microcystin occurrence is expected to increase with rising temperatures. Regions where summertime temperatures are near 22°C , promoting exceedance of microcystin thresholds, will shift poleward not just for the United States but globally (Extended Data Fig. 7). Potential seasonal shifts in periods with elevated microcystin concentrations can also be expected. The threat to water quality posed by microcystin will therefore grow globally as warming continues, especially in eutrophic systems.

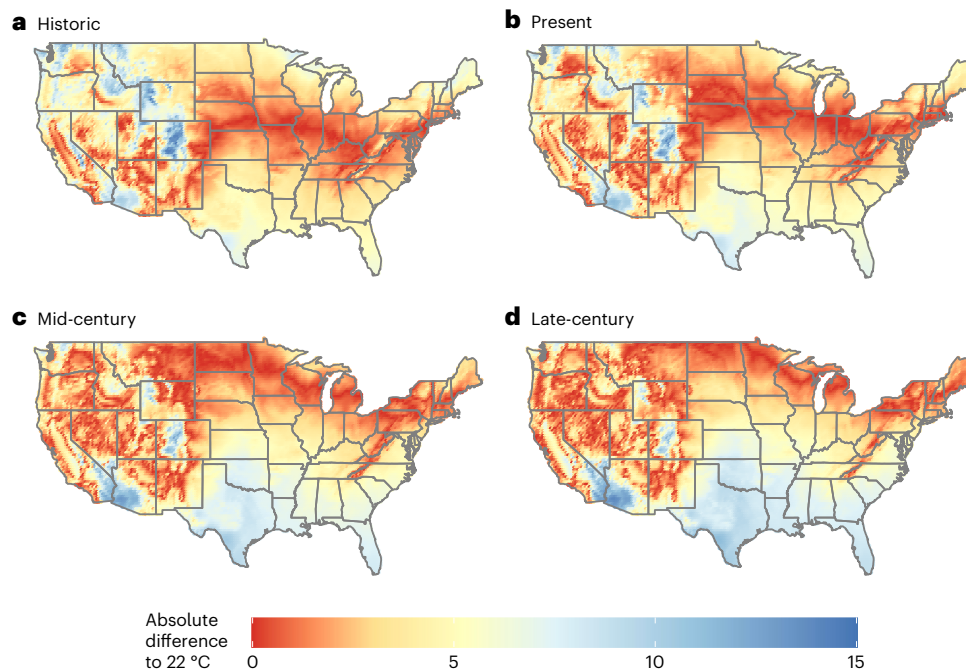


Fig. 7 | Areas with summertime temperatures associated with the highest risk of elevated microcystin concentrations are moving northward. a–d, The absolute difference of summertime air temperatures to 22 °C, the approximate

temperature with the highest risk, for the historic period (1950–1979) (a), present (1990–2019) (b), projected mid-century (2030–2059) (c) and projected late-century (2070–2099) (d).

Methods

Data

We analysed lakes across the continental United States that were part of the EPA National Lake Assessment (NLA) surveys conducted in 2007³¹, 2012⁷⁰ and 2017⁷¹ with microcystin concentrations detected by ELISA.

ELISA is an antibody-based analytical method widely applied in water treatment for the screening of toxic cyanobacterial metabolites^{32,33,59,60}. ELISA does not target individual microcystin congeners but rather measures the total concentration of microcystins (and nodularins) in a sample^{33,59,60} that share very similar toxicity and structure and are thus combined under the general term microcystins here. The ELISA used in the NLA surveys is calibrated against the microcystin-LR congener⁵⁹, which also forms the basis of the WHO drinking water guidelines³⁹. No statistically significant difference between concentrations could be identified when comparing ELISA to liquid chromatography/tandem mass spectrometry under various cell lysis techniques³³. The EPA uses freeze/thaw cycles to lyse cyanobacterial cells and extract toxins; as such this method not only quantifies dissolved toxins but also microcystin within cells⁵⁹. ELISA is the recommended method of the EPA to quantify total microcystins (and nodularins)⁵⁹ and as such many water monitoring efforts and drinking water supplies use ELISA kits to measure microcystins within their systems. Therefore, ELISA and the predicted microcystin concentrations derived from those measurements are in line with end users.

Lakes included in the NLA were selected by a stratified probabilistic sampling design to be representative of the >50,000 lakes in that project. Previous studies have analysed the 2007 and 2012 surveys^{27,28,34,45}, while our study incorporates the 2017 data.

From these data, we selected the environmental variables Chl-*a* ($\mu\text{g l}^{-1}$), surface water temperature *T* (°C), TN ($\mu\text{g l}^{-1}$) and TP ($\mu\text{g l}^{-1}$), DOC (mg l^{-1}), pH, Secchi depth (m), lake depth *D* (m) and area *A* (ha) for consideration for inclusion in the model. A very detailed description of data sampling can be found at EPA and within NLA references^{31,70,71}.

Although cyanobacterial abundance is also measured as part of this effort, we used Chl-*a* as a proxy for biomass because Chl-*a* is more

commonly measured and makes the model applicable to a wider range of water bodies beyond those examined in this study. Chl-*a*, TN, TP, DOC, *A* and *D* were all natural log-transformed. In total, 2,804 individual lakes had measurements of all the environmental variables considered here with paired microcystin measurements. As some lakes were sampled multiple times, the total number of observations was 3,027. In 1,092 of those observations microcystin was above the detection limit.

Statistical model

We used GAMLSS implemented via the R package `gamlss`^{36,72,73}. This is a distribution-based approach to nonlinear regression where not only the mean but all parameters of the conditional distribution, for example, the variance of the response variable, are related to environmental variables. In the easiest case, assuming the conditional distribution of the response is Gaussian, both the mean and the standard deviation are functions of covariates. Further, as all parameters of the distribution are modelled, the probability density function can be calculated for any combination of covariates. With this information, every quantile or probability to cross a chosen response value (microcystin threshold) can be calculated.

There is a detection limit of $0.1 \mu\text{g l}^{-1}$ for microcystin in these surveys, such that the observed microcystin concentration (*MC*) is not completely continuous. Because of that, we used a zero-adjusted model for $M = MC - 0.1$ that combines a logistic model (occurrence model) of microcystin detection or non-detection, with a continuous model to represent microcystin concentrations above the detection limit, that is, to model $M > 0$. For the continuous model, a log-normal distribution (LOGNO2) was selected as the optimal conditional distribution for $M > 0$ after extensive testing of >50 distributions currently available in GAMLSS⁷⁴ based on fit and parsimony via the Bayesian information criterion⁷⁵ (BIC). The log-normal distribution has two parameters, μ (the median of *M*) and σ (the standard deviation of $\log M$). The log-normal probability density function of $M > 0$ is defined as:

$$f_{M>0}(m) = \frac{1}{(2\pi)^{1/2} \sigma m} \exp\left\{-\frac{1}{2\sigma^2} [\ln(m) - \ln(\mu)]^2\right\} \quad (1)$$

where m represents any specific value of the random variable M . It is interesting that the log-normal distribution, which has a compounding effect between predictor variables (effects combine multiplicatively), performed best for microcystin concentrations out of the many distributions available in GAMLSS including, for example, a zero truncated t -distribution with identity link that does not lead to a compounding effect. This indicates that compounding effects between environmental variables are likely in predicting microcystin concentrations above the detection limit.

M was modelled using a zero-adjusted log-normal distribution, with a mixed (discrete-continuous) probability function. Hence, the mixed (discrete-continuous) probability function for microcystin MC is:

$$f_{MC(mc)} = \begin{cases} p_0 & \text{if } mc = 0.1 \\ \frac{(1-p_0)}{(2\pi)^{1/2}\sigma(mc-0.1)} \exp\left\{-\frac{1}{2\sigma^2}[\ln(mc-0.1) - \ln(\mu)]^2\right\} & \text{if } mc > 0.1 \end{cases} \quad (2)$$

where p_0 is the probability that microcystin is below the detection limit from the logistic model and where mc represents any specific value of the random variable MC.

The selection of environmental covariates used to estimate μ , σ of the log-normal model, and p_0 of the logistic model was based on minimizing the BIC in a stepwise selection process. Here, each environmental variable can be part of the distribution parameter model either as a penalized P-spline⁷² or a linear term, or not selected. BIC was also used locally for selecting the amount of smoothing, that is, the effective degrees of freedom used for smoothing, in the P-spline. The selection of environmental variables for the models was performed using the stepAIC() and stepAICAll.A() functions, respectively, from the R package gamlss⁷². We also checked for an interaction between TN and TP but these did not improve the fit as assessed by the BIC.

The final model (equation (2)) giving $f_{MC}(mc)$ for MC has the following parameters:

$$\begin{aligned} \ln\left(\frac{p_0}{(1-p_0)}\right) &= s(\ln(\text{TN})) - 0.31 \ln(\text{Chl-a}) + s(\ln(A)) - 0.42 \ln(\text{DOC}) \\ &\quad - 0.03 T - 0.45 \text{pH} + 6.89 \\ \ln(\mu) &= 0.68 \ln(\text{TN}) + 0.37 \text{pH} + s(T) + 0.28 \ln(\text{Chl-a}) + 0.34 \ln(D) - 9.81 \\ \ln(\sigma) &= 0.08 \ln(\text{TN}) + 0.06 \end{aligned} \quad (3)$$

The partial effects for $\ln((1-p_0)/p_0)$, $\ln(\mu)$ and $\ln(\sigma)$, are displayed in Fig. 3. Note that since Fig. 3 shows the partial effects for $\ln((1-p_0)/p_0) = -\ln(p_0/(1-p_0))$, the sign in the terms for $\ln(p_0/(1-p_0))$ in equation (3) is reversed in Fig. 3, for example, $-0.3 T$ in equation (3) becomes $0.3 T$ in Fig. 3.

Model validity was assessed using detrended quantile-quantile plots of the normalized quantile residuals⁷⁶, which show a good fit for both the logistic and log-normal model (Extended Data Fig. 8). Concurrency between model predictors was not concerning in both models as assessed via the R package gamlss.ggplots⁷⁷ using the function gamlss.ggplots::get_concurvity(). Also refitting the models using only linear predictors resulted in variance inflation factors never exceeding five, indicating no issues of multicollinearity⁷⁸ for those parameters. The partial effects plots, exceedance probability plots and detrended quantile-quantile plots were also created via the R package gamlss.ggplots⁷⁷.

Once all parameters are estimated from the final model, the probability of crossing a specified microcystin threshold under given environmental conditions can be calculated as one minus the cumulative probability up to that threshold. In contrast to studies conducting multiple logistic regressions to determine exceedance probabilities, this model can thus be applied to any threshold without refitting a new model at each threshold.

On the basis of the exceedance probabilities, we can create intervals for the percentage of samples in this study expected to cross a chosen threshold. For this, we treated each exceedance probability for a sample as a draw from a Bernoulli distribution⁷⁹ (a biased coin flip). We can do this for each sample to simulate whether it exceeds the threshold. We then sum the number of samples for which the threshold was exceeded and divide by the total number of samples. This gives us a simulated percentage of how often samples are expected to exceed the threshold under equivalent environmental conditions. We repeated this 10,000 times to calculate the 0.5th percentile and 99.5th percentile of this distribution and compare it to the observed data.

Spatially averaged detection probabilities and exceedance probabilities for 0.3, 1 and 8 $\mu\text{g l}^{-1}$ were calculated to find broad-scale patterns across the United States within HUC boundaries at the HUC2 and HUC6 level, as developed by the US Geological Survey⁸⁰ (Fig. 2). All base maps of this study have been created with the R package sf⁸¹.

Scenarios

We calculated how the risk of exceeding water quality guidelines changes under warming temperatures at each lake. To do so, we used projections from an ensemble of 15 GCMs participating in CMIP6 (ref. 37) (Extended Data Table 1) run for SSP 2-4.5 (ref. 38), which represents a 'likely' scenario for future climate regulation strategy given current policies. Data from the model projections were bias-corrected and spatially downscaled to 0.25° using the bias correction and spatial disaggregation^{82,83} method. For the bias correction process, the reference air temperature data in the historical period were collected from the Global Meteorological Forcing Dataset⁸⁴.

While the above climate models are predicting air temperature the model is based on water surface temperature. However, in the temperate zones and on a monthly scale, water surface and air temperature track very closely, especially in the range 5–30 °C during summer months^{85,86}.

Under this scenario we investigated how exceedance probabilities have changed and will change under the expected average local air temperature change in the summer months (June–September) for the present day (1990–2019), for the mid-century (2030–2059) and for the late-century (2070–2099) relative to the historical period (1950–1979). The relative change in exceedance probability was calculated as the ratio of the exceedance probability during each of the three periods to the exceeding probability during the historical period. The relative changes were again spatially averaged to identify broad-scale patterns of shifts in occurrence and exceedance probabilities at the HUC2 and HUC6 level. The percentage relative changes in exceedance probability (Fig. 6 and Extended Data Figs. 2 and 3) are defined as:

$$100 \left(\frac{P(\text{MC} > mc)_S - P(\text{MC} > mc)_H}{P(\text{MC} > mc)_H} \right) \% \quad (4)$$

where S represents one of the scenarios (present, mid-century or late-century) and H the historic reference.

The change in exceedance probability we report here between two scenarios can be seen and interpreted as the change between two snapshots that exhibit the average summer temperature of the assessed time periods, respectively.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Observations of microcystin concentrations and environmental parameters are part of the EPA NLA surveys conducted in 2007³¹, 2012⁷⁰ and 2017⁷¹ and are publicly available online at: <https://www.epa.gov/national-aquatic-resource-surveys/data-national-aquatic-resource-surveys>. CMIP6 model output was downloaded from the NASA Earth

Exchange Global Daily Downscaled Projections⁸² and merged with NLA data. Source data are provided with this paper.

Code availability

Data analysis was based on the R packages `gamlss`⁷² and the add-on `gamlss.ggplots`⁷⁷. Specific functions for model selection and formulas are given in the Methods. The model can be accessed in an online tool (<https://algal-risk.dge.carnegiescience.edu/>) where exceedance probabilities of individually selected thresholds of toxin concentration can be calculated under any combination of environmental conditions.

References

- Michalak, A. M. Study role of climate change in extreme threats to water quality. *Nature* **535**, 349–350 (2016).
- Brooks, B. W. et al. Are harmful algal blooms becoming the greatest inland water quality threat to public health and aquatic ecosystems? *Environ. Toxicol. Chem.* **35**, 6–13 (2016).
- Martins, J. C. & Vasconcelos, V. M. Microcystin dynamics in aquatic organisms. *J. Toxicol. Environ. Health B* **12**, 65–82 (2009).
- Meriluoto, J., Spoof, L. & Codd, G. A. *Handbook of Cyanobacterial Monitoring and Cyanotoxin Analysis* (Wiley, 2017).
- Svirčev, Z. et al. Molecular aspects of microcystin-induced hepatotoxicity and hepatocarcinogenesis. *J. Environ. Sci. Health C* **28**, 39–59 (2010).
- Trevino-Garrison, I. et al. Human illnesses and animal deaths associated with freshwater harmful algal blooms—Kansas. *Toxins* **7**, 353–366 (2015).
- Azevedo, S. M. F. O. et al. Human intoxication by microcystins during renal dialysis treatment in Caruaru—Brazil. *Toxicology* **181–182**, 441–446 (2002).
- Sarkar, S. et al. Microcystin exposure worsens nonalcoholic fatty liver disease associated ectopic glomerular toxicity via NOX-2-MIR21 axis. *Environ. Toxicol. Pharmacol.* **73**, 103281 (2020).
- He, J. et al. Health risks of chronic exposure to small doses of microcystins: an integrative metabolomic and biochemical study of human serum. *Environ. Sci. Technol.* **56**, 6548–6559 (2022).
- Ingested Nitrate and Nitrite and Cyanobacterial Peptide Toxins* (IARC, 2010).
- Woolway, R. I., Sharma, S. & Smol, J. P. Lakes in hot water: the impacts of a changing climate on aquatic ecosystems. *BioScience* **72**, 1050–1061 (2022).
- Woolway, R. I. et al. Global lake responses to climate change. *Nat. Rev. Earth Environ.* **1**, 388–403 (2020).
- Gleick, P. H. Water and conflict: fresh water resources and international security. *Int. Secur.* **18**, 79 (1993).
- Adrian, R. et al. Lakes as sentinels of climate change. *Limnol. Oceanogr.* **54**, 2283–2297 (2009).
- O'Reilly, C. M. et al. Rapid and highly variable warming of lake surface waters around the globe. *Geophys. Res. Lett.* **42**, 10773–10781 (2015).
- Lürling, M., van Oosterhout, F. & Faassen, E. Eutrophication and warming boost cyanobacterial biomass and microcystins. *Toxins* **9**, 64 (2017).
- Taranu, Z. E. et al. Acceleration of cyanobacterial dominance in north temperate-subarctic lakes during the Anthropocene. *Ecol. Lett.* **18**, 375–384 (2015).
- Paerl, H. W. & Huisman, J. Blooms like it hot. *Science* **320**, 57–58 (2008).
- 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).
- Kosten, S. et al. Warmer climates boost cyanobacterial dominance in shallow lakes. *Glob. Change Biol.* **18**, 118–126 (2012).
- Christiansen, G., Kurmayer, R., Liu, Q. & Börner, T. Transposons inactivate biosynthesis of the nonribosomal peptide microcystin in naturally occurring *Planktothrix* spp. *Appl. Environ. Microbiol.* **72**, 117–123 (2006).
- Yu, L. et al. The dynamics of *Microcystis* genotypes and microcystin production and associations with environmental factors during blooms in Lake Chaohu, China. *Toxins* **6**, 3238–3257 (2014).
- Davis, T. W., Berry, D. L., Boyer, G. L. & Gobler, C. J. The effects of temperature and nutrients on the growth and dynamics of toxic and non-toxic strains of *Microcystis* during cyanobacteria blooms. *Harmful Algae* **8**, 715–725 (2009).
- Bui, T., Dao, T.-S., Vo, T.-G. & Lürling, M. Warming affects growth rates and microcystin production in tropical bloom-forming *Microcystis* strains. *Toxins* **10**, 123 (2018).
- Wilhelm, S. W., Bullerjahn, G. S. & McKay, R. M. L. The complicated and confusing ecology of *Microcystis* blooms. *mBio* **11**, e00529–20 (2020).
- Hellweger, F. L. et al. Models predict planned phosphorus load reduction will make Lake Erie more toxic. *Science* **376**, 1001–1005 (2022).
- Taranu, Z. E., Gregory-Eaves, I., Steele, R. J., Beaulieu, M. & Legendre, P. Predicting microcystin concentrations in lakes and reservoirs at a continental scale: a new framework for modelling an important health risk factor. *Glob. Ecol. Biogeogr.* **26**, 625–637 (2017).
- Ho, J. C. & Michalak, A. M. Exploring temperature and precipitation impacts on harmful algal blooms across continental U.S. lakes. *Limnol. Oceanogr.* **65**, 992–1009 (2020).
- Hayes, N. M., Haig, H. A., Simpson, G. L. & Leavitt, P. R. Effects of lake warming on the seasonal risk of toxic cyanobacteria exposure. *Limnol. Oceanogr. Lett.* **5**, 393–402 (2020).
- Posch, T., Köster, O., Salcher, M. M. & Pernthaler, J. Harmful filamentous cyanobacteria favoured by reduced water turnover with lake warming. *Nat. Clim. Change* **2**, 809–813 (2012).
- National Aquatic Resource Surveys. National Lakes Assessment 2007* (EPA, 2007); <http://www.epa.gov/national-aquatic-resource-surveys/data-national-aquatic-resource-surveys>
- Gurbuz, F., Metcalf, J. S., Codd, G. A. & Karahan, A. G. Evaluation of enzyme-linked immunosorbent assays (ELISAs) for the determination of microcystins in cyanobacteria. *Environ. Forensics* **13**, 105–109 (2012).
- Loftin, K. A. et al. Cyanotoxins in inland lakes of the United States: occurrence and potential recreational health risks in the EPA National Lakes Assessment 2007. *Harmful Algae* **56**, 77–90 (2016).
- Buley, R. P., Correia, H. E., Abebe, A., Issa, T. B. & Wilson, A. E. Predicting microcystin occurrence in freshwater lakes and reservoirs: assessing environmental variables. *Inland Waters* **11**, 430–444 (2021).
- Paerl, H. W. & Otten, T. G. Blooms bite the hand that feeds them. *Science* **342**, 433–434 (2013).
- Rigby, R. A. & Stasinopoulos, D. M. Generalized additive models for location, scale and shape (with discussion). *J. R. Stat. Soc. C* **54**, 507–554 (2005).
- Eyring, V. et al. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.* **9**, 1937–1958 (2016).
- Fricko, O. et al. The marker quantification of the Shared Socioeconomic Pathway 2: a middle-of-the-road scenario for the 21st century. *Glob. Environ. Change* **42**, 251–267 (2017).
- Cyanobacterial Toxins: Microcystin-LR in Drinking-Water. Background Document for Development of WHO Guidelines for Drinking-Water Quality* (WHO, 2003).
- Recommendations for Cyanobacteria and Cyanotoxin Monitoring in Recreational Waters* (EPA, 2019).

41. Walls, J. T., Wyatt, K. H., Doll, J. C., Rubenstein, E. M. & Rober, A. R. Hot and toxic: temperature regulates microcystin release from cyanobacteria. *Sci. Total Environ.* **610–611**, 786–795 (2018).
42. Crettaz-Minaglia, M. et al. Effect of temperature on microcystin-LR removal and lysis activity on *Microcystis aeruginosa* (cyanobacteria) by an indigenous bacterium belonging to the genus *Achromobacter*. *Environ. Sci. Pollut. Res.* **27**, 44427–44439 (2020).
43. Van de Waal, D. B., Smith, V. H., Declerck, S. A. J., Stam, E. C. M. & Elser, J. J. Stoichiometric regulation of phytoplankton toxins. *Ecol. Lett.* **17**, 736–742 (2014).
44. Van Dam, B. R., Tobias, C., Holbach, A., Paerl, H. W. & Zhu, G. CO₂ limited conditions favor cyanobacteria in a hypereutrophic lake: an empirical and theoretical stable isotope study. *Limnol. Oceanogr.* **63**, 1643–1659 (2018).
45. Beaver, J. R., Tausz, C. E., Scotese, K. C., Pollard, A. I. & Mitchell, R. M. Environmental factors influencing the quantitative distribution of microcystin and common potentially toxigenic cyanobacteria in U.S. lakes and reservoirs. *Harmful Algae* **78**, 118–128 (2018).
46. Reynolds, C. S., Oliver, R. L. & Walsby, A. E. Cyanobacterial dominance: the role of buoyancy regulation in dynamic lake environments. *NZ J. Mar. Freshw. Res.* **21**, 379–390 (1987).
47. Gorham, E. & Boyce, F. M. Influence of lake surface area and depth upon thermal stratification and the depth of the summer thermocline. *J. Gt Lakes Res.* **15**, 233–245 (1989).
48. Jöhnk, K. D. et al. Summer heatwaves promote blooms of harmful cyanobacteria. *Glob. Change Biol.* **14**, 495–512 (2008).
49. Ihle, T., Jähnichen, S. & Benndorf, J. Wax and wane of *Microcystis* (Cyanophyceae) and microcystins in lake sediments: a case study in Quitzdorf reservoir (Germany). *J. Phycol.* **41**, 479–488 (2005).
50. Meerhoff, M. et al. Feedback between climate change and eutrophication: revisiting the allied attack concept and how to strike back. *Inland Waters* **12**, 187–204 (2022).
51. Poste, A. E., Hecky, R. E. & Guildford, S. J. Evaluating microcystin exposure risk through fish consumption. *Environ. Sci. Technol.* **45**, 5806–5811 (2011).
52. Tsoumalakou, E., Papadimitriou, T., Berillis, P., Kormas, K. A. & Levizou, E. Spray irrigation with microcystins-rich water affects plant performance from the microscopic to the functional level and food safety of spinach (*Spinacia oleracea* L.). *Sci. Total Environ.* **789**, 147948 (2021).
53. Codd, G. A., Metcalfe, J. S. & Beattie, K. A. Retention of *Microcystis aeruginosa* and microcystin by salad lettuce (*Lactuca sativa*) after spray irrigation with water containing cyanobacteria. *Toxicon* **37**, 1181–1185 (1999).
54. Zscheischler, J. et al. Future climate risk from compound events. *Nat. Clim. Change* **8**, 469–477 (2018).
55. Schindler, D. W. et al. Eutrophication of lakes cannot be controlled by reducing nitrogen input: results of a 37-year whole-ecosystem experiment. *Proc. Natl Acad. Sci. USA* **105**, 11254–11258 (2008).
56. Paerl, H. The cyanobacterial nitrogen fixation paradox in natural waters. *F1000Research* **6**, 244 (2017).
57. Paerl, H. W., Gardner, W. S., McCarthy, M. J., Peierls, B. L. & Wilhelm, S. W. Algal blooms: noteworthy nitrogen. *Science* **346**, 175–175 (2014).
58. Gobler, C. J. et al. The dual role of nitrogen supply in controlling the growth and toxicity of cyanobacterial blooms. *Harmful Algae* **54**, 87–97 (2016).
59. Zaffiro, A., Rosenblum, L. & Wendelken, S. C. *Method 546: Determination of Total Microcystins and Nodularins in Drinking Water and Ambient Water by Adda Enzyme-linked Immunosorbent Assay* (EPA, 2016).
60. Fischer, W. J. et al. Congener-independent immunoassay for microcystins and nodularins. *Environ. Sci. Technol.* **35**, 4849–4856 (2001).
61. Heussner, A. H., Altaner, S., Kamp, L., Rubio, F. & Dietrich, D. R. Pitfalls in microcystin extraction and recovery from human blood serum. *Chem. Biol. Interact.* **223**, 87–94 (2014).
62. Metcalfe, J. S., Beattie, K. A., Pflugmacher, S. & Codd, G. A. Immuno-crossreactivity and toxicity assessment of conjugation products of the cyanobacterial toxin, microcystin-LR. *FEMS Microbiol. Lett.* **189**, 155–158 (2000).
63. He, X., Stanford, B. D., Adams, C., Rosenfeldt, E. J. & Wert, E. C. Varied influence of microcystin structural difference on ELISA cross-reactivity and chlorination efficiency of congener mixtures. *Water Res.* **126**, 515–523 (2017).
64. Mantzouki, E. et al. Temperature effects explain continental scale distribution of cyanobacterial toxins. *Toxins* **10**, 156 (2018).
65. Harke, M. J. et al. A review of the global ecology, genomics and biogeography of the toxic cyanobacterium, *Microcystis* spp. *Harmful Algae* **54**, 4–20 (2016).
66. Xu, W. et al. Rising temperature more strongly promotes low-abundance *Paramecium* to remove *Microcystis* and degrade microcystins. *Environ. Pollut.* **291**, 118143 (2021).
67. Rouco, M., López-Rodas, V., Flores-Moya, A. & Costas, E. Evolutionary changes in growth rate and toxin production in the cyanobacterium *Microcystis aeruginosa* under a scenario of eutrophication and temperature increase. *Microb. Ecol.* **62**, 265–273 (2011).
68. Zhang, M., Duan, H., Shi, X., Yu, Y. & Kong, F. Contributions of meteorology to the phenology of cyanobacterial blooms: implications for future climate change. *Water Res.* **46**, 442–452 (2012).
69. Thomas, M. K. & Litchman, E. Effects of temperature and nitrogen availability on the growth of invasive and native cyanobacteria. *Hydrobiologia* **763**, 357–369 (2016).
70. *National Aquatic Resource Surveys. National Lakes Assessment 2012* (EPA, 2012); <http://www.epa.gov/national-aquatic-resource-surveys/data-national-aquatic-resource-surveys>
71. *National Aquatic Resource Surveys. National Lakes Assessment 2017* (EPA, 2017); <http://www.epa.gov/national-aquatic-resource-surveys/data-national-aquatic-resource-surveys>
72. Stasinopoulos, D. M. & Rigby, R. A. Generalized additive models for location scale and shape (GAMLSS) in R. *J. Stat. Softw.* **23**, 1–46 (2007).
73. Stasinopoulos, M. D., Rigby, R. A., Heller, G. Z., Voudouris, V. & Bastiani, F. D. *Flexible Regression and Smoothing: Using GAMLSS in R* (Chapman and Hall/CRC, 2017).
74. Rigby, R. A., Stasinopoulos, M. D., Heller, G. Z. & De Bastiani, F. *Distribution for Modelling Location, Scale and Shape: Using GAMLSS in R* (CRC Press, 2020).
75. Schwarz, G. Estimating the dimension of a model. *Ann. Stat.* **6**, 461–464 (1978).
76. van Buuren, S. & Fredriks, M. Worm plot: a simple diagnostic device for modelling growth reference curves. *Stat. Med.* **20**, 1259–1277 (2001).
77. Stasinopoulos, M. D., Rigby, R. A., Bastiani, F. D. & Merder, J. *gamlls.ggplots: Plotting generalised additive model for location, scale and shape. R package version 2.1-2* (2022).
78. Sheather, S. *A Modern Approach to Regression with R* (Springer, 2009).
79. Johnson, N. L., Kotz, S., Kemp, A. W. & Johnson, N. L. *Univariate Discrete Distributions* (Wiley, 1992).
80. Jones, K. A., Niknami, L. S., Buto, S. G. & Decker, D. *Federal Standards and Procedures for the National Watershed Boundary Dataset (WBD)* (USGS, 2022).
81. Pebesma, E. Simple features for R: standardized support for spatial vector data. *R J.* **10**, 439 (2018).
82. Thrasher, B. et al. NASA global daily downscaled projections, CMIP6. *Sci. Data* **9**, 262 (2022).

83. Wood, A. W. Long-range experimental hydrologic forecasting for the eastern United States. *J. Geophys. Res.* **107**, 4429 (2002).
84. Sheffield, J., Goteti, G. & Wood, E. F. Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. *J. Clim.* **19**, 3088–3111 (2006).
85. Livingstone, D. M. & Lotter, A. F. The relationship between air and water temperatures in lakes of the Swiss Plateau: a case study with palaeolimnological implications. *J. Paleolimnol.* **19**, 181–198 (1998).
86. Armitage, D. W. Global maps of lake surface water temperatures reveal pitfalls of air-for-water substitutions in ecological prediction. *Ecography* **2023**, e06595 (2023).

Acknowledgements

The lakes 2007, 2012 and 2017 survey data were a result of the collective efforts of dedicated field crews, laboratory staff, data management and quality control staff, analysts and many others from the EPA, states, tribes, federal agencies, universities and other organizations. Please contact nars-hq@epa.gov with any questions regarding data. Funding for this work was provided by the authors' home institutions.

Author contributions

J.M., A.M.M. and T.H. designed the study. G.Z. processed the CMIP6 data. J.M., A.M.M., D.M.S., R.A.R., G.Z. and T.H. conducted the formal analysis. J.M., A.M.M. and T.H. wrote the original draft of the manuscript. J.M., A.M.M., T.H., G.Z., D.M.S. and R.A.R. contributed to the writing of the final version of the manuscript. A.M.M. supervised the study.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s44221-023-00138-w>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s44221-023-00138-w>.

Correspondence and requests for materials should be addressed to Julian Merder or Anna M. Michalak.

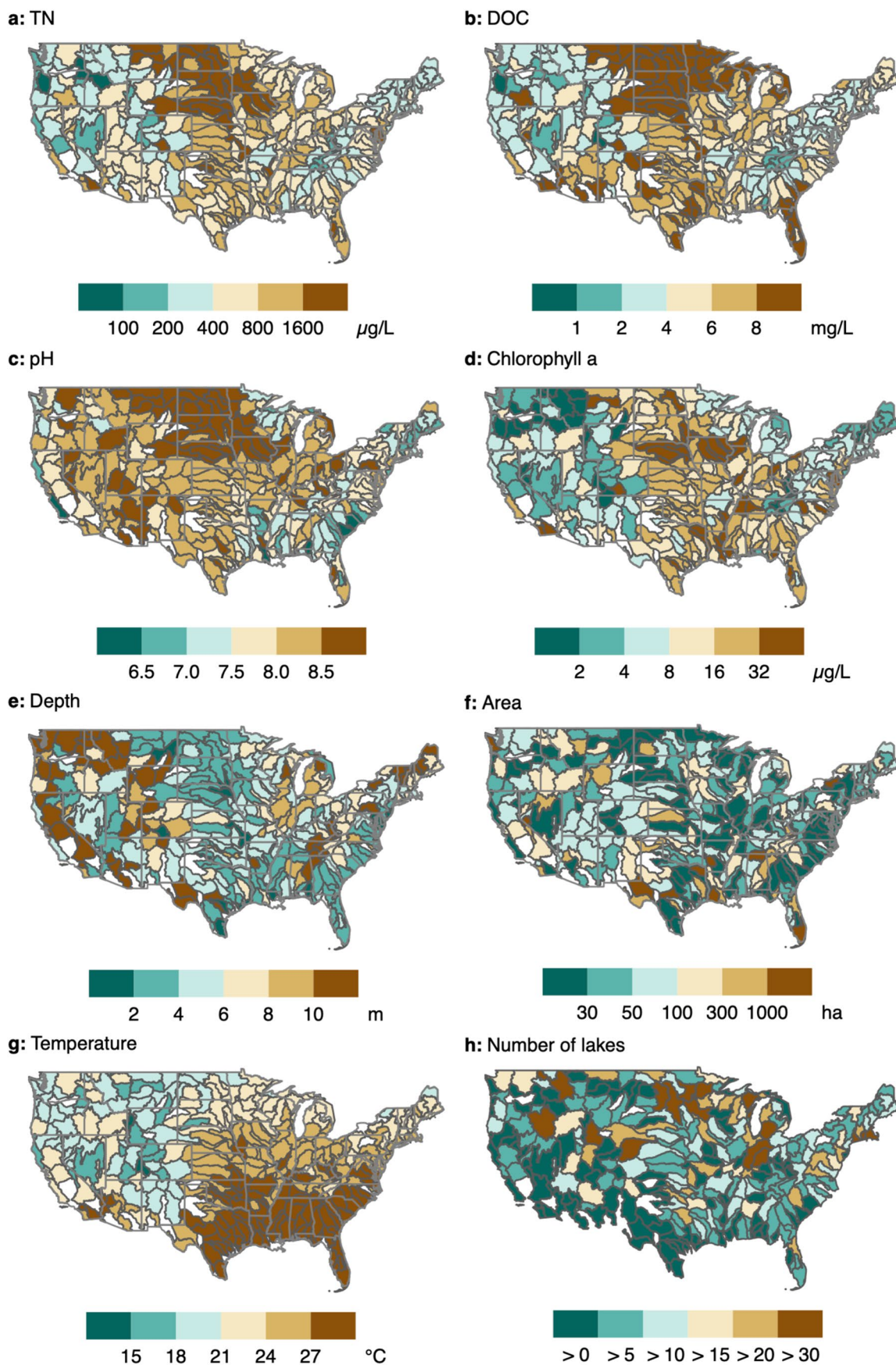
Peer review information *Nature Water* thanks Jolanda Verspagen and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

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

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

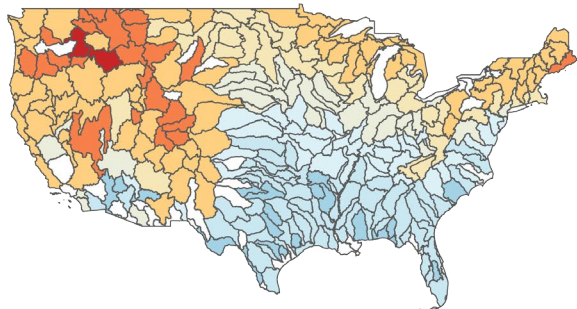
© The Author(s) 2023



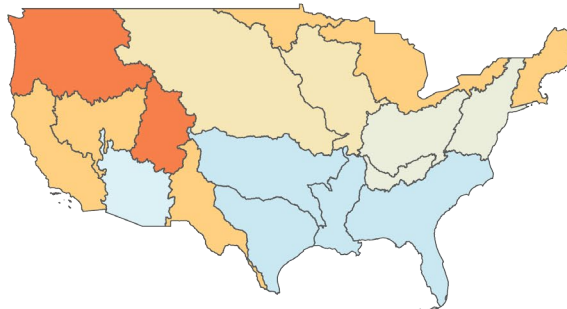
Extended Data Fig. 1 | Spatial distribution of environmental variables selected for inclusion in the model for sampled lakes, averaged to the HUC6 basin scale. These include (a) TN concentration, (b) dissolved organic carbon

concentration, (c) pH, (d) chlorophyll a concentration, (e) depth, (f) area and (g) temperature. Panel (h) shows the number of lakes present in each HUC6 basin. White areas represent basins with no sampled lakes.

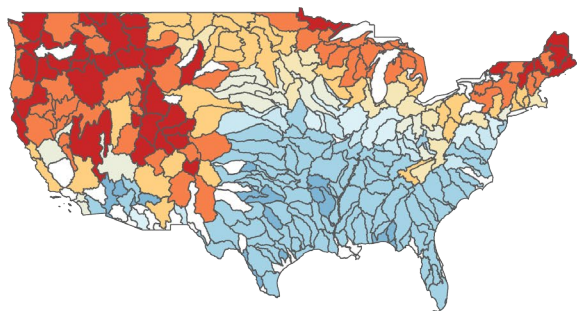
a: Present — historic



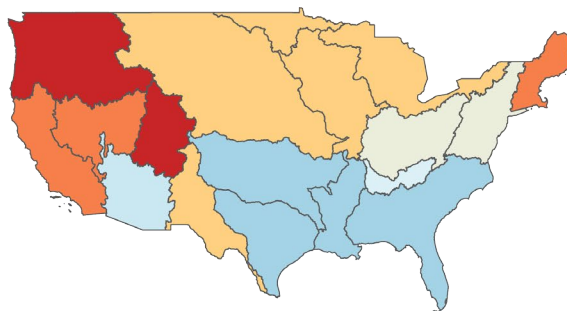
b: Present — historic



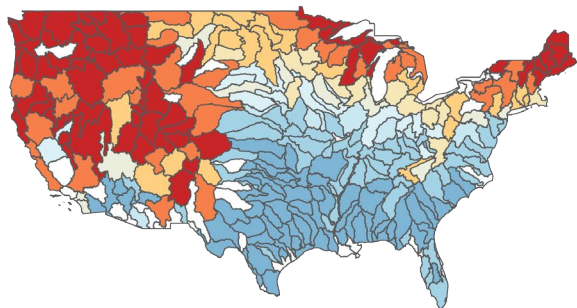
c: Mid-century — historic



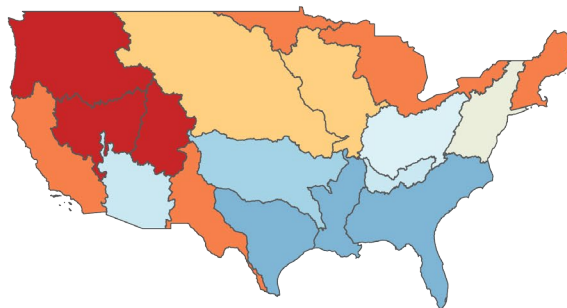
d: Mid-century — historic



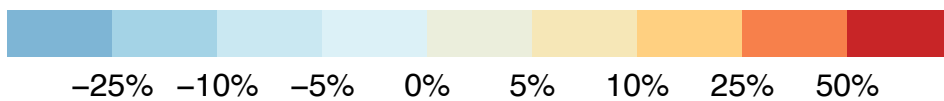
e: Late-century — historic



f: Late-century — historic



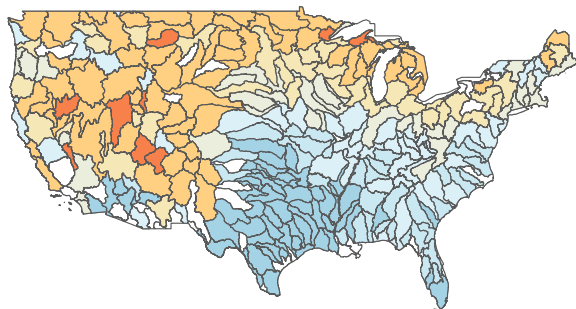
Relative change in exceedance probability



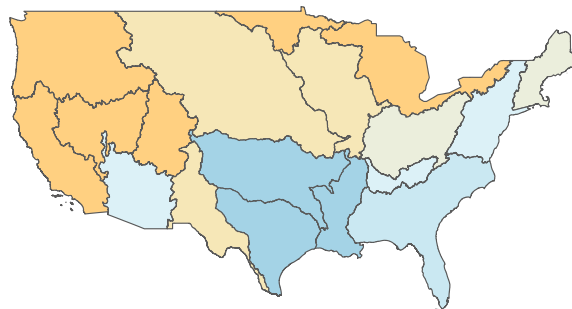
Extended Data Fig. 2 | Temperatures will change the exceedance probability of the adult drinking water guideline (1 $\mu\text{g/L}$). Panels show the relative increase in exceedance probabilities due to warming relative to the historic period (1950–1979), averaged across lakes within individual HUC6 basins (a, c, e)

and HUC2 regions (b, d, f), for (a, b) present (1990–2019), (c, d) projected mid-century (2030–2059) and (e, f) projected late-century (2070–2099) summertime temperatures. All other variables are held constant at their values at the time of sampling. White areas represent basins with no sampled lakes.

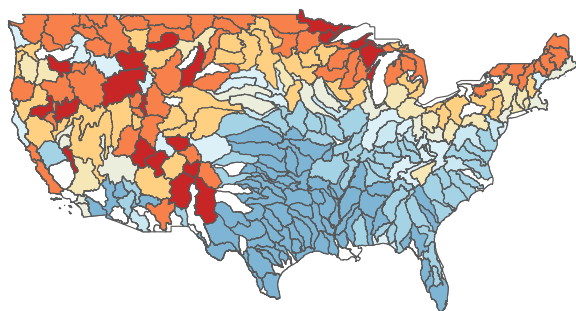
a: Present — historic



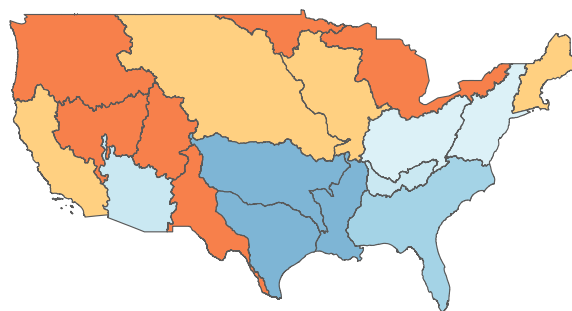
b: Present — historic



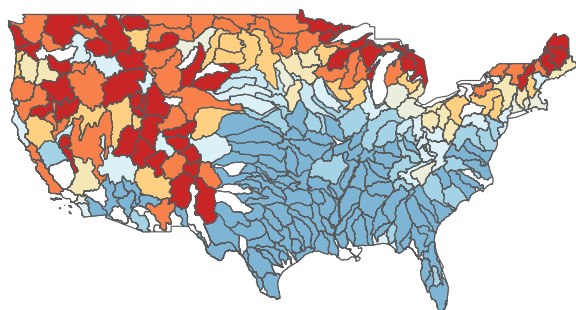
c: Mid-century — historic



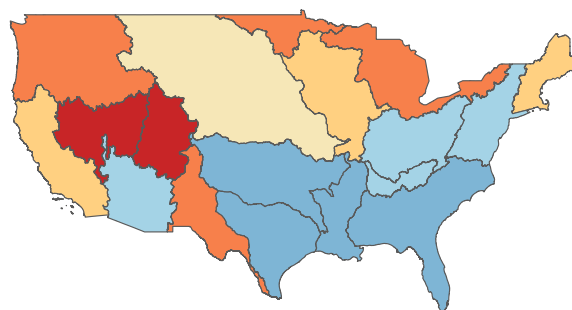
d: Mid-century — historic



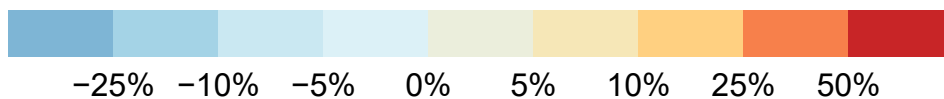
e: Late-century — historic



f: Late-century — historic

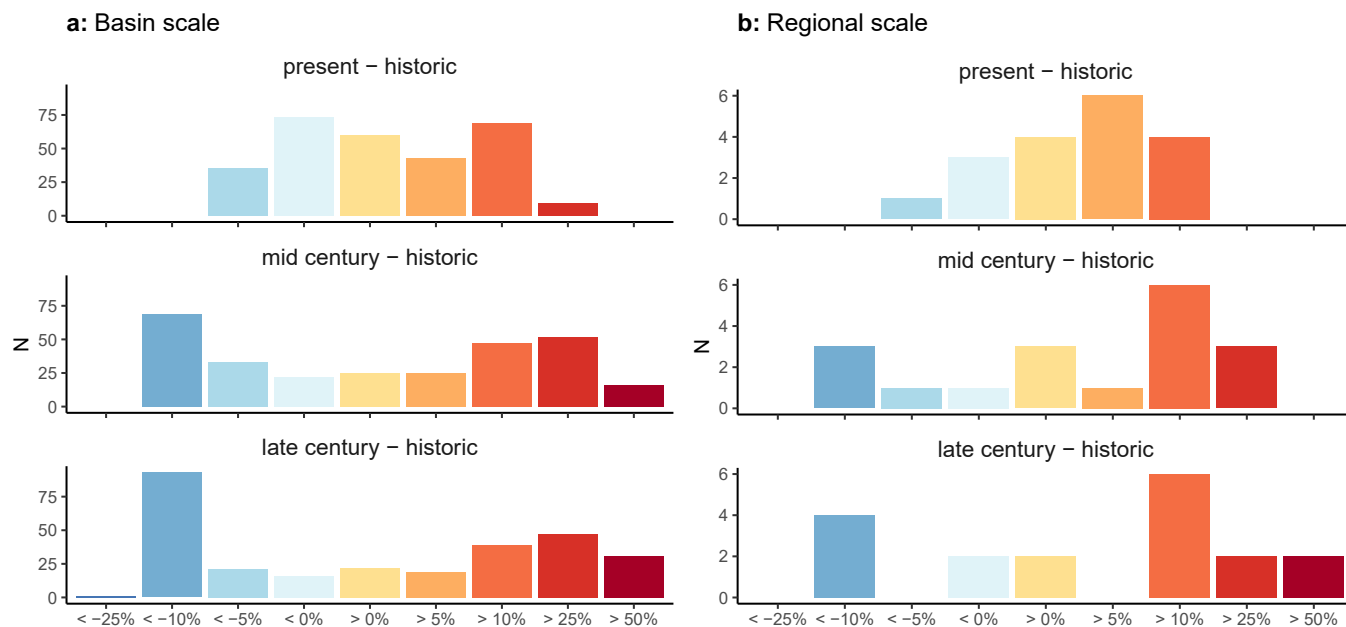


Relative change in exceedance probability



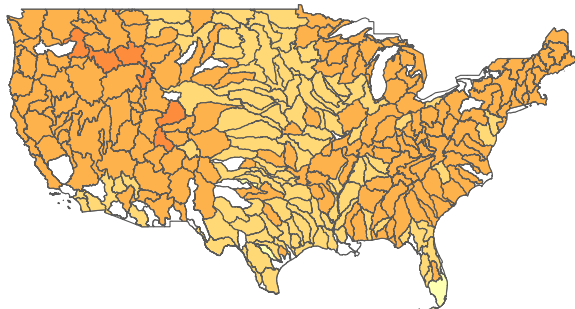
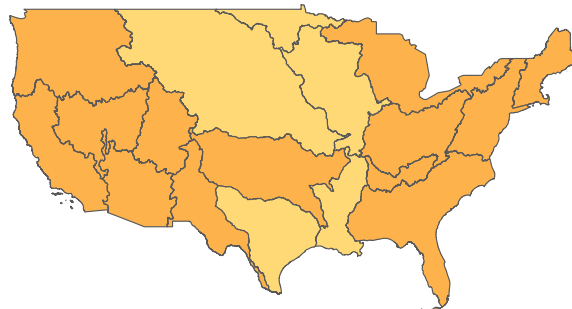
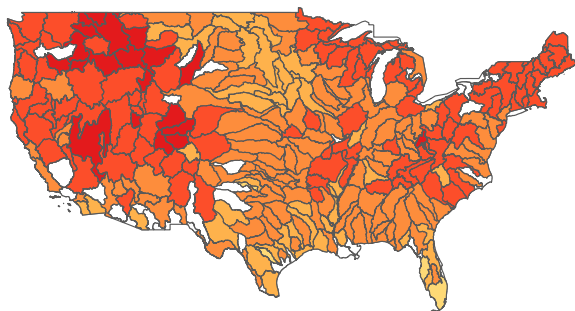
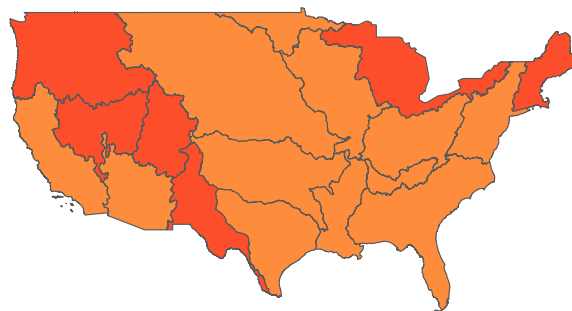
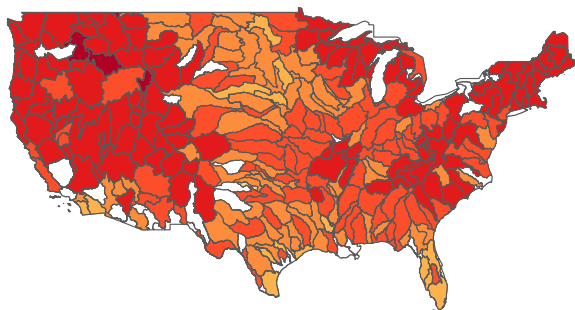
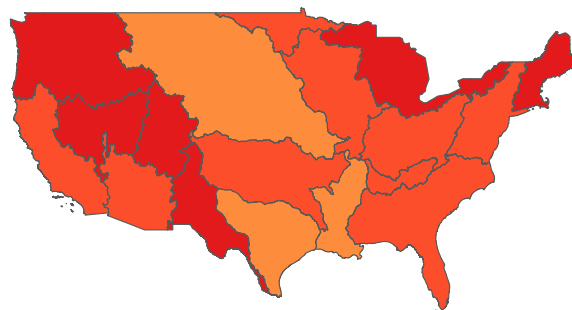
Extended Data Fig. 3 | Temperatures will change the exceedance probability of the recreational water quality criterion (8 $\mu\text{g/L}$). Panels show the relative increase in exceedance probabilities due to warming relative to the historic period (1950–1979), averaged across lakes within individual HUC6 basins (a, c, e)

and HUC2 regions (b, d, f), for (a, b) present (1990–2019), (c, d) projected mid-century (2030–2059) and (e, f) projected late-century (2070–2099) summertime temperatures. All other variables are held constant at their values at the time of sampling. White areas represent basins with no sampled lakes.

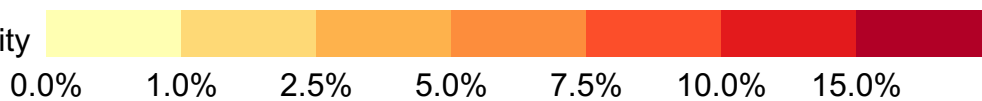


Extended Data Fig. 4 | More areas will experience increases than decreases in the probability of exceeding the microcystin drinking water guideline for children ($0.3 \mu\text{g/L}$) under projected warming. Frequency distribution of percentage relative changes in exceedance probabilities for (a) HUC6 basins and (b) HUC2 regions. Panels show the relative increase (positive) or decrease

(negative) in exceedance probabilities due to warming relative to the historic period (1950–1979), averaged across lakes (a) within individual HUC6 basins and (b) HUC2 regions, for present (1990–2019), projected mid-century (2030–2059) and projected late-century (2070–2099) summertime temperatures. Colours are the same as in Fig. 6.

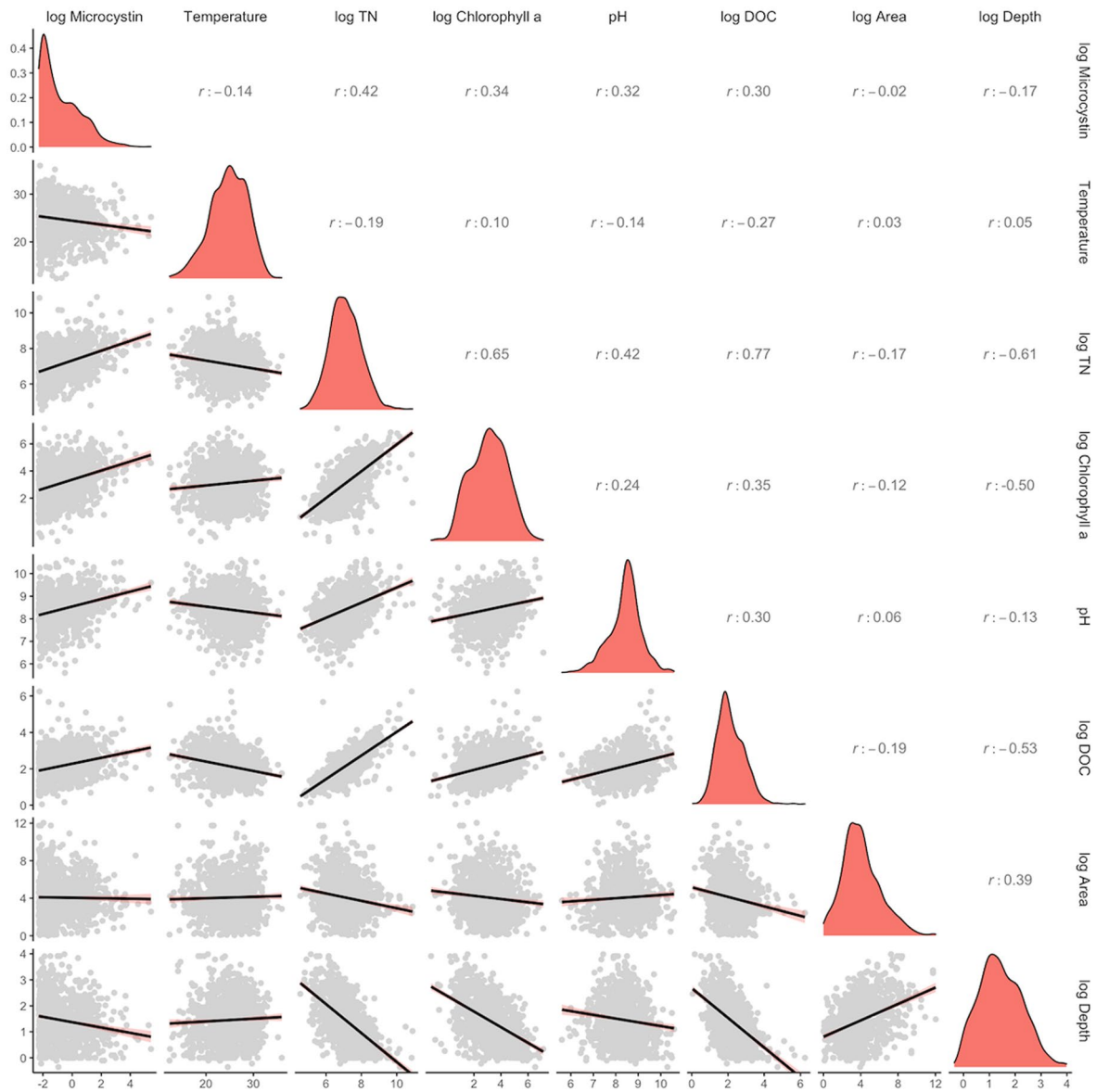
a: Present — historic**b: Present — historic****c: Mid-century — historic****d: Mid-century — historic****e: Late-century — historic****f: Late-century — historic**

Relative change in
exceedance probability



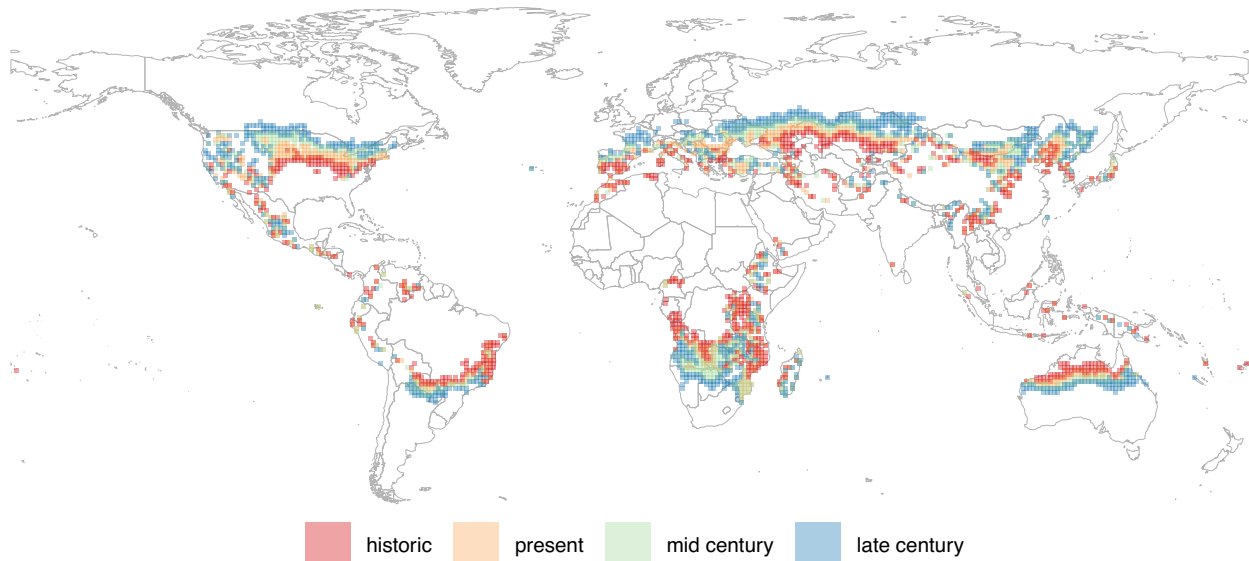
Extended Data Fig. 5 | Rising temperatures will increase the probability of microcystin occurrence. Panels show the relative increase in the probability of microcystin occurrence (above the detection limit of $0.1 \mu\text{g/L}$) due to warming relative to the historic period (1950–1979), averaged across lakes within individual HUC6 basins (**a, c, e**) and HUC2 regions (**b, d, f**), for (**a, b**) present

(1990–2019), (**c, d**) projected mid-century (2030–2059) and (**e, f**) projected late-century (2070–2099) summertime temperatures. All other variables are held constant at their sampled values at the time of sampling. White areas represent basins with no sampled lakes.

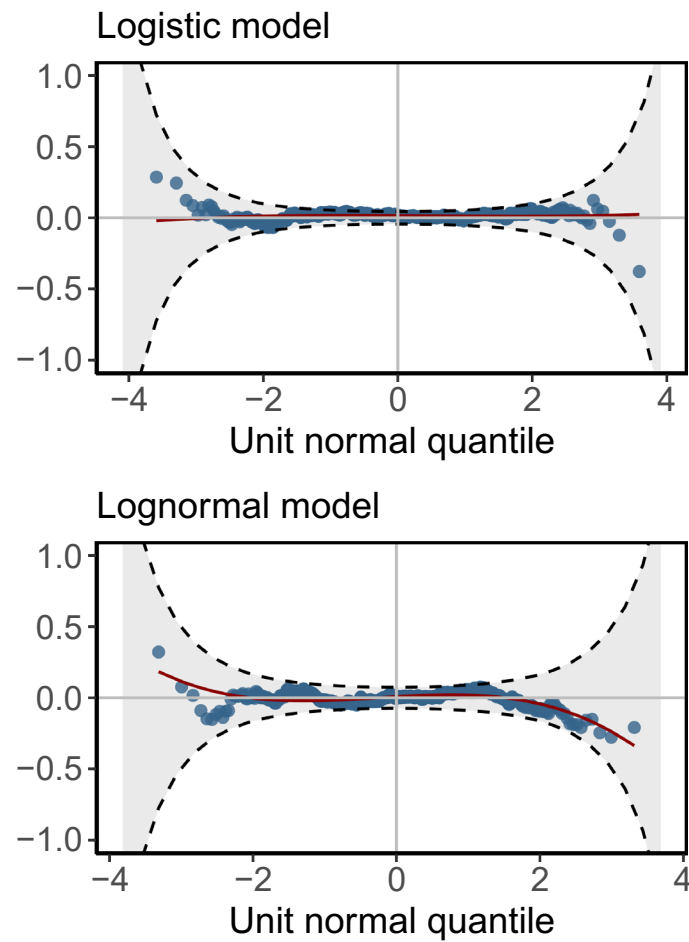


Extended Data Fig. 6 | Correlations between microcystin predictors. Pearson correlation (upper triangle), distribution (diagonal) and linear relationship (lower triangle) between environmental parameters included in the model

measured at sampled lakes, including temperature (°C), log TN (µg/L), log chlorophyll a (µg/L), pH, log DOC (mg/L), log area (ha), log depth (m) and log microcystin (where microcystin > 0.1 µg/L).



Extended Data Fig. 7 | The regions with temperatures with the highest microcystin risk are moving poleward globally. Areas indicating where local summertime average air temperatures range from 21 °C to 23 °C for the historic (1950–1979; pink), present (1990–2019; peach), mid-century (2030–2059; green) and late-century (2070–2099; blue).



Extended Data Fig. 8 | Model validation based on detrended quantile-quantile plots for the logistic and the lognormal model. Points show the deviations of the ordered normalized quantile residuals from their expected normal quantiles plotted against the expected normal quantiles. The shaded

area corresponds to an approximate 95% confidence region around the zero line, within which the actual residual deviations should fall. The red line corresponds to a polynomial fit and should be close to a horizontal line at zero.

Extended Data Table 1 | List of 15 climate models participating in the Coupled Model Inter-comparison Project phase 6 (CMIP6) that were used for projections of future temperatures

Model name	Modeling center	Nominal resolution
BCC-CSM2-MR	Beijing Climate Center, China Meteorological Administration, China	1.125°x1.125°
CanESM5	Canadian Centre for Climate Modelling and Analysis, Canada	2.813°x2.813°
IITM-ESM	Centre for Climate Change Research-Indian Institute of Tropical Meteorology, India	1.915°x1.875°
CNRM-ESM2-1	Center National de Recherches Météorologiques–Center Européen de Recherche et de Formation Avancée en Calcul Scientifique, France	1.406°x1.406°
ACCESS-ESM1-5	Commonwealth Scientific and Industrial Research Organisation, Australia	1.250°x1.875°
MPI-ESM1-2-HR	German Climate Computing Centre, Germany	0.938°x0.938°
INM-CM5-0	Institute for Numerical Mathematics, Russia	1.500°x2.000°
IPSL-CM6A-LR	L'Institut Pierre-Simon Laplace, France	1.259°x2.500°
MIROC-ES2L	Model for Interdisciplinary Research on Climate, Japan	2.813°x2.813°
HadGEM3-GC31-LL	Met Office Hadley Center, United Kingdom	1.250°x1.875°
UKESM1-0-LL	Met Office Hadley Center, United Kingdom	1.250°x1.875°
MRI-ESM2-0	Meteorological Research Institute, Japan	1.125°x1.125°
KACE-1-0-G	National Institute of Meteorological Sciences-Korea Meteorological Administration, South Korea	1.250°x1.875°
GFDL-ESM4	NOAA/Geophysical Fluid Dynamics Laboratory, USA	1.000°x1.250°
NESM3	Nanjing University of Information Science and Technology, China	1.875°x1.875°

The selection of models was based on data availability on the NASA Earth Exchange Global Daily Downscaled Projections (NEX-GDDP-CMIP6⁸²) accessed on September 1st, 2022. For institutes that supply multiple models (for example, ACCESS-CM2 vs ACCESS-ESM1-5), we chose the model with an earth system model (ESM) component (for example, we selected ACCESS-ESM1-5).

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Data analysis

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Data of microcystin concentrations and environmental parameters is part of the Environmental Protection Agency (EPA) National Lake Assessment (NLA) surveys conducted in 2007, 2012 and 2017 and publicly available online under: <https://www.epa.gov/national-aquatic-resource-surveys/data-national-aquatic-resource-surveys>.

Data for climate simulations was downloaded from the NASA Earth Exchange Global Daily Downscaled Projections (see: <https://doi.org/10.7917/OFSG3345> and <https://doi.org/10.1038/s41597-022-01393-4> and references in the manuscript) and merged with NLA data.

Human research participants

Policy information about [studies involving human research participants and Sex and Gender in Research](#).

Reporting on sex and gender	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Data analysis of microcystin risk in the US and changes in risk under global warming.
Research sample	Microcystin concentrations paired with environmental conditions at several thousand lakes in the US measured during the National Lake Assessment (NLA).
Sampling strategy	Survey design of the NLA data is described in: https://www.epa.gov/national-aquatic-resource-surveys/what-waters-are-included-national-aquatic-resource-surveys#nla and https://www.epa.gov/national-aquatic-resource-surveys/national-lakes-assessment-design-documents Data of climate models can be found under: https://doi.org/10.1038/s41597-022-01393-4
Data collection	see Data section
Timing and spatial scale	Information about spatial scale and timing of data from the NLA can be found at: https://www.epa.gov/national-aquatic-resource-surveys/national-lakes-assessment-design-documents
Data exclusions	Only data points with complete measurements of environmental predictors and microcystin concentrations were used. Wrong entries such as surface temperatures of 273 °C were removed from the data.
Reproducibility	Predictions from our model can be done here: https://algal-risk.dge.carnegiescience.edu/ All packages and functions used are publicly available (see Code section).
Randomization	N/A

Blinding

N/A

Did the study involve field work? Yes No

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

- | n/a | Involvement in the study |
|-------------------------------------|--|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Antibodies |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Eukaryotic cell lines |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Animals and other organisms |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern |

Methods

- | n/a | Involvement in the study |
|-------------------------------------|---|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |