

# Water Resources Research

## RESEARCH ARTICLE

10.1029/2018WR023088

### Key Points:

- Winter precipitation and summer temperature are important climate predictors of water quality in thousands of north temperate lakes
- Machine learning approaches enable predictions of lake water quality with climate metrics even when data availability is limited
- Predicted responses of lakes to climate change are highly variable, suggesting strong context dependency

### Supporting Information:

- Supporting Information S1

### Correspondence to:

S. M. Collins,  
sarah.collins@uwyo.edu

### Citation:

Collins, S. M., Yuan, S., Tan, P. N., Oliver, S. K., Lapierre, J. F., Cheruvilil, K. S., et al. (2019). Winter precipitation and summer temperature predict lake water quality at macroscales. *Water Resources Research*, 55, 2708–2721. <https://doi.org/10.1029/2018WR023088>












Received 10 APR 2018

Accepted 2 MAR 2019

Accepted article online 13 MAR 2019

Published online 4 APR 2019

## Winter Precipitation and Summer Temperature Predict Lake Water Quality at Macroscales

S. M. Collins<sup>1,2</sup> , S. Yuan<sup>3</sup> , P. N. Tan<sup>3</sup> , S. K. Oliver<sup>1,4</sup> , J. F. Lapierre<sup>5,6</sup> , K. S. Cheruvilil<sup>7,8</sup> , C. E. Fergus<sup>9</sup> , N. K. Skaff<sup>7</sup> , J. Stachelek<sup>7</sup> , T. Wagner<sup>10</sup> , and P. A. Soranno<sup>7</sup> 

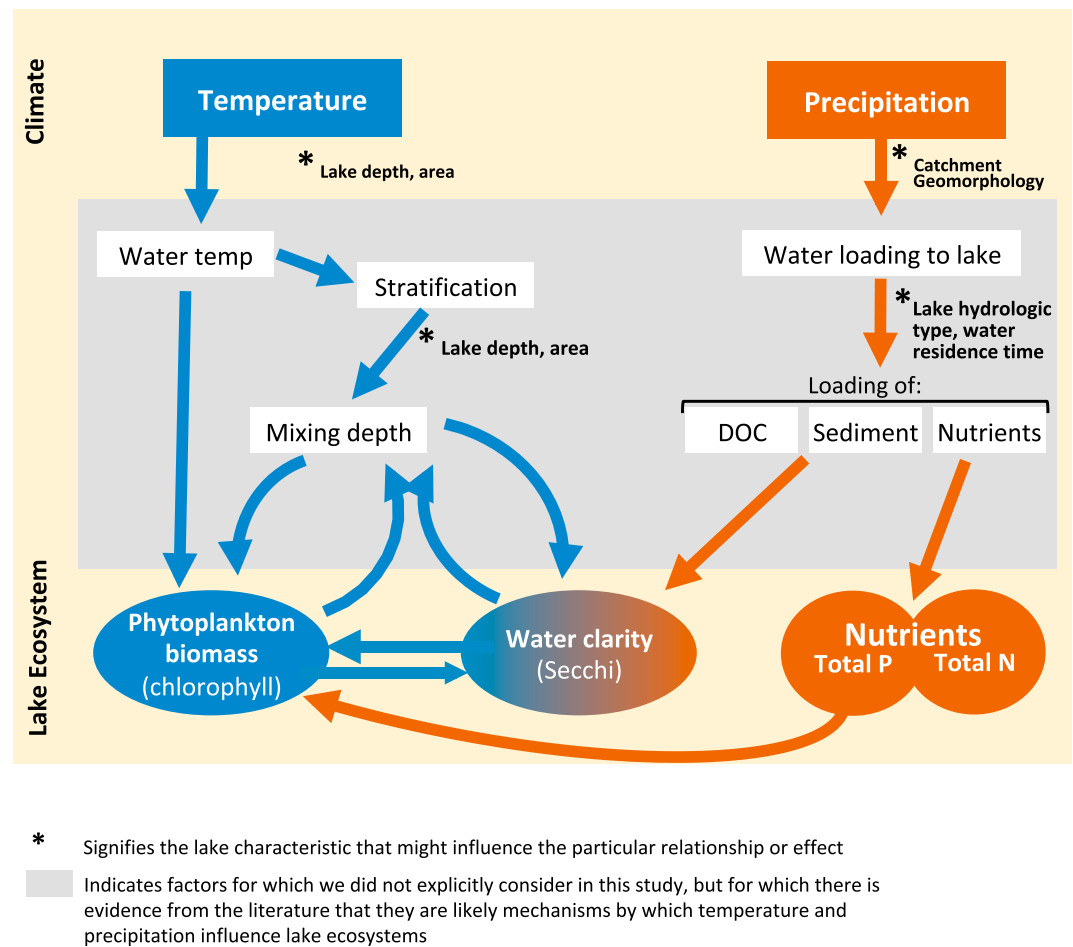
<sup>1</sup>Center for Limnology, University of Wisconsin, Madison, WI, USA, <sup>2</sup>Department of Zoology and Physiology, University of Wyoming, Laramie, WY, USA, <sup>3</sup>Department of Computer Science and Engineering, Michigan State University, East Lansing, MI, USA, <sup>4</sup>Upper Midwest Water Science Center, U.S. Geological Survey, Middleton, WI, USA, <sup>5</sup>Département de Sciences Biologiques, Université de Montréal, Montréal, QC, Canada, <sup>6</sup>Groupe de Recherche Interuniversitaire en Limnologie (GRIL), Montréal, Québec, Canada, <sup>7</sup>Department of Fisheries and Wildlife, Michigan State University, East Lansing, MI, USA, <sup>8</sup>Lyman Briggs College, Michigan State University, East Lansing, MI, USA, <sup>9</sup>The National Research Council, Environmental Protection Agency, Corvallis, OR, USA, <sup>10</sup>U.S. Geological Survey, Pennsylvania Cooperative Fish and Wildlife Unit, The Pennsylvania State University, University Park, PA, USA

**Abstract** Climate change can have strong effects on aquatic ecosystems, including disrupting nutrient cycling and mediating processes that affect primary production. Past studies have been conducted mostly on individual or small groups of ecosystems, making it challenging to predict how future climate change will affect water quality at broad scales. We used a subcontinental-scale database to address three objectives: (1) identify which climate metrics best predict lake water quality, (2) examine whether climate influences different nutrient and productivity measures similarly, and (3) quantify the potential effects of a changing climate on lakes. We used climate data to predict lake water quality in ~11,000 north temperate lakes across 17 U.S. states. We developed a novel machine learning method that jointly models different measures of water quality using 48 climate metrics and accounts for properties inherent in macroscale data (e.g., spatial autocorrelation). Our results suggest that climate metrics related to winter precipitation and summer temperature were strong predictors of lake nutrients and productivity. However, we found variation in the magnitude and direction of the relationship between climate and water quality. We predict that a likely future climate change scenario of warmer summer temperatures will lead to increased nutrient concentrations and algal biomass across lakes (median ~3%–9% increase), whereas increased winter precipitation will have highly variable effects. Our results emphasize the importance of heterogeneity in the response of individual ecosystems to climate and are a caution to extrapolating relationships across space.

## 1. Introduction

Air temperature and precipitation can have strong effects on ecosystem structure and function at a variety of spatiotemporal scales. For example, changes in climate have led to species range shifts, changes in primary production and elemental cycling, and alterations to species interactions in terrestrial, freshwater, and marine ecosystems (Grimm et al., 2013; Hoegh-Guldberg & Bruno, 2010; Parmesan, 2006; Walther et al., 2002). Ecologists are increasingly being asked to predict ecosystem responses to a changing climate at macroscales—regions, continents, and the globe. However, much of our understanding of ecosystem change is based on a relatively small number of long-term studies of individual or groups of ecosystems within regions. Therefore, to forecast the effects of future climate change at macroscales, we must extrapolate the effects of climate change to broader scales and capture the full range of variation in individual ecosystem responses. In this study, we address this problem by using a spatially and temporally extensive database and a novel machine learning method to simultaneously make predictions about thousands of lakes using climate data.

Lakes are ideal models for examining the effects of climate at macroscales because they have discrete boundaries, making them well suited for comparisons across individual ecosystems, and they are low points in the landscape that integrate environmental and climate impacts on their watersheds, making them sentinels of climate change (Adrian et al., 2009; Williamson et al., 2009). Understanding the relationship between climate and lake water quality may be particularly important if a changing climate has a strong influence on water quality, thus altering the capacity for freshwaters ecosystems to provide important services. Many



**Figure 1.** Conceptual figure showing pathways for temperature and precipitation effects on lake water quality response variables (productivity, water clarity, and total nutrients).

connections between climate and lake ecosystems have been made based on conceptual reviews (e.g., Jeppesen et al., 2014; Moss, 2012; Whitehead et al., 2009; Williamson et al., 2009) or studies of a small number of well-studied lakes (e.g., Bertahas et al., 2006; Crossman et al., 2013; Haldna et al., 2008; Pierson et al., 2013; Preston et al., 2016), but analyses at regional to global scales have become more common in recent years (e.g., O'Reilly et al., 2015; Rose et al., 2017).

In this study, our measures of water quality (lake nutrients and productivity) are directly and indirectly affected by climate-mediated processes. Previous work suggests that both precipitation and temperature are mechanistically linked to lake ecosystems (Figure 1), precipitation through the delivery of nutrients from watersheds to lakes (e.g., Arvola et al., 2015; Rose et al., 2017), and temperature by influencing processing rates (e.g., decomposition or primary production) and physical aspects of lakes like stratification (e.g., Kraemer et al., 2015; Kraemer, Chandra, et al., 2017) that can have a strong influence on biology and nutrient cycles. Because different measures of nutrients and productivity are often correlated, modeling them jointly can improve predictions (Wagner & Schliep, 2018).

Based on prior studies and first principles (Figure 1), we expect some climate metrics like winter precipitation (Pierson et al., 2013) and summer temperature (Kraemer, Chandra, et al., 2017) to be important predictors of nutrients and productivity in many lakes. However, recent studies have found heterogeneity in lake responses to climate such that these general predictions might not apply to all lakes. For example, the response of lake water temperature and chlorophyll to changes in air temperature varies globally across lakes (Kraemer, Mehner, et al., 2017; O'Reilly et al., 2015). Similarly, the effects of precipitation on water clarity can interact with factors such as land use and lake morphometry (Rose et al., 2017). This

complexity highlights the need for models that can account for differences across lakes, regions, and ecosystem response variables.

Our goal in this study was to examine whether nutrients and productivity in thousands of lakes across a heterogeneous landscape could be predicted from climate metrics quantified at annual, seasonal, and monthly time scales, including precipitation, air temperature, and several climate indices. Quantifying these relationships is the first step in being able to forecast the response of lake water quality to climate change at macroscales, that is, across regions with broad gradients in climate, land use, and other characteristics. Our three specific objectives were to (1) determine which climate metrics measured at which temporal scales are most important for predicting nutrients and productivity in lakes, (2) evaluate the level of heterogeneity across lakes in their responses to climate, and (3) examine how anticipated climate changes might alter water quality across a diverse population of lakes.

To achieve these objectives, we developed a machine learning method to predict multiple nutrient and productivity variables simultaneously, even in lakes with missing data. Our model was designed to take advantage of two common characteristics of macroscale data—multicollinearity of response variables and spatial autocorrelation (e.g., Cheruvilil et al., 2013; Lapierre et al., 2018). We expected that temperature would have a strong effect on physical processes that control lake productivity, and precipitation would have a strong effect on the delivery of nutrients to the water column (hypotheses outlined in Figure 1), and that seasonal-scale climate metrics would be more important than annual-scale data. Second, we expected to observe a high degree of heterogeneity in the specific climate metrics that predict water quality across lakes given the variety of responses documented in the literature and the diversity of lakes in our study area.

## 2. Methods

### 2.1. Limnological and Climate Data

We conducted our analysis on 11,882 lakes with annual nutrient, water clarity, or primary productivity data in the LAKE GeOSpatial Database (LAGOS; Soranno et al., 2015, 2017). Measures of algal biomass (chlorophyll-*a*, henceforth referred to as chlorophyll), water clarity (measured as Secchi disk depth), total phosphorus (TP), and total nitrogen (TN) were obtained from LAGOS-NE<sub>LIMNO</sub> v. 1.087.1, which spans a 17-state, 1,800,000-km<sup>2</sup> region of the Midwestern and Northeastern United States (Soranno & Cheruvilil, 2017). We used median observations of chlorophyll (as a proxy for primary productivity), Secchi, TP, and TN from the summer stratified season (15 June to 15 September) of each year from 1980 to 2011. We used the inverse of Secchi measurements so they were positively correlated with other response variables, and the direction of effects all different responses were the same (i.e., positive effects on response variables indicate increases in nutrients and algal biomass and declines in water clarity).

We quantified 48 climate metrics at a range of temporal resolutions (Table 1), related to temperature, precipitation, and climate indices. Additional climate-related factors (e.g., wind) can also be important, but we limited our focus to temperature and precipitation for this analysis because data were readily available in LAGOS-NE. First, we calculated annual, seasonal, and monthly averages for precipitation and temperature for the summer in which each water quality response variable was sampled, as well as for the previous spring, winter, and fall. For seasonal data, summer was defined as June, July, and August; Fall as September, October, and November; Winter as December, January, and February; and Spring as March, April, and May. Temperature and precipitation data were from the PRISM climate database and were calculated at the watershed scale (hydrologic unit code-12; HUC-12).

Indices describing anomalies in climate conditions can also be important for lake ecosystems and water quality (Marce et al., 2010; Straile, 2002), so we also examined three climate indices: Palmer Hydrological Drought Index (PHDI), ENSO precipitation index (ESPI), and North Atlantic Oscillation (NAO). PHDI data were acquired from the National Oceanographic and Atmospheric Administration (NOAA) at the climate division scale, which is the finest spatial unit at which they are available. We used PHDI rather than other available drought metrics because PHDI reflects hydrological rather than meteorological drought and thus may better represent groundwater, lake, and reservoir conditions (Karl & Knight, 1985). We assigned PHDI data to each lake by identifying the climate division encompassing the lake or, in instances where the lake was along the U.S. border and did not fall in a climate division, by calculating the distance to the nearest

**Table 1**  
*A Description of the Climate Metrics Used in the Analysis*

		Time periods	Variables
Annual data	Sample year	Annual	Mean temperature, maximum temperature, minimum temperature, precipitation, El Nino Southern Oscillation, and North Atlantic Oscillation
	Previous year	Annual	Mean temperature, maximum temperature, minimum temperature, precipitation, El Nino Southern Oscillation, and North Atlantic Oscillation
Seasonal data	Sample year	Winter, spring, and summer	Mean temperature, precipitation, and Palmer drought index
	Previous year	Fall	Mean temperature, precipitation, and Palmer drought index
Monthly data	Sample year	January, February, March, April, May, June, July, August	Mean temperature and precipitation
	Previous year	September, October, November, and December	Mean temperature and precipitation

*Note.* Sample year indicates that data were from the same calendar year as the response variable observation; previous year indicates that data were from the calendar year before the response variable observation.

climate division (Bivand et al., 2013; Pebesma & Bivand, 2005). ESPI describes the strength of the El Nino Southern Oscillation based on precipitation anomalies in two regions in the Pacific Ocean. NAO is based on anomalies in sea level pressure at a station in the North Atlantic Ocean and can be associated with climate patterns in Europe and North America. ESPI and NAO data are not spatially explicit and were calculated at the annual scale. All features were normalized prior to analysis to have zero mean and unit variance. A full list of monthly, seasonal, and annual climate predictors is summarized in Table 1 and data are archived with the Environmental Data Initiative (Collins et al., 2018).

## 2.2. Prediction With Multitask Learning

We used a novel machine learning model to predict lake nutrients, chlorophyll, and water clarity (Secchi) using the 48 climate metrics. Although this analysis included data from many lakes at a broad spatial extent and up to 30 years of data in some lakes, data availability was unbalanced across both space and time. Our modeling approach was designed to overcome the challenges of limited or missing data in making predictions about lake nutrients and productivity with climate data.

Specifically, multitask learning (MTL) allows a complex prediction problem to be decomposed into smaller subproblems or *tasks* that are solved simultaneously by exploiting the joint relationship between the tasks (Caruana, 1997). For example, in our data set with  $N$  lakes and  $M$  response variables, the modeling of a water quality response variable (i.e., chlorophyll, nitrogen, phosphorus, or water clarity) in each lake is considered a separate task. The conventional approach would be to solve each modeling task separately; assuming a linear relationship between climate predictors and a response variable, the regression coefficients for each response variable in each lake are often derived independently of other lakes or response variables. However, our approach incorporates two factors that we expect to greatly improve predictions of water quality in lakes: first that different water quality response variables can be correlated (Dillon & Rigler, 1974; Downing & McCauley, 1992; Filstrup et al., 2017) and second that water quality response variables are known to be spatially autocorrelated (Lapierre et al., 2015, 2018). Our model leverages both spatial relationships and joint responses to improve model performance. Spatial autocorrelation is estimated based on the spatial distance, though it can also be inferred from their correlation in the training data. If the spatial autocorrelation is low or does not exist, then the MTL approach reduces to training the model for each lake independently. Nevertheless, the beauty of our proposed formulation is that it can still exploit the dependencies between correlated response variables to improve the model, which is especially helpful when there are limited training samples available for some response variable.

Our multitask learning approach models the response variables for all the lakes by optimizing the following joint objective function:

$$\min_{(\beta)} \sum_{j=1}^M \sum_{i=1}^N \|X_i \beta_{ij} - y_{ij}\|^2 + \Omega(\beta)$$

where  $X_i$  is a matrix containing the time series of climate predictors for the  $i$ th lake,  $y_{ij}$  is the time series for the  $j$ th response variable of the  $i$ th lake, and  $\beta_{ij}$  is the regression coefficient vector for the  $i$ th lake and  $j$ th

response variable. Note that the first term in the objective function corresponds to the sum of residual errors between the observed and predicted response values, whereas the second term enforces a set of constraints on the regression coefficients found for the different tasks. Specifically, we designed  $\Omega(\beta)$  to ensure that the estimated regression coefficients would preserve the spatial autocorrelation between lakes and the correlation between response variables:

$$\Omega(\beta) = \lambda_1 \sum_{i=1}^N \sum_{j=1}^M \sum_{k=1}^M c_{jk} \|\beta_{ij} - \beta_{ik}\|^2 + \lambda_2 \sum_{j=1}^M \sum_{p=1}^N \sum_{q=1}^N a_{pq} \|\beta_{pj} - \beta_{qj}\|^2 + \lambda_3 \|\beta\|_1$$

where  $c_{jk}$  is the absolute correlation between the  $j$ th and  $k$ th response variables and  $a_{pq}$  is the spatial proximity between the  $p$ th and  $q$ th lakes. Conceptually, the first term of  $\Omega(\beta)$  restricts the regression coefficients to be similar if two response variables are strongly correlated with each other and the second term ensures that the regression coefficients for two nearby lakes should not differ substantially. The last term in  $\Omega(\beta)$  corresponds to a lasso regularization term, which is used to control sparsity of the regression coefficients (i.e., it helps accommodate the fact that many climate predictors are correlated with one another). Increasing the value of  $\lambda_3$  would zero out many of the regression coefficients, which is beneficial to improve model interpretability and control model overfitting. However, if  $\lambda_3$  is too large, this may underfit the model and lead to higher prediction error. We tested a range of  $\lambda_3$  values and selected an intermediate value (0.03) to avoid model overfitting and underfitting.

To evaluate the performance of the model, we randomly chose two thirds of the data as our training data set and the remaining one third as our test data set. Lakes with only a single observation were included in the test set. We iterated this procedure 10 times to evaluate how sensitive the model was to the train-test data partition. We found that the standard deviation of the root-mean-square error (RMSE) over those 10 runs was low, suggesting low sensitivity (results in supporting information Table S1). We performed cross validation on the training set to determine the hyperparameter values that minimized RMSE over the four response variables. The MTL model allowed us to estimate the regression coefficients and make predictions even when there were no training data for an individual lake. We identified which climate metrics were important predictors across all lakes by summing the number of times each climate metric was among the most important (i.e., had one of the three largest absolute value standardized regression coefficients) for each lake using the MTL model. To determine whether all lakes had similar responses to climate, we examined the lake-specific regression coefficients from the MTL graphically using heat maps. We also clustered the coefficients using two clustering approaches ( $k$ -means and hierarchical) to determine whether lakes could be clustered into groups that had a similar relationship with climate predictors. Our attempts at clustering included a wide range of possible  $k$  values in  $k$ -means and cutoff heights for the hierarchical approach. Finally, we mapped the lake-specific RMSE for each response variable to determine whether model fit varied geographically and correlated RMSE to the number of observations for each lake and to the median value of the response variable.

To quantify the likely effect of climate change on lake water quality by 2050, we summarized the direction and magnitude of the most important climate predictors and used those predictors to quantify the expected percent change in lake ecosystem properties. First, we calculated the median summer temperature or winter precipitation for each lake across the study sites. Then, we added a conservative estimate of anticipated temperature (1.7 °C) or precipitation (15%) change for our study area to the medians, based on downscaled climate projections for the Great Lakes region based on an RCP4.5 emissions scenario in Byun and Hamlet (2018). Finally, using the sum of the winter precipitation or summer temperature regression coefficients from the MTL model, we calculated the percent change in each response variable for each lake under this climate change scenarios. This is a highly simplified approach intended to give context to our results and estimate how lakes might respond to changing climate based on the strength of the relationship between climate and water quality. More complex modeling exercises to make lake-specific projections for all climate predictors would be useful extensions of this work.

Data manipulation, preparation, and summaries of the model results were conducted in R (R Core Team, 2017), and the code and data are available on Github (<https://github.com/smc322/LAGOSclimate>). The MTL modeling was conducted in Matlab and is also available on Github (<https://github.com/shuaiyuan-msu/csi-mtmr>).



**Table 2***Descriptive Statistics for Lake Water Quality Data Used in the Analysis*

	Chlorophyll	Secchi depth	TP	TN
Number of observations	35,820	70,218	33,713	8,828
Number of lakes	7,393	10,247	8,245	2,398
Mean	17 $\mu\text{g/L}$	3.1 m	33 $\mu\text{g/L}$	819 $\mu\text{g/L}$
Median	5.6 $\mu\text{g/L}$	2.7 m	16 $\mu\text{g/L}$	541 $\mu\text{g/L}$
Maximum	696 $\mu\text{g/L}$	18 m	1123 $\mu\text{g/L}$	19,691 $\mu\text{g/L}$
Minimum	0 $\mu\text{g/L}$	0 m	0 $\mu\text{g/L}$	0 $\mu\text{g/L}$
Average number of years with at least one observation	5	7	4	4
Lakes with 2+ years of observations	5,383	7,065	4,424	1,315
Lakes with 5+ years of observations	2,490	4,271	2,079	584
Lakes with 15+ years of observations	456	1,738	559	91

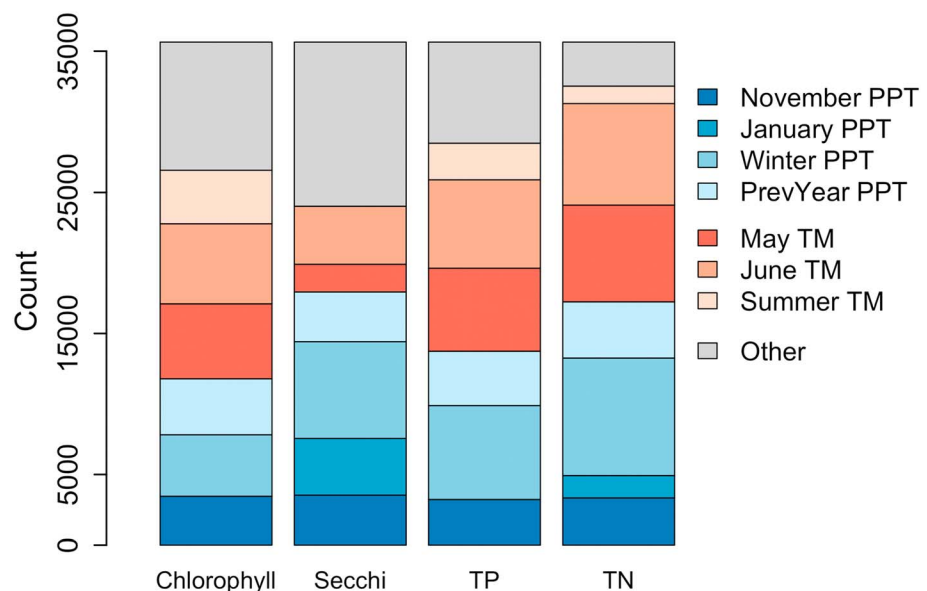
*Note.* TP is total phosphorus, and TN is total nitrogen.

### 3. Results

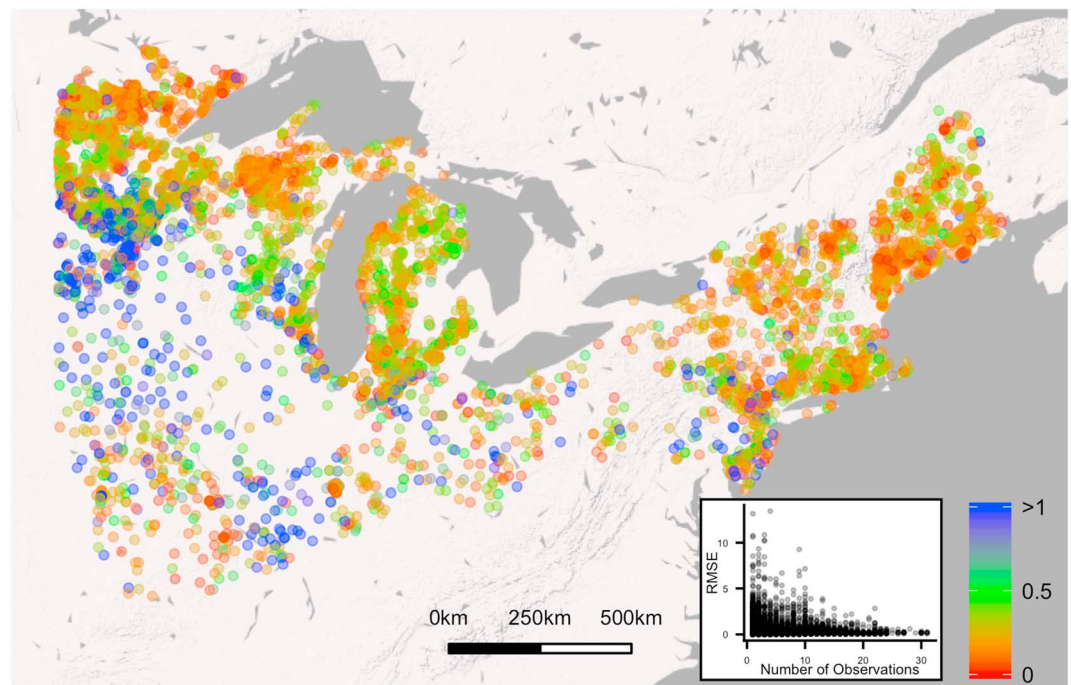
We observed a wide range of nutrient concentrations, algal biomass, and water clarity in the study lakes (Table 2). Data availability varied across the four response variables, with the most observations for Secchi depth, followed by an intermediate number for TP and chlorophyll, and the smallest number for TN (Table 2). Despite having samples from many lakes, most response variables included only a few years of observations per lake and few lakes included long time series of 15 or more years (Table 2).

#### 3.1. Comparison of Climate Metrics

We found that a small number of the 48 climate metrics calculated at different temporal scales were consistently among the top predictors across all lakes (Figure 2). These included metrics associated with precipitation from the previous winter (November and January monthly precipitation and winter seasonal precipitation) and temperature during the summer in which nutrients and productivity were sampled (May and June monthly mean temperatures and summer seasonal mean temperature, Figure 2).



**Figure 2.** Count of the most important predictors (i.e., the three predictors with highest effect sizes for each lake) across all lakes for each response variable. Climate metrics related to summer temperature (i.e., mean temperature, abbreviated TM, for May and June, and the summer season, shown in shades of orange) and winter precipitation (i.e., monthly precipitation, abbreviated PPT, for January and November, and precipitation for the winter season, shown in shades of blue) were consistently the top predictors for many lakes. For each lake in our study population (~12,000), three values for the top three effect sizes are represented in each bar, for a total count of ~35,000 in this analysis.



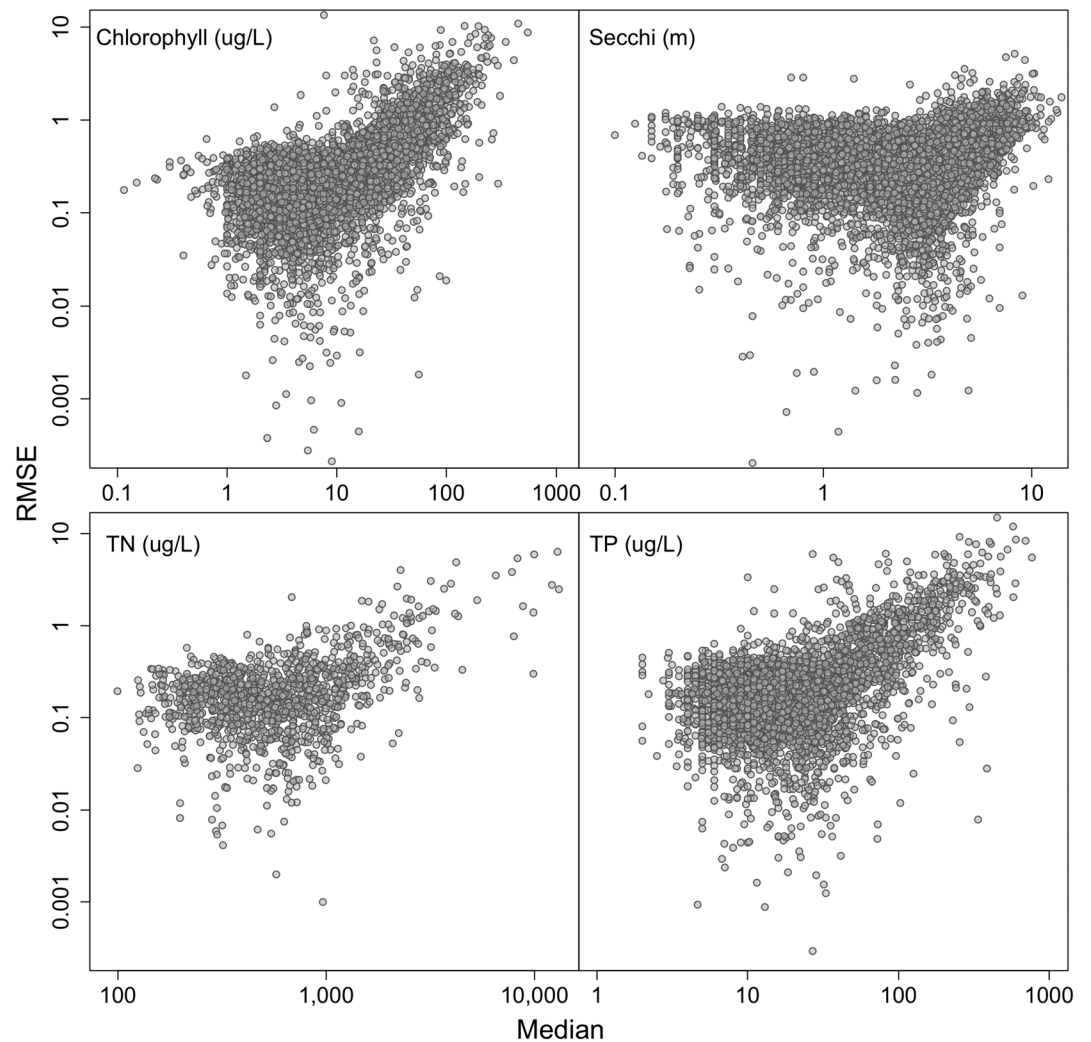
**Figure 3.** Root-mean-square error (RMSE) for predictions in each lake for chlorophyll, with an inset showing the relationship between RMSE and the number of observations for each lake. Corresponding figures for other response variables are included in supporting information.

### 3.2. Heterogeneity of Lake Responses to Climate and Projected Climate Change Effect on Lakes

The effects of climate on lake water quality varied geographically, and we found spatial patterns in the quality of predictions. The lake-specific RMSE from the MTL model shows large differences in the amount of error associated with predictions across lakes and regions (Figure 3). In particular, for chlorophyll, TP, and TN, RMSE was relatively high for lakes located in the Midwestern U.S. (Figures 3 and S1–S3). RMSE was not significantly correlated to the number of observations for each individual lake (Figures 3 and S1–S3) but was related to the median concentration of each response variable (Figure 4). Specifically, error was lowest in lakes with low nutrient or chlorophyll concentrations and lowest in lakes with intermediate Secchi values.

Although monthly or seasonal predictors related to winter precipitation and summer temperature were consistently important for all water quality responses, there was large variation in the identity of important predictors across lakes, and in the direction of relationships (Figure 5). Despite this variation across lakes, both *k*-means and hierarchical clustering on the regression coefficients assigned all lakes into a single cluster. This result suggests that all lakes responded similarly enough to key climate predictors of lake water quality that there are no uniquely distinguishable groups. Despite the fact that groups of lakes could not be distinguished with clustering, some lakes had more obvious, strong responses to the dominant climate predictors, while others had more uniform responses to all predictors (Figure 5). There was also more cross-lake variation in the importance of individual predictors to Secchi than there was for nutrients (both TN and TP), which were more uniformly associated with the consistently important winter precipitation and summer temperature metrics (Figure 5).

The sum of the effect sizes for the most important winter precipitation and summer temperature metrics demonstrates a cross-lake difference in the direction and variability of the effects of winter precipitation versus summer temperature (Figure 6a). Specifically, combined summer temperature effects are positive for the four response variables in most lakes. This is in contrast with the generally negative winter precipitation effects for the four response variables that also showed more variability in directionality and effect size (Figure 6a).



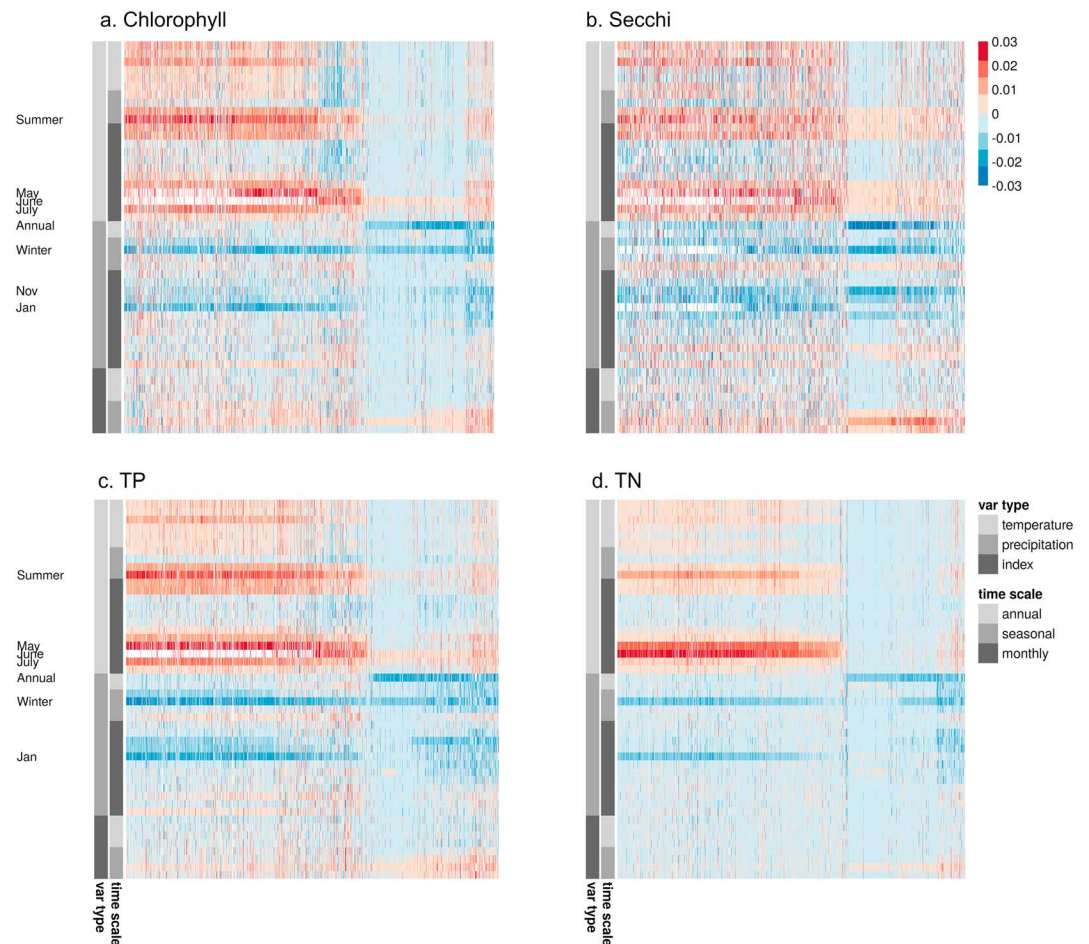
**Figure 4.** Relationships between median water quality condition (nutrient or chlorophyll concentrations, or Secchi depth) and RMSE of climate predictions. Error was lowest in relatively oligotrophic lakes based on nutrient or chlorophyll concentrations and lowest in lakes with intermediate water clarity based on Secchi depth.

We also found large variation in the direction and magnitude of potential percent changes in lake water quality as a result of anticipated changes in summer temperature or winter precipitation (Figure 6). Our model results suggest that on average, lakes are likely to experience positive changes in all response variables with increasing summer temperature (median 3.3% to 8.6% across the four response variables) but that individual lakes vary substantially in their predicted response (Figure 6b). In contrast, increasing precipitation may have a more neutral average effect on all responses (median  $-1.1\%$  to  $-0.5\%$  across the four response variables), but that variation in magnitude and directionality remained large (Figure 6b). The negative effects of precipitation in many lakes suggest that there may be context dependency in the relationship, and perhaps, the overall amount of precipitation influences the relationship between precipitation and water quality.

#### 4. Discussion

Our results demonstrate that summer temperature and winter precipitation are consistently important predictors of lake water quality but that there is also heterogeneity in responses across lakes. Studies of individual lakes and theory (summarized in Figure 1) provide strong support for these summer temperature and winter precipitation effects. For example, summer temperature can have strong effects on physical processes

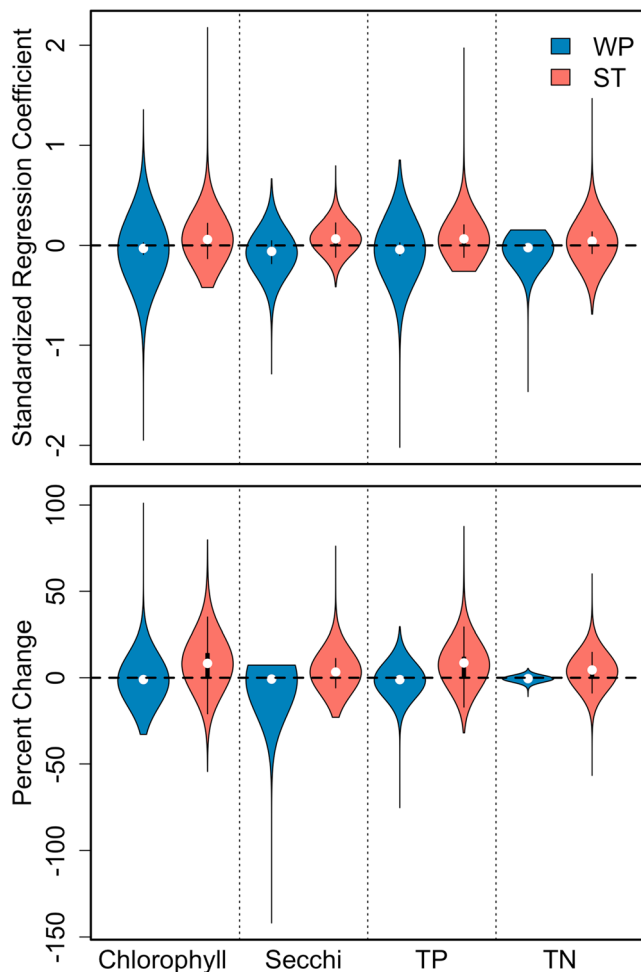




**Figure 5.** Heat maps showing the standardized regression coefficients from the MTL model for each climate predictor (rows) in each lake (columns, arranged by longitude) for (a) chlorophyll, (b) Secchi depth, (c) TP, and (d) TN. The types of climate predictors are indicated by time scale (annual, monthly, and seasonal) and type (temperature, precipitation, and index). The most important climate predictors are labeled on the left. TP is total phosphorus, and TN is total nitrogen. Secchi effects are reversed because inverse of data were used in the model.

like water column stratification, which in turn can influence mixing depth and phytoplankton biomass. Previous work confirms this mechanism, demonstrating that abnormally high summer temperatures can lead to very high thermal stability and hypolimnetic oxygen depletion (Jankowski et al., 2006). Summer temperature is also related to rates of primary production and other biological processes that have a strong effect on primary production, N and P cycling, and water clarity.

Moreover, higher temperatures have been linked to altered timing and reduced duration of ice cover (Magnuson et al., 2000) and winter conditions can influence subsequent summer concentrations of nutrients (Hampton et al., 2017). Winter precipitation, on the other hand, can affect the amount of light available in the water under ice (Hampton et al., 2017; Leppäranta, 2015) and is strongly related to spring nutrient loading, which is directly related to summer nutrient concentrations, primary production, and water clarity. Consequently, future changes in summer temperature and winter precipitation are most likely to alter nutrient dynamics and production in lakes, which we quantified through a simplistic forecasting exercise. Although these general patterns appear to apply to most lakes in our heterogeneous study area, we also observed relatively high across lake variation in relationships and model fit, suggesting that caution should be exercised when extrapolating results from individual regions or lakes to broad spatial scales. In particular, our results suggest that error is highest in lakes with high nutrient and chlorophyll concentrations. This increased error might relate to the number of other stressors (e.g., land use and increased nutrient loading) that obscures the effects of climate on water quality.



**Figure 6.** Violin plot of the standardized regression coefficients for each lake for (a) the effect of summer temperature (orange, ST) and the effect of winter precipitation (blue, WP) on water quality response variables and (b) the percent change in each response variable expected under anticipated changes in summer temperature and winter precipitation, colored the same as in panel a. TP is total phosphorus, and TN is total nitrogen. Secchi effects are reversed because inverse data were used in the model to accommodate for positive correlations between response variables.

#### 4.1. Heterogeneity in the Effect Sizes of Climate Predictors Across Lakes

Our results confirm our expectation that seasonal climate metrics (including both the seasonal average and monthly data during those seasons) are the most relevant temporal scale for predicting water quality in lake ecosystems. However, there is substantial variation in predictions across lakes, suggesting that extrapolation from one ecosystem to another may be difficult. This large variability suggests that some lakes may be much more sensitive to climate change than others, emphasizing the need for lake-specific predictions about the effects of climate. Further, the variation in directionality means that the effects of temperature and precipitation on lakes may be contradictory and could obscure each other. Nonetheless, a simple forecasting exercise revealed that expected changes in summer temperature and winter precipitation may result in ecologically meaningful changes in lake water quality. The effect sizes from the model suggest that changes in temperature under climate change scenarios projected to 2050 might lead to 3%–8% changes in water quality. Changes in precipitation led to more variable changes, with some lakes increasing and some decreasing. A more detailed forecasting exercise that considers the different climate projections for lakes over space and a full suite of climate predictors would be required to quantify water quality responses to climate change accurately, but this simplified version demonstrates a consistent effect of summer temperature across thousands of lakes, which may result in increased lake nutrients and productivity within a few decades.

While we did identify variation in the magnitude and direction of the effects of climate predictors across lakes, it was not possible to organize those effects into groups that we could associate with lake or watershed characteristics or regions. While climate-related predictors are one important driver of lake water quality, there are other known drivers that were not included in our model and may obscure effects of climate, including land use (e.g., Arbuckle & Downing, 2001; Collins et al., 2017; Fraterrigo & Downing, 2008), lake morphometry and internal lake processes (e.g., Read et al., 2015), atmospheric deposition (e.g., Crowley et al., 2012), and hydrological connectivity (e.g., Cardille et al., 2007; Fergus et al., 2017). Incorporating those factors into our MTL approach was not possible because they are relatively static through time or have undocumented temporal variation, which is required for treating individual lakes as separate tasks in the model.

If clustering approaches had grouped lakes into more than one cluster, we might have been able to relate cluster membership to spatial variables (e.g., land use or land cover), but we were not able to distinguish groups of lakes.

Our modeling approach required temporal data and was designed to make predictions about water quality data with temporally resolved climate data, but other spatial variables are also likely important. Unpublished analyses of our study lakes suggest much stronger spatial variation than temporal variation in lake and landscape variables (e.g., climate, land use, and atmospheric deposition). Despite documented increases in temperature and precipitation in this region, most lakes have remained relatively unchanged in nutrients or chlorophyll in recent decades (Oliver et al., 2017), and temporal patterns in water clarity are less predictable than spatial patterns (Lottig et al., 2017), suggesting that spatial relationships are stronger than temporal ones.

The effects of climate on lakes can interact with eutrophication (Moss et al., 2011), and climate effects are likely to vary depending on whether a lake is influenced by other stressors (Berthoin et al., 2014). Agricultural land use, and consequently nutrient loading, varies widely in our study area, but the mostly

agricultural region has a relatively small gradient in climate conditions. The error around our predictions was highest in lakes with high nutrients and chlorophyll, suggesting that the relationship between climate and water quality is muted in eutrophic lakes or lakes in agricultural landscapes. Other characteristics could also mediate the strength of the relationship between climate and water quality; for example, deep waters of small lakes have shown muted responses to warming because of lower physical mixing (Winslow et al., 2015). Because other variables are known to moderate the effects of climate change on lakes, and the effects of climate are expected to operate in a relatively homogenous way at broad spatial scales (Gavin et al., 2018; Lapierre et al., 2018), we hypothesize that the effects of climate will be most clearly expressed in regions with little temporal change in land use and land cover. We were not able to test these hypotheses in the current study, but future work should test these ideas more formally.

The climate metrics we examined in this study at annual, seasonal, and monthly time scales are effective in characterizing climate at coarser temporal resolutions but likely did not capture extreme temperature or precipitation events that can last on the order of days. In the future, the frequency of extreme climate events is likely to increase in our study region (Hayhoe et al., 2010; Horton et al., 2014; Pryor et al., 2014) and can have strong effects on lake carbon and nutrients (Carpenter et al., 2015; Jennings et al., 2012; Strock et al., 2016). We were able to include several indices intended to capture anomalies, but they are calculated at relatively coarse spatial and temporal scales, so may be unlikely to reflect the signal of some extreme events. Event-focused climate metrics may capture crucial periods of carbon and nutrient loading that disproportionately affect lake ecosystems (Havens et al., 2016; Loecke et al., 2017; Zwart et al., 2016). Extreme events are rare and require longer-term and higher-frequency records to capture their effects; additionally, these effects are harder to forecast and validate in the future and remain one of the greater uncertainties of climate change (Palmer & Ralsanen, 2002).

#### 4.2. Incorporating Spatial Autocorrelation and Covariance to Improve Predictions at Macroscales

Modeling lakes in the climate system can be challenging (MacKay et al., 2009), and previous studies examining spatial variation in the effect of climate drivers on lake characteristics have primarily used clustering or geospatial interpolation techniques (Jensen et al., 2007; O'Reilly et al., 2015). The results of such studies form the basis of existing predictions regarding potential change in lake characteristics with continued climate change (Jensen et al., 2007; Straile, 2002). Our approach makes predictions at macroscales by incorporating two important factors that are common in many ecological systems. We demonstrate that accounting for covariance among response variables in combination with spatial autocorrelation can produce accurate and precise predictions of lake water quality with climate metrics, even in lakes where data availability is limited.

Both climate metrics (e.g., mean annual temperature and precipitation) and lake ecosystem data (nutrients, productivity, and water clarity) exhibit high to moderate levels of spatial autocorrelation (Lapierre et al., 2018), making them well suited for our modeling approach. Many nutrient and productivity response variables are inherently correlated with one another, partly as a result of coupled biogeochemical cycles (Schlesinger et al., 2011). In addition, others have argued that the connections between biogeochemical cycles of different elements make it important to consider multiple nutrients when evaluating the effects of climate on ecosystems (Whitehead & Crossman, 2012). We found that leveraging information about correlated water quality variables improved precision and accuracy compared to traditional approaches that fail to account for covariance, which is consistent with related studies (Wagner & Schliep, 2018). Correlations among predictors, responses, and space are often viewed as problems in the data that limit interpretation and must be dealt with prior to analysis, yet we describe a model that uses those characteristics to inform the relationship between climate and lake ecosystems outside the spatial and temporal scope of individual lake data sets.

Another benefit of our approach relates to incomplete data, which is a common problem in ecology at macroscales. Climate studies often require multiple decades of relatively complete time series data to explain variation or make predictions. Our modeling framework overcame this constraint because it leveraged data from many lakes over broad gradients in climate and other ecological context characteristics to make predictions with incomplete data. Since very few lakes have long-term data on ecosystem properties, this approach allowed us to examine the relationship between climate and ecosystems in systems that would not be compatible with modeling approaches that require complete time series spanning many years. This method

identified large variation in the magnitude and direction of climate effects on lake ecosystems, suggesting that extrapolation from a small number of well-studied systems is challenging, and it also provides a useful methodological approach to apply in other macroscale research questions where data are incomplete over space or time.

## 5. Conclusions

Our results demonstrate large variation across lakes in how climate predicts ecosystem properties, including the identity of important climate predictors and the magnitude and direction of their effect on water quality response variables. Despite this variation, summer temperature and winter precipitation are consistently important predictors across thousands of lakes, and summer temperature has, on average, positive effects on nutrients and productivity, while winter precipitation effects are, on average, slightly negative with both positive and negative effects across lakes. Consequently, we predict that increases in temperature are likely to increase lake nutrients and productivity, while potential changes in precipitation are likely to vary in direction across lakes. Our results emphasize the context dependency of climate effects on lakes and the need to consider individual ecosystems separately to make accurate predictions. Further, water quality is related to numerous ecosystem services, and understanding predicted changes in water quality across lakes will lead to more effective mitigation of climate change at local to global scales.

## Acknowledgments

We thank the CSI-Limnology group, particularly Caren Scott, for providing ideas and feedback on this analysis. This research was funded by the Macrosystems Biology Program in the Emerging Frontiers Division of the Biological Sciences Directorate at the U.S. National Science Foundation (EF-1065786, EF-1065818, EF-1638679, EF-1638554, and EF-1638539) and a U.S. National Science Foundation Postdoctoral Research Fellowship in Biology to S. M. Collins (DBI-1401954). P. A. Soranno was also supported by the USDA National Institute of Food and Agriculture, Hatch project 1013544. Limnological and geographic data for LAGOS-NE are published as a data paper (Soranno et al., 2017), and monthly, seasonal, and annual climate data to accompany the LAGOS-NE data set are published with the Environmental Data Initiative (Collins et al., 2018). Code and data for the MTL method are available on Github (<https://github.com/shuaiyuan-msu/csi-mtmmr>) as are code and data to prepare model input and process results (<https://github.com/smc322/LAGOSclimate>).

## References

- Adrian, R., O'Reilly, C. M., Zagarese, H., Baines, S. B., Hessen, D. O., Keller, W., et al. (2009). Lakes as sentinels of climate change. *Limnology and Oceanography*, 54(6), 2283–2297. [https://doi.org/10.4319/lo.2009.54.6\\_part\\_2.2283](https://doi.org/10.4319/lo.2009.54.6_part_2.2283)
- Arbuckle, K. E., & Downing, J. A. (2001). The influence of watershed land use on lake N: P in a predominantly agricultural landscape. *Limnology and Oceanography*, 46(4), 970–975. <https://doi.org/10.4319/lo.2001.46.4.0970>
- Arvola, L., Einola, E., & Järvinen, M. (2015). Landscape properties and precipitation as determinants for high summer nitrogen load from boreal catchments. *Landscape Ecology*, 30(3), 429–442. <https://doi.org/10.1007/s10980-015-0166-8>
- Bertahas, I., Dimitriou, E., Karaouzas, I., Laschou, S., & Zacharias, I. (2006). Climate change and agricultural pollution effects on the trophic status of a Mediterranean lake. *Acta Hydrochimica et Hydrobiologica*, 34(4), 349–359. <https://doi.org/10.1002/aheh.200500637>
- Berthon, V., Alric, B., Rimet, F., & Perga, M.-E. (2014). Sensitivity and responses of diatoms to climate warming in lakes heavily influenced by humans. *Freshwater Biology*, 59(8), 1755–1767. <https://doi.org/10.1111/fwb.12380>
- Bivand, R. S., Pebesma, E., & Gomez-Rubio, V. (2013). *Applied spatial data analysis with R* (2nd ed.). New York: Springer. <http://www.asdar-book.org/>, <https://doi.org/10.1007/978-1-4614-7618-4>
- Byun, K., & Hamlet, A. F. (2018). Projected changes in future climate over the Midwest and Great Lakes region using downscaled CMIP5 ensembles. *International Journal of Climatology*, 38, e531–e553. <https://doi.org/10.1002/joc.5388>
- Cardille, J. A., Carpenter, S. R., Coe, M. T., Foley, J. A., Hanson, P. C., Turner, M. G., & Vano, J. A. (2007). Carbon and water cycling in lake-rich landscapes: Landscape connections, lake hydrology, and biogeochemistry. *Journal of Geophysical Research*, 112, G02031. <https://doi.org/10.1029/2006JG000200>
- Carpenter, S. R., Booth, E. G., Kucharik, C. J., & Lathrop, R. C. (2015). Extreme daily loads: Role in annual phosphorus input to a north temperate lake. *Aquatic Sciences*, 77(1), 71–79. <https://doi.org/10.1007/s00027-014-0364-5>
- Caruana, R. (1997). Multi-task learning. *Machine Learning*, 28(1), 41–75. <https://doi.org/10.1023/A:1007379606734>
- Cheruvilil, K. S., Soranno, P. A., Webster, K. E., & Bremigan, M. T. (2013). Multi-scaled drivers of ecosystem state: Quantifying the importance of the regional spatial scale. *Ecological Applications*, 23(7), 1603–1618. <https://doi.org/10.1890/12-1872.1>
- Collins, S. M., McCullough, I. M., Oliver, S. K., & Skaff, N. K. (2018). LAGOS-NE annual, seasonal, and monthly climate data for lakes and watersheds in a 17-state region of the U.S. Environmental Data Initiative. <https://doi.org/10.6073/pasta/4abe86a2c00dc9a628924aa149d7bf34>
- Collins, S. M., Oliver, S. K., Lapierre, J.-F., Stanley, E. H., Jones, J. R., Wagner, T., & Soranno, P. A. (2017). Lake nutrient stoichiometry is less predictable than nutrient concentrations at regional and sub-continental scales. *Ecological Applications*, 27(5), 1529–1540. <https://doi.org/10.1002/eap.1545>
- Crossman, J., Futter, M. N., Oni, S. K., Whitehead, P. G., Jin, L., Butterfield, D., et al. (2013). Impacts of climate change on hydrology and water quality: Future proofing management strategies in the Lake Simcoe watershed, Canada. *Journal of Great Lakes Research*, 39(1), 19–32. <https://doi.org/10.1016/j.jglr.2012.11.003>
- Crowley, K. F., McNeil, B. E., Lovett, G. M., Canham, C. D., Driscoll, C. T., Rustad, L. E., et al. (2012). Do nutrient limitation patterns shift from nitrogen toward phosphorus with increasing nitrogen deposition across the northeastern United States? *Ecosystems*, 15(6), 940–957. <https://doi.org/10.1007/s10021-012-9550-2>
- Dillon, P. J., & Rigler, F. H. (1974). The phosphorus-chlorophyll relationship in lakes. *Limnology and Oceanography*, 19(5), 767–773. <https://doi.org/10.4319/lo.1974.19.5.0767>
- Downing, J. A., & McCauley, E. (1992). The nitrogen:phosphorus relationship in lakes. *Limnology and Oceanography*, 37(5), 936–945. <https://doi.org/10.4319/lo.1992.37.5.0936>
- Fergus, C. E., Lapierre, J.-F., Oliver, S. K., Skaff, N. K., Cheruvilil, K. S., Webster, K., et al. (2017). The freshwater landscape: Lake, wetland, and stream abundance and connectivity at macroscales. *Ecosphere*, 8(8), e01911–e01923. <https://doi.org/10.1002/ecs2.1911>
- Filstrup, C. T., Wagner, T., Oliver, S. K., Stow, C. A., Webster, K. E., Stanley, E. H., & Downing, J. A. (2017). Evidence for regional nitrogen stress on chlorophyll *a* in lakes across large landscape and climate gradients. *Limnology and Oceanography*, 46, 970–916.
- Fraterrigo, J. M., & Downing, J. A. (2008). The influence of land use on lake nutrients varies with watershed transport capacity. *Ecosystems*, 11(7), 1021–1034. <https://doi.org/10.1007/s10021-008-9176-6>



- Gavin, A. L., Nelson, S. J., Klemmer, A. J., Fernandez, I. J., Strock, K. E., & McDowell, W. H. (2018). Acidification and climate linkages to increased dissolved organic carbon in high-elevation lakes. *Water Resources Research*, 54, 5376–5393. <https://doi.org/10.1029/2017WR020963>
- Grimm, N. B., Chapin, F. S. III, Bierwagen, B., Gonzalez, P., Groffman, P. M., Luo, Y., et al. (2013). The impacts of climate change on ecosystem structure and function. *Frontiers in Ecology and the Environment*, 11(9), 474–482. <https://doi.org/10.1890/120282>
- Haldna, M., Milius, A., Laugaste, R., & Kangur, K. (2008). Nutrients and phytoplankton in Lake Peipsi during two periods that differed in water level and temperature. *Hydrobiologia*, 599(1), 3–11. <https://doi.org/10.1007/s10750-007-9208-9>
- Hampton, S. E., Galloway, A. W. E., Powers, S. M., Ozersky, T., Woo, K. H., Batt, R. D., et al. (2017). Ecology under lake ice. *Ecology Letters*, 20(1), 98–111. <https://doi.org/10.1111/ele.12699>
- Havens, K., Paerl, H., Philips, E., Zhu, M., Beaver, J., & Srifa, A. (2016). Extreme weather events and climate variability provide a lens to how shallow lakes may respond to climate change. *Water*, 8(6), 229–218. <https://doi.org/10.3390/w8060229>
- Hayhoe, K., VanDorn, J., Croley, T. II, Schlegal, N., & Wuebbles, D. (2010). Regional climate change projections for Chicago and the US Great Lakes. *Journal of Great Lakes Research*, 36, 7–21. <https://doi.org/10.1016/j.jglr.2010.03.012>
- Hoegh-Guldberg, O., & Bruno, J. F. (2010). The impact of climate change on the world's marine ecosystems. *Science*, 328(5985), 1523–1528. <https://doi.org/10.1126/science.1189930>
- Horton, R., Yohe, G., Easterling, W., Kates, R., Ruth, M., Sussman, E., et al. (2014). Ch. 16: Northeast. Climate change impacts in the United States: The third National Climate Assessment. J. M. Melillo, Terese (T.C.) Richmond, and G. W. Yohe, Eds., U.S. Global Change Research Program, 16–1-nn.
- Jankowski, T., Livingstone, D. M., Buhner, H., Forster, R., & Niederhauser, P. (2006). Consequences of the 2003 European heat wave for lake temperature profiles, thermal stability, and hypolimnetic oxygen depletion: Implications for a warmer world. *Limnology and Oceanography*, 51(2), 815–819. <https://doi.org/10.4319/lo.2006.51.2.0815>
- Jennings, E., Jones, S., Arvola, L., Staehr, P. A., Gaiser, E., Jones, I. D., et al. (2012). Effects of weather-related episodic events in lakes: An analysis based on high-frequency data. *Freshwater Biology*, 57(3), 589–601. <https://doi.org/10.1111/j.1365-2427.2011.02729.x>
- Jensen, O. P., Benson, B. J., Magnuson, J. J., Card, V. M., Futter, M. N., Soranno, P. A., & Stewart, K. M. (2007). Spatial analysis of ice phenology trends across the Laurentian Great Lakes region during a recent warming period. *Limnology and Oceanography*, 52(5), 2013–2026. <https://doi.org/10.4319/lo.2007.52.5.2013>
- Jeppesen, E., Meerhoff, M., Davidson, T. A., Trolle, D., Søndergaard, M., Lauridsen, T. L., et al. (2014). Climate change impacts on lakes: An integrated ecological perspective based on a multi-faceted approach, with special focus on shallow lakes. *Journal of Limnology*, 73, 1–24.
- Karl, T. R., & Knight, R. W. (1985). *Atlas of monthly Palmer hydrological drought indices (1931-1983) for the contiguous United States, Historical Climatology Series* (pp. 3–7). Asheville, NC: National Climatic Data Center.
- Kraemer, B. M., Anneville, O., Chandra, S., Dix, M., Kuusisto, E., Livingstone, D. M., et al. (2015). Morphometry and average temperature affect lake stratification responses to climate change. *Geophysical Research Letters*, 42, 4981–4988. <https://doi.org/10.1002/2015GL064097>
- Kraemer, B. M., Chandra, S., Dell, A. I., Dix, M., Kuusisto, E., Livingstone, D. M., et al. (2017). Global patterns in lake ecosystem responses to warming based on the temperature dependence of metabolism. *Global Change Biology*, 23(5), 1881–1890. <https://doi.org/10.1111/gcb.13459>
- Kraemer, B. M., Mehner, T., & Adrian, R. (2017). Reconciling the opposing effects of warming on phytoplankton biomass in 188 large lakes. *Scientific Reports*, 7(1), 10762. <https://doi.org/10.1038/s41598-017-11167-3>
- Lapierre, J.-F., Collins, S. M., Seekell, D. A., Cheruvilil, K. S., Tan, P.-N., Skaff, N. K., et al. (2018). Similarity in spatial structure constrains ecosystem relationships: Building a macroscale understanding of lakes. *Global Ecology and Biogeography*, 27(10), 1251–1263. <https://doi.org/10.1111/geb.12781>
- Lapierre, J.-F., Seekell, D. A., & del Giorgio, P. A. (2015). Climate and landscape influence on indicators of lake carbon cycling through spatial patterns in dissolved organic carbon. *Global Change Biology*, 21(12), 4425–4435. <https://doi.org/10.1111/gcb.13031>
- Leppäranta, M. (2015). *Freezing of lakes and the evolution of their ice cover*. Berlin, Heidelberg: Springer. <https://doi.org/10.1007/978-3-642-29081-7>
- Loecke, T. D., Burgin, A. J., Riveros-Iregui, D. A., Ward, A. S., Thomas, S. A., Davis, C. A., & Clair, M. A. S. (2017). Weather whiplash in agricultural regions drives deterioration of water quality. *Biogeochemistry*, 133(1), 7–15. <https://doi.org/10.1007/s10533-017-0315-z>
- Lottig, N. R., Tan, P.-N., Wagner, T., Cheruvilil, K. S., Soranno, P. A., Stanley, E. H., et al. (2017). Macroscale patterns of synchrony identify complex relationships among spatial and temporal ecosystem drivers. *Ecosphere*, 8(12), e02024–e02011. <https://doi.org/10.1002/ecs2.2024>
- MacKay, M. D., Neale, P. J., Arp, C. D., de Senerpont Domis, L. N., Fang, X., Gal, G., et al. (2009). Modeling lakes and reservoirs in the climate system. *Limnology and Oceanography*, 54(6part2), 2315–2329. [https://doi.org/10.4319/lo.2009.54.6\\_part\\_2.2315](https://doi.org/10.4319/lo.2009.54.6_part_2.2315)
- Magnuson, J. J., Robertson, D. M., Benson, B. J., Wynne, R. H., Livingstone, D. M., Arai, T., et al. (2000). Historical trends in lake and river ice cover in the Northern Hemisphere. *Science*, 289(5485), 1743–1746. <https://doi.org/10.1126/science.289.5485.1743>
- Marce, R., Rodriguez-Arias, M. À., Garcia, J. C., & Armengol, J. (2010). El Niño Southern Oscillation and climate trends impact reservoir water quality. *Global Change Biology*, 16(10), 2857–2865. <https://doi.org/10.1111/j.1365-2486.2010.02163.x>
- Moss, B. (2012). Cogs in the endless machine: Lakes, climate change and nutrient cycles—A review. *Science of the Total Environment*, 434, 130–142. <https://doi.org/10.1016/j.scitotenv.2011.07.069>
- Moss, B., Kosten, S., Meerhoff, M., Battarbee, R. W., Jeppesen, E., Mazzeo, N., et al. (2011). Allied attack: Climate change and eutrophication. *Inland Waters*, 1(2), 101–105. <https://doi.org/10.5268/IW-1.2.359>
- Oliver, S. K., Collins, S. M., Soranno, P. A., Wagner, T., Stanley, E. H., Jones, J. R., et al. (2017). Unexpected stasis in a changing world: Lake nutrient and chlorophyll trends since 1990. *Global Change Biology*, 23(12), 5455–5467. <https://doi.org/10.1111/gcb.13810>
- O'Reilly, C. M., Sharma, S., Gray, D. K., Hampton, S. E., Read, J. S., Rowley, R. J., et al. (2015). Rapid and highly variable warming of lake surface waters around the globe. *Geophysical Research Letters*, 42, 10,773–10,781. <https://doi.org/10.1002/2015GL066235>
- Palmer, T. N., & Ralsanen, J. (2002). Quantifying the risk of extreme seasonal precipitation events in a changing climate. *Nature*, 415(6871), 512–514. <https://doi.org/10.1038/415512a>
- Parnesan, C. (2006). Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology, Evolution, and Systematics*, 37(1), 637–669. <https://doi.org/10.1146/annurev.ecolsys.37.091305.110100>
- Pebesma, E. J., & Bivand, R. S. (2005). Classes and methods for spatial data in R. *R News*, 5(2), 9–13. <https://cran.r-project.org/doc/Rnews/>
- Pierson, D. C., Samal, N. R., Owens, E. M., Schneiderman, E. M., & Zion, M. S. (2013). Changes in the timing of snowmelt and the seasonality of nutrient loading: Can models simulate the impacts on freshwater trophic status? *Hydrological Processes*, 27(21), 3083–3093. <https://doi.org/10.1002/hyp.9894>



- Preston, D. L., Caine, N., McKnight, D. M., Williams, M. W., Hell, K., Miller, M. P., et al. (2016). Climate regulates alpine lake ice cover phenology and aquatic ecosystem structure. *Geophysical Research Letters*, 43, 5353–5360. <https://doi.org/10.1002/2016GL069036>
- Pryor, S., Scavia, D., Downer, C., Gaden, M., Iverson, L., Nordstrom, R., et al. (2014). Climate change impacts in the United States: The third National Climate Assessment. U.S. Global Change Research Program (pp. 418–440). Washington, DC. <https://doi.org/10.7930/J0J1012N.On>
- R Core Team (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Read, E. K., Patil, V. P., Oliver, S. K., Hetherington, A. L., Brenttrup, J. A., Zwart, J. A., et al. (2015). The importance of lake-specific characteristics for water quality across the continental United States. *Ecological Applications*, 25, 943–955.
- Rose, K. C., Greb, S. R., Diebel, M., & Turner, M. G. (2017). Annual precipitation regulates spatial and temporal drivers of lake water clarity. *Ecological Applications*, 27(2), 632–643. <https://doi.org/10.1002/eap.1471>
- Schlesinger, W. H., Cole, J. J., Finzi, A. C., & Holland, E. A. (2011). Introduction to coupled biogeochemical cycles. *Frontiers in Ecology and the Environment*, 9(1), 5–8. <https://doi.org/10.1890/090235>
- Soranno, P. A., Bacon, L. C., Beauchene, M., Bednar, K. E., Bissell, E. G., Boudreau, C. K., et al. (2017). LAGOS-NE: A multi-scaled geospatial and temporal database of lake ecological context and water quality for thousands of US lakes. *GigaScience*, 6(12), 1–22. <https://doi.org/10.1093/gigascience/gix101>
- Soranno, P. A., Bissell, E. G., Cheruvilil, K. S., Christel, S. T., Collins, S. M., Fergus, C. E., et al. (2015). Building a multi-scaled geospatial temporal ecology database from disparate data sources: Fostering open science and data reuse. *GigaScience*, 4, 1–15.
- Soranno, P. A., & Cheruvilil, K. S. (2017). LAGOS-NE-LIMNO v1.087.1: A module for LAGOS-NE, a multi-scaled geospatial and temporal database of lake ecological context and water quality for thousands of U.S. lakes: 1925–2013. Environmental Data Initiative. <https://doi.org/10.6073/pasta/b1b93ccf3354a7471b93ecca484d506>
- Straile, D. (2002). North Atlantic Oscillation synchronizes food-web interactions in central European lakes. *Proceedings of the Royal Society B: Biological Sciences*, 269(1489), 391–395. <https://doi.org/10.1098/rspb.2001.1907>
- Strock, K. E., Saros, J. E., Nelson, S. J., Birkel, S. D., Kahl, J. S., & McDowell, W. H. (2016). Extreme weather years drive episodic changes in lake chemistry: Implications for recovery from sulfate deposition and long-term trends in dissolved organic carbon. *Biogeochemistry*, 127(2–3), 353–365. <https://doi.org/10.1007/s10533-016-0185-9>
- Wagner, T., & Schliep, E. M. (2018). Combining nutrient, productivity, and landscape-based regressions improves predictions of lake nutrients and provides insight into nutrient coupling at macroscales. *Limnology and Oceanography*, 63(6), 2372–2383. <https://doi.org/10.1002/lno.10944>
- Walther, G.-R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T. J. C., et al. (2002). Ecological responses to recent climate change. *Nature*, 416(6879), 389–395. <https://doi.org/10.1038/416389a>
- Whitehead, P. G., & Crossman, J. (2012). Macronutrient cycles and climate change: Key science areas and an international perspective. *Science of the Total Environment*, 434, 13–17. <https://doi.org/10.1016/j.scitotenv.2011.08.046>
- Whitehead, P. G., Wilby, R. L., Battarbee, R. W., Kernan, M., & Wade, A. J. (2009). A review of the potential impacts of climate change on surface water quality. *Hydrological Sciences Journal*, 54(1), 101–123. <https://doi.org/10.1623/hysj.54.1.101>
- Williamson, C. E., Saros, J. E., Vincent, W. F., & Smol, J. P. (2009). Lakes and reservoirs as sentinels, integrators, and regulators of climate change. *Limnology and Oceanography*, 54(6part2), 2273–2282. [https://doi.org/10.4319/lo.2009.54.6\\_part\\_2.2273](https://doi.org/10.4319/lo.2009.54.6_part_2.2273)
- Winslow, L. A., Read, J. S., Hansen, G. J. A., & Hanson, P. C. (2015). Small lakes show muted climate change signal in deepwater temperatures. *Geophysical Research Letters*, 42, 355–361. <https://doi.org/10.1002/2014GL062325>
- Zwart, J. A., Sebestyen, S. D., Solomon, C. T., & Jones, S. E. (2016). The influence of hydrologic residence time on lake carbon cycling dynamics following extreme precipitation events. *Ecosystems*, 20, 1000–1014.