Science of the Total Environment xxx (xxxx) xxx

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Quantifying uncertainty cascading from climate, watershed, and lake models in harmful algal bloom predictions

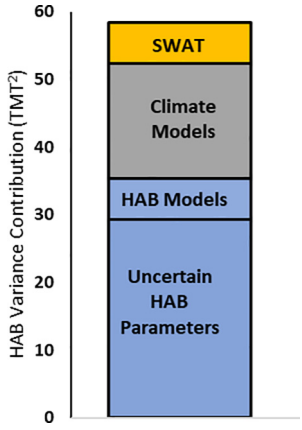
Donald Scavia [a,*], Yu-Chen Wang [a], Daniel R. Obenour [b], Anna Apostel [c], Samantha J. Basile [d], Margaret M. Kalcic [c], Christine J. Kirchhoff [e], Lorrayne Miralha [f], Rebecca L. Muenich [f], Allison L. Steiner [g]

[a] School for Environment and Sustainability, University of Michigan, Ann Arbor, MI 48104, USA
[b] Department of Civil, Construction & Environmental Engineering, NC State University, Raleigh, NC 27695, USA
[c] Department of Food, Agricultural and Biological Engineering and Translational Data Analytics Institute, The Ohio State University, Columbus, OH 43210, USA
[d] National Climate Assessment, ICF, 1725 I St NW, Washington, DC 20006, USA
[e] Department of Civil and Environmental Engineering, University of Connecticut, Storrs, CT 06269, USA
[f] School of Sustainable Engineering and the Built Environment, Arizona State University, Tempe, AZ 85281, USA
[g] Climate and Space Sciences and Engineering, University of Michigan, Ann Arbor, MI 48109, USA

HIGHLIGHTS

- Forecasting harmful algal blooms (HAB) in an uncertain future is important
- Quantifying sources of HAB forecast uncertainty is important
- Integrated climate, watershed, and HAB models to forecast future HABs
- Climate-related changes in nutrient loads will decrease future HABs
- Uncertainty from all model categories are important

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 7 August 2020
Received in revised form 22 October 2020
Accepted 29 October 2020
Available online xxxx

Dr. Damia Barcelo

Keywords:
Integrating models
Error analysis
Monte Carlo
Forecasting

ABSTRACT

In response to increased harmful algal blooms (HABs), hypoxia, and nearshore algae growth in Lake Erie, the United States and Canada agreed to phosphorus load reduction targets. While the load targets were guided by an ensemble of models, none of them considered the effects of climate change. Some watershed models developed to guide load reduction strategies have simulated climate effects, but without extending the resulting loads or their uncertainties to HAB projections. In this study, we integrated an ensemble of four climate models, three watershed models, and four HAB models. Nutrient loads and HAB predictions were generated for historical (1985–1999), current (2002–2017), and mid-21st-century (2051–2065) periods. For the current and historical periods, modeled loads and HABs are comparable to observations but exhibit less interannual variability. Our results show that climate impacts on watershed processes are likely to lead to reductions in future loading, assuming land use and watershed management practices are unchanged. This reduction in load should help reduce the magnitude of future HABs, although increases in lake temperature could mitigate that decrease. Using Monte-Carlo analysis to attribute sources of uncertainty from this cascade of models, we show that the uncertainty associated with each model is significant, and that improvements in all three are needed to build confidence in future projections.

* Corresponding author.
E-mail address: scavia@umich.edu (D. Scavia).

https://doi.org/10.1016/j.scitotenv.2020.143487
0048-9697/

1. Introduction

In response to increased harmful algal blooms (HABs), hypoxia, and nearshore algae growth in Lake Erie (Scavia et al., 2014), the United States and Canada agreed to a 40% reduction in annual total phosphorus (TP) load from 2008 levels (GLWQA, 2016). The Maumee River, the primary source of the nutrients driving Lake Erie HABs (Scavia et al., 2016), has an additional target to reduce the March–July TP and dissolved reactive phosphorus (DRP) loads by 40% in 9 out of 10 years. These targets were guided by an ensemble of lake models (Scavia et al., 2016), including three that focused on HABs (Bertani et al., 2016; Stumpf et al., 2016, and Verhamme et al., 2016). However, none of the target-setting planning scenarios accounted for climate change. And, while an ensemble of Maumee watershed models showed that large-scale implementation of multiple agricultural conservation practices could achieve the 40% reduction on average (Scavia et al., 2017a; Martin et al., 2019), these simulations also did not account for climate change.

Most HAB models used to explore climate effects (e.g., Ralson and Moore, 2020, Glibert et al., 2010) focus on temperature because of its importance for HAB development. Toxin-forming cyanobacteria taxa (primarily *Microcystis* in Lake Erie) are favored by higher temperatures (Paerl and Huisman, 2008) and a warmer climate would lead warmer waters, enhanced stratification, and earlier and likely more intense blooms (Gobler, 2020; Del Giudice et al., in press). However, how climate-influenced changes in nutrient loads and how their uncertainties propagate to HAB development has received less attention. For example, Salmaso et al. (2018) showed that warming in deep lakes can lead to oligotrophication in spite of high nutrient loads.

A number of studies explored potential impacts of climate on nutrient loads to Lake Erie; some based on adjusted historical climate data (e.g., Daloğlu et al., 2012; Bosch et al., 2014; Jarvie et al., 2017) and others based on output from climate models (e.g., Johnson et al., 2015; Verma et al., 2015; Culbertson et al., 2016; Kalcic et al., 2019; Kujawa et al., 2020). Some of these studies suggest P loads will increase in response to climate change, while others suggest a decrease. The diversity of projected climate impacts on P loads could be attributed in part to the range of time scales used in these regional analyses. For example, Johnson et al. (2015), Bosch et al. (2014), Daloğlu et al. (2012), Jarvie et al. (2017), and Kujawa et al. (2020) evaluated responses at annual scales, whereas Kalcic et al. (2019) analyzed spring (March–July) loads and Culbertson et al. (2016) and Verma et al. (2015) analyzed monthly loads. This diversity of results could also be influenced by the climate and watershed models chosen for analysis, whether or not model outputs have been bias corrected, or how modeling techniques were linked. Moreover, to our knowledge, no modeling studies explicitly propagated load changes and their uncertainties onto HAB projections.

In addition to the conflicting load projections, uncertainties arise from the use of different time frames and forms of P as inputs to HAB models. The timeframe crucial to Lake Erie HABs is approximately March through July (Scavia et al., 2016) because that is a period of intensive flow and nutrient loading prior to the primary cyanobacteria growth period. But, precisely which months and which forms of P are used depends on the HAB model. For example, Stumpf et al. (2016) found the strongest correlation with peak HAB abundance was with the March–July load, Ho and Michalak (2015) found the strongest relationship with the April–July load, whereas Obenour et al. (2014) and Bertani et al. (2016) found the strongest relationships with roughly half of the February load plus the March–June load. To complicate these relationships further, Obenour's relationship was with TP, Ho's was with DRP, and Stumpf and Bertani were with bioavailable P estimated as different combinations of TP and DRP.

To approach this variability in the impacts of load estimation systematically, we recalibrated four existing HAB models with a consistent set of historical P loads and HAB extent estimates (2002–2017), and used them to explore changes in HAB extent among historical (1985–1999), current (2002–2017), and future (2051–2065) time periods. Simulated loads were generated from three watershed models that were driven by either observed climate data or a suite of climate models that have been shown to perform well in the study region (Basile et al., 2017). We first use the modeling results to understand the changes between historical and current periods, and then to explore potential loads and HAB extents under a future climate. Following that, we use Monte-Carlo analysis to quantify and attribute sources of uncertainty from this cascade of models.

2. Materials and methods

2.1. Study area

The 17,000 km^2 Maumee River watershed (Fig. 1a) is characterized by low-sloping to flat topography and heavy, clayey soils with poor natural drainage. Productive agriculture in the Midwest/Great Lakes regions has benefited from widespread installation of subsurface drainage tiles (Evans and Fausey, 2015). Row crop agriculture is dominant (~70%), consisting primarily of rotations of corn, soybean, and winter wheat (Fig. 1b).

The Maumee River flows into the 3000 km^2 western basin of Lake Erie that is relatively shallow, averaging 7.4 m (Millie et al., 2009) with a water residence time of 20–40 days, and subject to intermittent stratification and wind-driven mixing (Wynne et al., 2011). The Detroit and Maumee rivers contribute 41% and 48% of the TP load to the western basin, respectively. While the P load from the Detroit River is comparable to that of the Maumee, its high flow and low P concentration tend to dilute or deflect the Maumee-driven HAB (Michalak et al., 2013).

Seasonal precipitation in the study area has a unimodal pattern with a summer maximum (Basile et al., 2017). For 1980–1999, average annual precipitation was 908.2 mm, with a February minimum (46.7 mm month^{-1}) and July maximum (97.8 mm month^{-1}). While the highest fraction of total annual precipitation occurs in summer, over one-fourth of annual precipitation occurs in spring when evapotranspiration is lower.

2.2. Climate models and bias correction

Climate models (Fig. 1c,d) used to provide historical (1985–1999) and mid-century (2051–2065) inputs to the watershed models were the 4 selected from 18 Coupled Model Intercomparison Project (CMIP5, Taylor et al., 2012) models that minimized seasonal biases in temperature, precipitation, snow fall, and snow cover over the western Lake Erie basin (Miralha et al., In Press): CCSM4, CNRM-CM5 (CNRM), IPSL-CM5A-MR (IPSL), and MPI-ESM-MR (MPI). While these models minimize seasonal bias, the models still over predict precipitation and under predict temperature. As a result, they also over predict DRP and TP loads when used as input to the watershed models. Miralha et al. (in press) tested different climate bias-correction methods and concluded that the Delta, Quantile Delta Mapping (QDM), and N-dimensional probability density function (MBC-N) methods performed best, with MBC-N performing best overall. So, we applied all three methods for our change analyses (Section 2.8). However, to keep the number of climate models (4) similar to the number of watershed (3) and HAB models (4) for the uncertainty attribution assessment (Section 2.9), we only used MBC-N. Additional details on the evaluation and selection of climate models and bias correction methods are in Miralha et al. (in press) and in the Supporting Information.

2.3. Watershed models

The three watershed models used here were previously developed, calibrated, and validated for this watershed using the Soil and Water Assessment Tool (SWAT) models. The Apostel version used by Miralha et al. (in press) was developed at the agricultural field-scale resolution

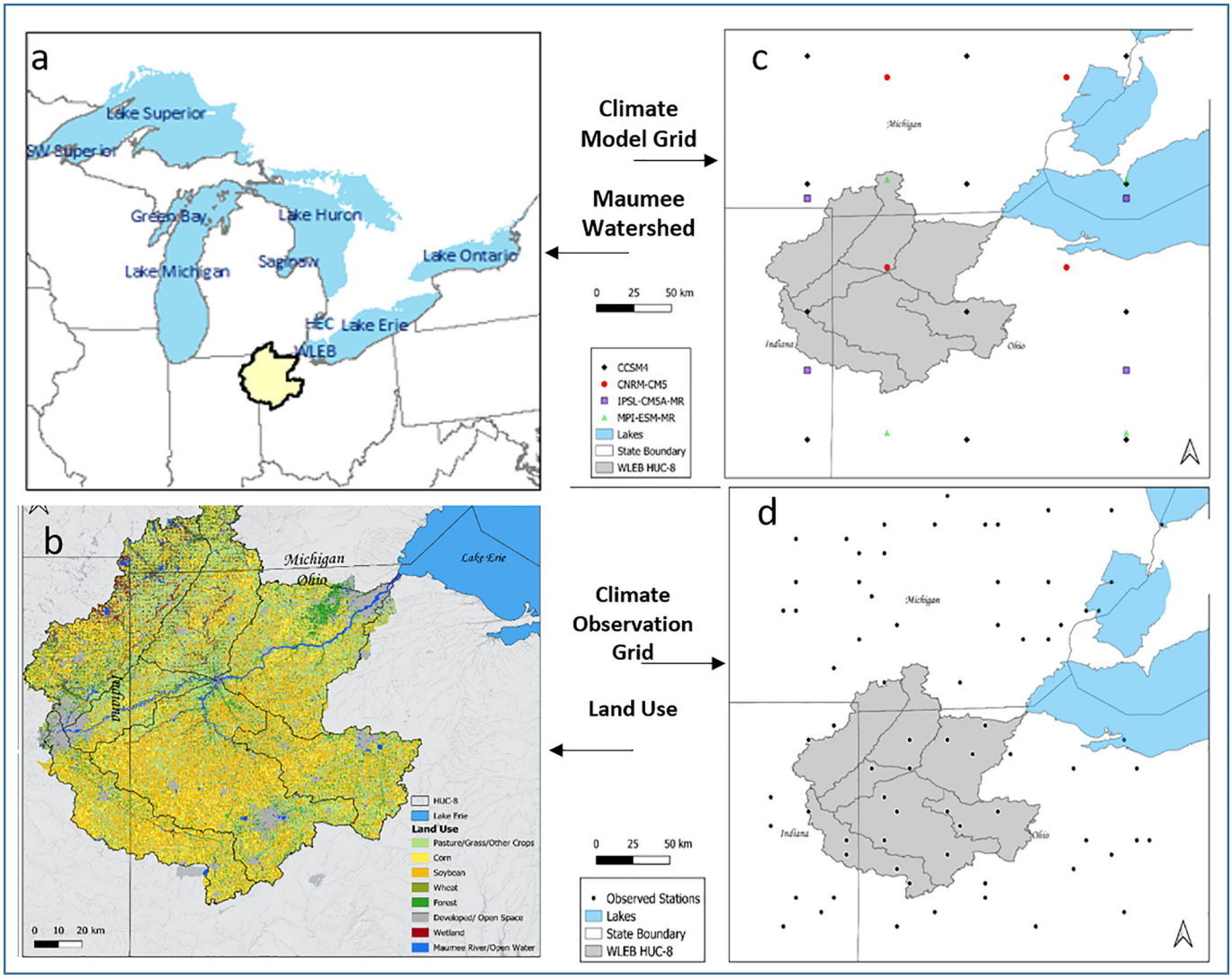


Fig. 1. Study site showing (a) Maumee Watershed, (b) Land use, (c) climate model grid, and (d) climate observation grid for the western Lake Erie region.

(ca. 74 ha). A second SWAT model was developed through a stakeholder-engaged process to study land management and climate change (Kalcic et al., 2016). The third model (Kujawa et al., 2020) is an updated version of the Kalcic model with new inputs, crop management, and parameterizations that was used in an ensemble modeling project. The calibration and validation statistics for all three models were all rated 'good' or better (Apostel et al., in review, Table S1) against standards (Moriasi et al., 2007). The three SWAT models represent a range of process implementations, land management inputs, calibrations, time periods, and parameterizations as described in Table S2 and Supplemental Information. As such, the three models represent alternative plausible representations of the Maumee watershed, and the variability among them provides a measure of watershed model uncertainty.

2.4. Phosphorus load estimates

For HAB model calibrations, daily TP and DRP concentrations were downloaded from Heidelberg University's National Center for Water Quality Research (https://ncwqr.org/monitoring/data/), and daily flow data were downloaded from the United States Geological Survey (USGS, http://www.usgs.gov/water) station USGS-04193500 at Waterville, OH through 2017. Daily loads were calculated and aggregated to monthly loads as described in Obenour et al. (2014).

For simulating HAB responses to historical (1985–1999), current (2002–2017), and mid-century (2051–2065) conditions, median monthly loads for each 15-year time period were produced with the SWAT models driven by climate model outputs as described in Sections 2.2 and 2.3, and then aggregated to the loading periods and forms of P required by the individual HAB models as described in Section 2.6. When observed climate data were used, they were from Menne et al. (2012). Because we are interested in isolating the effect of climate on these loads, the point sources, crop rotations, and other agricultural practices used in SWAT were the same for the historical, current, and future simulations.

To generate future HAB-model-specific P load distributions for the Monte Carlo analysis (Section 2.9), we tested gamma, lognormal, and Weibull distribution fits to SWAT monthly output over the 2051–2065 period for each of the SWAT models, using the R package "fitdistrplus" (version 1.1–1) to fit the gamma distribution by maximum likelihood (Delignette-Muller and Dutang, 2015), and found the gamma model provided the best fit (Figs. S1, S2).

2.5. HAB extent estimates

HAB extent, expressed as dry weight biomass, was obtained from three independent datasets: two satellite-derived (Stumpf et al., 2016; Manning et al., 2019) and one a geostatistical synthesis of in

situ observations (Fang et al., 2019). The Stumpf et al. (2016) cyanobacteria index (CI) is based on processing satellite image spectra specific for cyanobacteria, whereas the Manning biomass estimates are based on chlorophyll-specific spectra (Sayers et al., 2016). These satellite-derived estimates give similar results for relatively high chlorophyll concentrations, but both are constrained to near-surface observations. The Fang estimates are based on in situ chlorophyll observations, which are sparser but allow for full water-column estimates. Because the Stumpf estimates are lake-wide, whereas the other two estimates cover only the western basin, the Stumpf estimates were clipped to the western basin as described by Fang et al. (2019). Because of the differences among these approaches and because we are most interested in inter-annual variability, we scaled the Sayers and Stumpf estimates to the Fang estimates using the ratio of their averages. Results are presented in 1000 metric tons (TMT, hereafter) dry weight based on conversion factors described in Fang et al. (2019).

2.6. HAB models

We used four previously published models that predict HAB extent as a function of Maumee River P loads. The Obenour et al. (2014) model predicts peak western basin bloom biomass (Z, TMT) as a function of spring TP load (W) and a temporal trend representing increased HAB sensitivity to loads:

$$Z = \beta_b + \beta_0 + \beta_w W + \beta_t T + e + m \quad \text{for } (\beta_0 + \beta_w W + \beta_t T) > 0$$
$$Z = \beta_b + e + m \quad \text{for } (\beta_0 + \beta_w W + \beta_t T) < 0$$

where β_i are model coefficients, T is the prediction year relative to 2007 (i.e., year - 2007), e is model prediction error, and m is measurement error across the three extent estimates (see Section 2.5). W is the sum of weighted monthly loads with the weights estimated along with the other model coefficients during calibration (Section 2.7). The model, originally calibrated for 2002–2013 with two independent HAB extent estimates (one from in situ observations and one from remote sensing), resulted in an r^2 of 0.92 and a cross-validation r^2 of 0.84 (Obenour et al., 2014).

The Bertani et al. (2016) model is a modification of Obenour et al. that replaced TP load with bioavailable P load defined as:

$$W = DRP + \theta \times (TP - DRP),$$

with Θ a calibration parameter (Section 2.7). This model was originally calibrated to 2002–2014, resulting in an r^2 of 0.84 and a cross-validation r^2 of 0.70 (Bertani et al., 2016).

The Stumpf et al. (2016) model is based on bioavailable P and calibrated to CI for 2002–2015. Their formulation was:

$$\mathrm{Log}_{10}(Z) = \log_{10}(\beta_b) + \alpha W + e + m$$

where α and β are regression coefficients, and W is March–July bioavailable P, calculated as:

$$W = DRP + (1 - S) \times B \times (TP - DRP),$$

where B is the bioavailable proportion of particulate P (TP-DRP) (0.26, Baker et al., 2014a) and S is the proportion of bioavailable P that settles from the water column (0.7, Baker et al., 2014b). They found that applying March–July loads, with the exclusion of July in years with cold Junes, best described the observations. Stumpf et al. (2016) did not report r^2 or perform cross-validation, but reported the mean absolute deviation and standard deviation of observed HAB extent against the model regression line (excluding 2012 and 2015) of 1.87 and 2.26 in their CI units that range between 1.3 and 29.9.

The Ho and Michalak (2015) model was calibrated to the average of three remote sensing areal extent estimates (km^2) from 1984 to 2015, using the current year's spring (April–July) DRP load (W) and the most recent cumulative 9-year DRP load.

$$Z = \beta_w W + \beta_t DRP_{9-year} + \beta_b + e + m$$

That implementation resulted in an r^2 of 0.75 and a cross-validation r^2 of 0.70.

To separate uncertainty due to HAB measurement error and model prediction error, we used a hierarchical modeling approach to expand the original model formulations and simultaneously calibrate the model to all three sets of HAB extent estimates (Obenour et al., 2014). The three individual extent estimates in each year *i* are modeled with error "m" arising from a probability distribution with mean y_i and standard deviation σ_m. In this formulation, y_i represents the true, unknown HAB extent in year *i* and is itself modeled as arising from a probability distribution with mean equal to the deterministic model prediction and standard deviation σ_e.

2.7. HAB model calibration

The models were calibrated with data from a common period (2002–2017) based on Bayesian inference using a Markov Chain Monte Carlo (MCMC) sampling algorithm implemented within WinBUGS interfaced with the R package, R2WinBUGS (Lunn et al., 2000; R Core Team, 2015; Sturtz et al., 2005). Detailed information on the MCMC algorithm settings, chain convergence evaluation, and parameter prior distributions can be found in Obenour et al. (2014) and Bertani et al. (2016). We tested models using different probability distributions (normal, lognormal, and gamma) to represent "e" and "m", and found the gamma model performed best for the Obenour, Bertani, and Ho models and the normal model performed best for the log-transformed Stumpf model (i.e., residuals follow a lognormal distribution).

2.8. HAB simulations

Each HAB model uses a unique representation of P loading. Ho, Obenour, and Bertani include terms to represent presumed changes in HAB sensitivity to loads. Ho used the cumulative DRP load defined as the sum of the previous nine water years (October–September) through March of the year being forecast. For historical and current period simulations, we used the measured 9-year cumulative loads. For future simulations we used the SWAT-forecast load scaled with the ratio of observed historical mean to the SWAT historical hindcast mean (2011 MT). For comparison, we also used the average observations from the current (2002–2017) period (4111 MT) (Fig. S3).

Obenour and Bertani include a time term based on the prediction year. For calibrating and simulating the current period, we used the actual prediction years. For the historical period, we used $T = -2.5$, which is the average from 2002 to 2007 representing a period of lower sensitivity. Because it is not clear how this sensitivity would change in the future, we used the average from 2008 to 2017 (5.5), the period after the increase in the cumulative load (Fig. S3). For comparison, we also used the mean value for the entire current period (2.5).

Stumpf suggested there was no sensitivity change during their calibration period (2002–2015) but instead accounted for the potential impact of colder springs by excluding July loads for the two years (2003 and 2008) that had a relatively cool June. When we followed that same approach during recalibration, the model's fit was poorer than when all years were included. So, we included all years for calibration and historical simulations. As Stumpf et al. (2016) suggested, we did not make "cold-June" adjustments for future climate simulations.

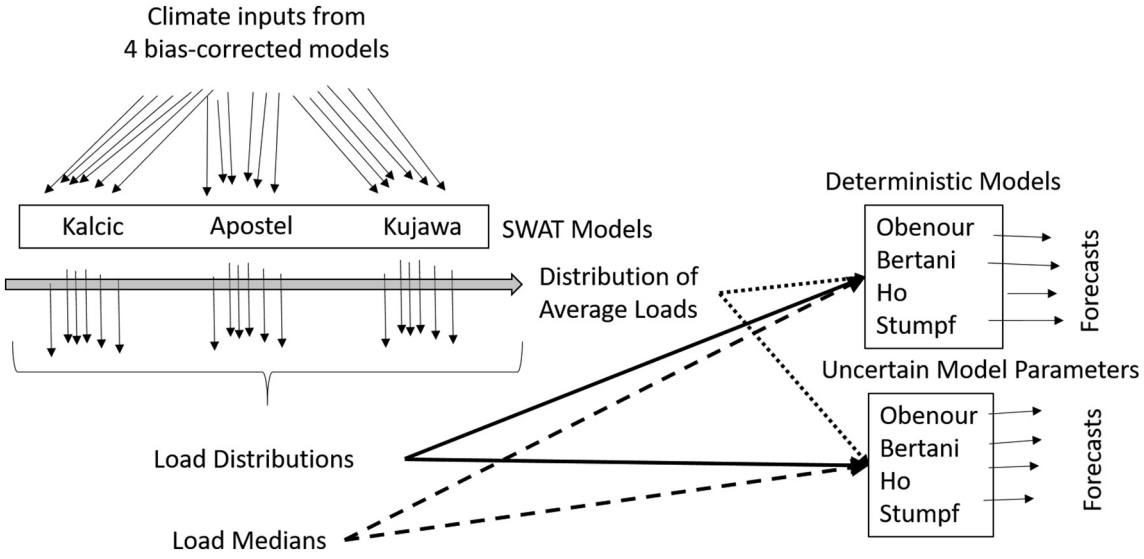


Fig. 2. Monte Carlo runs for estimating and attributing HAB projection uncertainties.

2.9. Uncertainty attribution

We used Monte Carlo simulations to explore sources of uncertainty in the HAB predictions. We compared the variances of future loads and HAB extents from the HAB models run with fixed HAB parameter values and fixed loads to variances from a series of probabilistic model runs (Fig. 2). HAB model parameter uncertainties were represented by sampling from the WinBUGS-generated joint posterior parameter distributions. When exploring model parameter uncertainty, we also include uncertainty in the sensitivity terms (Section 2.8) by drawing values from their uniform distributions with minima and maxima from the current time period. In addition, we performed simulations with and without sampling from the HAB model predictive and measurement error distributions.

For measures of combined climate, SWAT, and HAB model uncertainties, we used the 15 years of future monthly loads from the 3 SWAT models that were driven by the 4 bias-corrected climate models ($n = 15 \times 3 \times 4 = 180$, Fig. 2). To remove watershed model uncertainties, we averaged the outputs from the 3 SWAT models across each year and climate simulation ($n = 4 \times 15 = 60$) before fitting the load distributions. For all probabilistic simulations, loads were drawn from gamma probability distributions fit to the HAB-model-specific P loads calculated from the ensemble of SWAT model runs. When removing both climate- and watershed-model uncertainties, we used median loads from the full uncertainty case.

The means, medians, and variances of predictive HAB distributions converged after about 4000 samples. So, to be conservative, for each Monte Carlo analysis we performed 10,000 simulations drawing from the appropriate load, parameter, and error distributions. To eliminate the occasional extreme and highly unlikely (<1%) HAB extent estimates generated from combinations of values from the tails of these distributions, we analyzed results that fell within the 99% confidence interval of the Monte Carlo outputs.

3. Results

3.1. HAB model recalibration and response curves

The four HAB models were calibrated to the three sets of HAB extent estimates, with resulting r^2 values of 0.63, 0.84, 0.78, and 0.76 for Stumpf, Ho, Bertani, and Obenour models, respectively (Fig. 3). These r^2 values are slightly lower than those reported for the original models; however, if the apparent outlier year 2013 is removed, the values increase to 0.78, 0.90, 0.85, and 0.82, respectively, similar to those reported for the original models.

Load response curves were developed for each HAB model (Fig. 4) based on calibrated parameter estimates (Table S3) and model-specific spring loads (see Section 2.6). The 95% credible intervals for Obenour and Bertani are larger than for Ho because the uncertainty comes from four-five parameters compared to only three for Ho. The

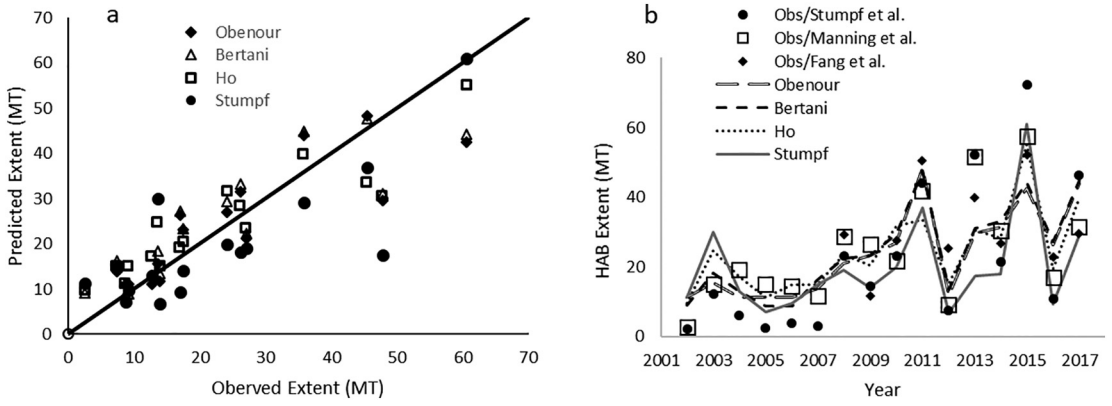


Fig. 3. Predicted vs. observed HAB extents for the four HAB models: (a) each model plotted against the average of the three observed extent estimates; (b) time course of model results and the three sets of observations.

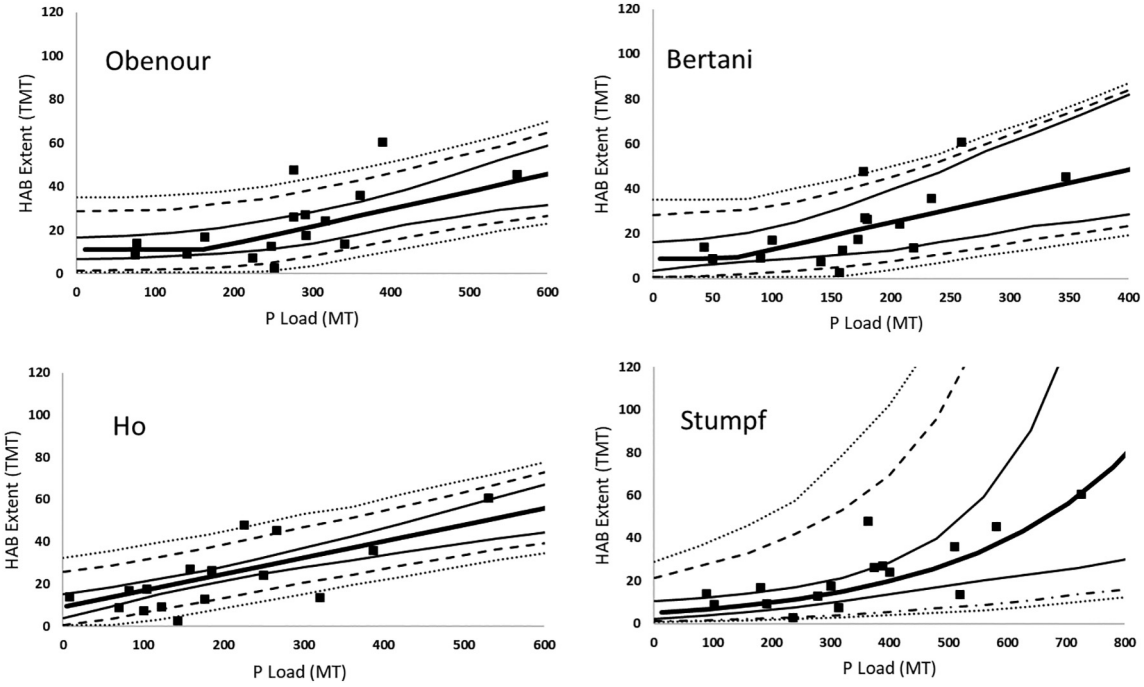


Fig. 4. Response curves (thick solid lines) showing HAB extent as a function of model-specific loads, average extent observations (boxes), 95% credible intervals for parameter uncertainty (thin solid lines), parameter and prediction errors (dashed lines), and parameter, prediction, and measurement error (dotted lines). These response curves use the 2008 time term for the Obenour and Bertani models, and the average 9-year cumulative DRP load for the Ho model.

Stumpf credible intervals grow quite large at higher loads because of the exponential form of the model. In all cases, model prediction error (σ_e) and measurement error (σ_m) were comparable, with ratios between 0.94 and 1.2 (Table S3). These uncertainties characterize inter-annual environmental and observational stochasticity that is not accounted for in the deterministic relationships of the HAB models.

For the simulations discussed below (Figs. 5-7), results are displayed as box plots for each HAB and SWAT model output, representing the 12 climate cases (4 models × 3 bias corrections) when modeled climate was used or the 15 years when observed climate or loads were used for current and historical periods. Measured and modeled loads are averaged HAB-model-specific loads.

3.2. Current simulations

Simulated 2002–2017 March–July TP and DRP loads from the SWAT models were not significantly different from observed loads (Wilcoxon p-values were 0.36–0.9 from each paired group comparison, Fig. S4a, c). The SWAT models, driven by observed climate data, also captured the features of the HAB-model-specific measured loads (Fig. 5a, b). Mean HAB-model-specific SWAT loads based on observed climate (515 MT, Fig. 5a, Table 1) were similar to those calculated from measured loads (537 MT, Fig. 5b). Averaging across HAB models, the HAB extents based on measured loads (23 TMT, Fig. 5d, Table 1) and based on SWAT loads using observed climate inputs (23 TMT, Fig. 5c, Table 1) also align with observed extents (25 TMT).

The mean, median, and interquartile range of HAB extents from Obenour, Bertani, and Ho, when driven with measured loads (Fig. 5d), were consistent with observed extents; Stumpf's were underestimates. HAB models driven with simulated loads varied across SWAT models, but most straddled the observations, with the Stumpf model tending to underestimate overall (Fig. 5c).

3.3. Historical simulations

Simulated 1985–1999 March–July TP loads, driven by climate model output, were not significantly different from measured loads for the Apostel and Kalcic SWAT models (Wilcoxon p-values = 0.92 and 0.27, respectively), but were underestimated by 19% with the Kujawa model ($p = 0.01$, Fig. S4b, d). The three models overestimated measured DRP loads by 77–114%, likely because the models did not adjust for changes in management practices between the historical period and current period for which the models were calibrated.

Historical loads simulated with observed climate (349 MT, Fig. 6b) fell between measured loads (376 MT, Fig. 6c) and those simulated with modeled climate (294 MT, Fig. 6a) (Table 1). While the simulated load for the Obenour model (Fig. 6a,b), which is based only on TP, was below the measured load (Fig. 6c), the simulated loads for the three models that use combinations of TP and DRP were comparable to measured loads. HAB extents were low and similar across all three loading cases and HAB models (10–12 TMT). Historical loads driven with climate model outputs and their associated HAB extents (Fig. 6a, d) are comparable to, but less variable than, those driven by observed climate (Fig. 6b, e) or driven by observed loads (Fig. 6c, f).

3.4. Future projections

Simulated future P loads based on climate-model inputs resulted in an average HAB extent of 13 TMT (Table 1, Fig. 7). These values are similar to those from the historical period and lower than those from the current period.

Averaging across all climate, SWAT, and HAB models, projected P loads and HAB extents decreased 46% and 44% from current conditions, respectively. Future loads decreased 13% and HAB extent increased 17% compared to the historical period. This divergence in the direction of change between loads and HABs can be influenced by assumptions about changes in HAB-load sensitivity built into the Obenour, Bertani, and Ho models. These simulations assumed that the Obenour and Bertani future sensitivities remain the same as the later part of the current period (time term = 5.5), representing increased sensitivity, and that the 9-year cumulative DRP load for the Ho model is represented by future SWAT loads (2011 MT). If we set the Obenour and Bertani sensitivity term to the average of the current period (2.2), the future forecast decreases from 19 TMT to 14 TMT. However, if we set the Ho

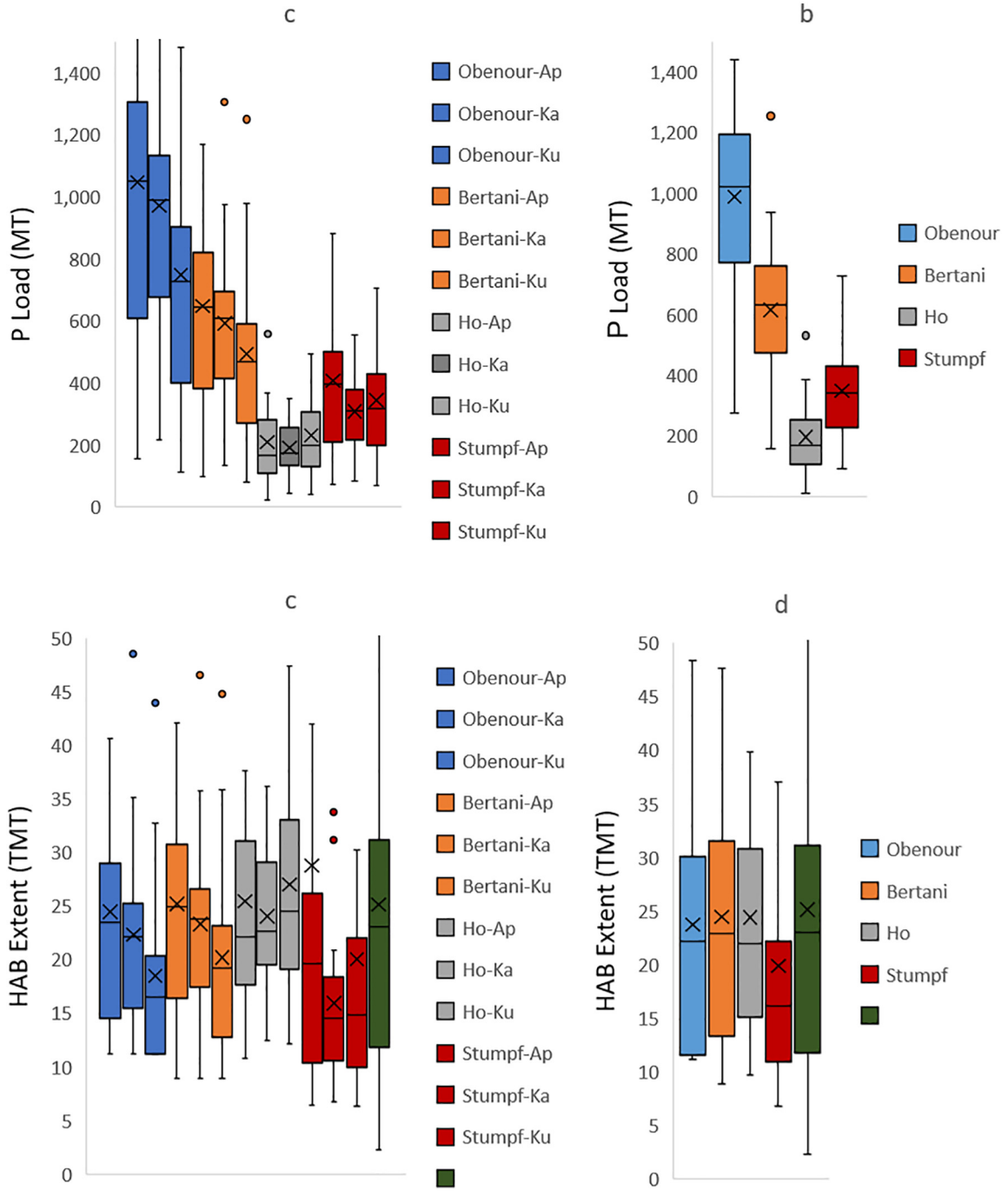


Fig. 5. Current period (2002-2017) phosphorus load simulations (a) and observations (b), as required by the HAB models, and the corresponding HAB extent observations and estimates based on simulated (c) and observed (d) loads. Obenour (blue), Bertani (orange), Ho (grey), and Stumpf (red) are the four HAB models. Obenour loads are TP, Bertani and Stumpf loads are different estimates of bioavailable P, and Ho loads are DRP. The Obenour and Bertani load window is 60-70% of March plus April-June, the Ho load window is March-July, and the Stumpf load window is April-July. Apostel (Ap), Kalcic (Ka), and Kujawa (Ku) are the three SWAT models.

sensitivity term to the average values for the current period (4111 MT), the HAB forecast increases from 8 TMT to 21 TMT. As a result, the forecasts averaged across all 4 HAB models change little, increasing from 13 to 14 TMT, a 40% decrease from the current period and a 26% increase from the historical period. Uncertainties associated with the sensitivity terms are also addressed in Section 3.5.

3.5. Sources of uncertainty

There was considerable within- and across-model uncertainty. For example, the coefficients of variation (standard deviation/mean) for simulated loads in the historical and current periods were 70% and 83% when simulated with observed climate (Table 1) and 54% and 59% for historical and future periods simulated with climate models. Similarly, HAB simulation coefficients of variation ranged between 16% and 58%. Given these levels of uncertainty, we explored various contributions to overall load and HAB uncertainties.

Simulated future HAB extent and load medians from the Monte Carlo analysis were relatively stable across all uncertainty cases (HAB mean = 12 TMT, load mean = 278 MT, Fig. 8) and similar to the deterministic models described above (13 TMT, 279 MT, Table 1). Because loads for the individual HAB models are composed of different combinations of

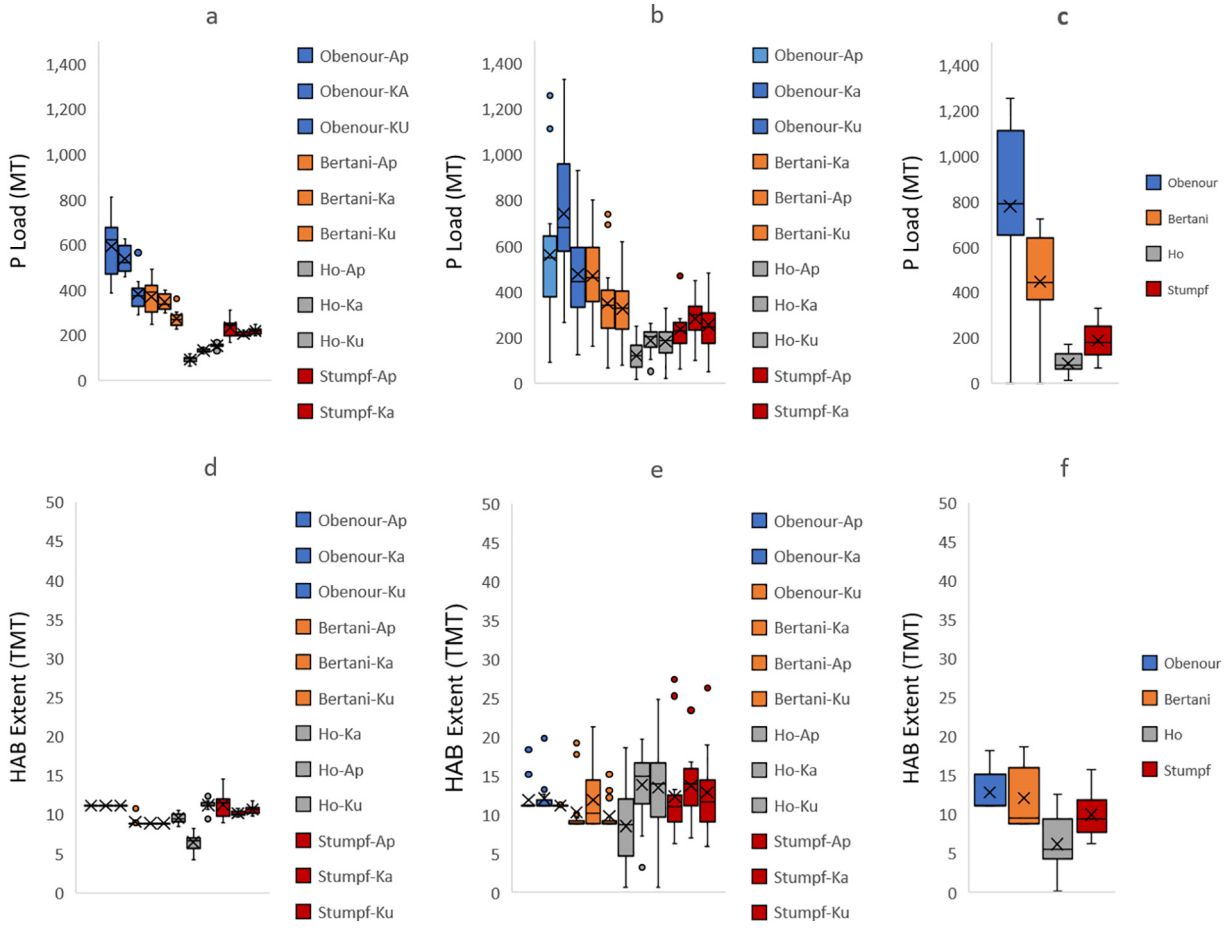


Fig. 6. Phosphorus load simulations for the historical time period (1985-1999) based on (a) climate model input to SWAT, (b) observed climate input to SWAT, and (c) measured loads, as required by the HAB models. Panels (d), (e), and (f) show HAB model predictions corresponding to the loads in (a), (b), and (c), respectively. Obenour, Bertani, Ho, and Stumpf are the four HAB models; Apostel (Ap), Kalcic (Ka), and Kujawa (Ku) are the three SWAT models.

months, TP, and DRP, we compare their coefficients of variation (CoV). Obenour and Bertani values (45–55%) were somewhat higher than Ho and Stumpf (ca. 30–35%). With fixed loads and uncertain model parameters, the Obenour and Bertani models generated higher HAB variance (first set of bars in Fig. 8a) because they have more parameters than Ho and Stumpf. As additional sources of uncertainty are added, HAB extent variance increases for all HAB models until the case that includes prediction and measurement errors when they all approach a variance of 100 TMT^2 (final set of bars in Fig. 8a).

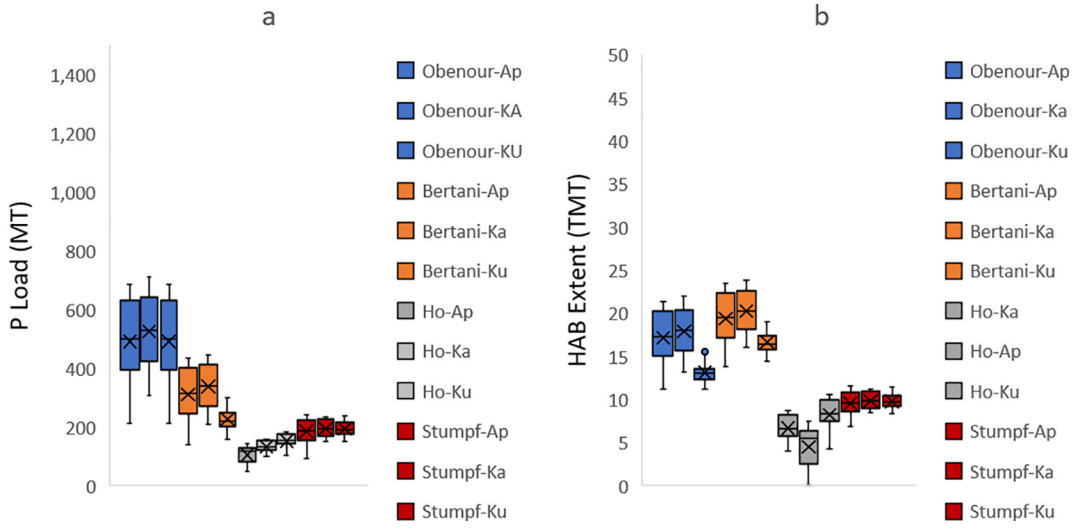


Fig. 7. Future mid-century (2051-2065) phosphorus load simulations based on climate model input to SWAT (a) and the corresponding HAB extents (b). Obenour, Bertani, Ho, and Stumpf are the four HAB models; Apostel (Ap), Kalcic (Ka), and Kujawa (Ku) are the three SWAT models.

Table 1
Comparison of simulated and measured averages and coefficients of variation (CoV) for loads (MT) and HABs (TMT). Load and HAB simulations are shown for both observed and modeled climate inputs. Simulation results reflect an aggregation across the climate, SWAT, and HAB models. *Measured loads and HABs are excluded from the averages.

	Averages*			CoV		
	Historical	Current	Future	Historical	Current	Future
Load						
Observed climate	349	515		70%	83%	
Modeled climate	294		279	54%		59%
Simulated*	321	515	279			
Measured	376	537				
HAB						
Observed climate	12	23		36%	58%	
Modeled climate	10		13	16%		43%
Simulated*	11	23	13			
Measured load	10	23				
Measured HAB		25				

These Monte Carlo simulations were performed with different combinations of uncertainty sources to allow us to isolate their impacts on overall HAB forecast variance. The results of source combinations are described here, and the isolated sources are discussed in Section 4. The variance across the four deterministic models (Fig. 8b), representing uncertainty in HAB model structure, was relatively low (6.0 TMT^2), reflecting the fact that the forecasts are well within the range of the calibration data. When HAB model parameter uncertainty or climate model uncertainty are added, variance increases to 35 TMT^2 and 23 TMT^2, respectively. Adding both Climate and SWAT uncertainties to the HAB model structure uncertainty only increased variance from 23 to 29 TMT^2.

Adding HAB model prediction error to HAB model structure uncertainty results in a variance of 45 TMT^2; adding HAB measurement error to that increases the variance to 98 TMT^2. These latter two variance sources represent uncertainties associated with interannual stochastic properties that are not accounted for in the deterministic (structural) relationships of the HABs models. As such, they do not affect the mean HAB predictions, although if these errors were reduced, HAB model parameter uncertainties would also be lower.

4. Discussion

4.1. Climate impacts on HAB-relevant spring loads

Simulated current P loads, driven with observed climate data, match measured loads well (Fig. 5), with some variation among SWAT models. Historical loads, driven by either observed or modeled climate, also match measured loads with some differences among SWAT models (Fig. 6). Inter-annual load variability, when driven by climate models, was considerably less than the variability in measured loads or loads simulated with observed climate. This would be expected because the bias-corrected climate outputs tend to converge toward common, observed values, thus reducing the SWAT output variability.

The simulated substantial increase between historical and current loads is consistent with Daloğlu et al. (2012), who suggested the interaction of climate change and agricultural practices over the last half of the 20th century was a potential cause of increased DRP loads. Jarvie et al. (2017) also suggested that increased DRP loads in the early 2000s were due to increases in both runoff and flux of DRP from the land.

Our mid-century load projections that are slightly lower than the historical period and 44% lower than the current period (Fig. 7, Table 1) are also consistent with earlier work. Using climate models and change factors applied to observed data, Verma et al. (2015) projected Maumee loads would decrease by 2045–2055. Driving a watershed model with an ensemble of 15 global climate models that were bias-corrected and downscaled, Culbertson et al. (2016) projected P loads may decrease due to increased plant uptake and decreased winter surface runoff. Kalcic et al. (2019) also projected less nutrient runoff by mid-century due to increased evapotranspiration and decreased snowfall, despite projected moderate increases in intensity and overall precipitation. Miralha et al. (in press) also associated mid-century load reductions with reduced early spring rainfall and higher temperatures and evapotranspiration leading to decreased flow.

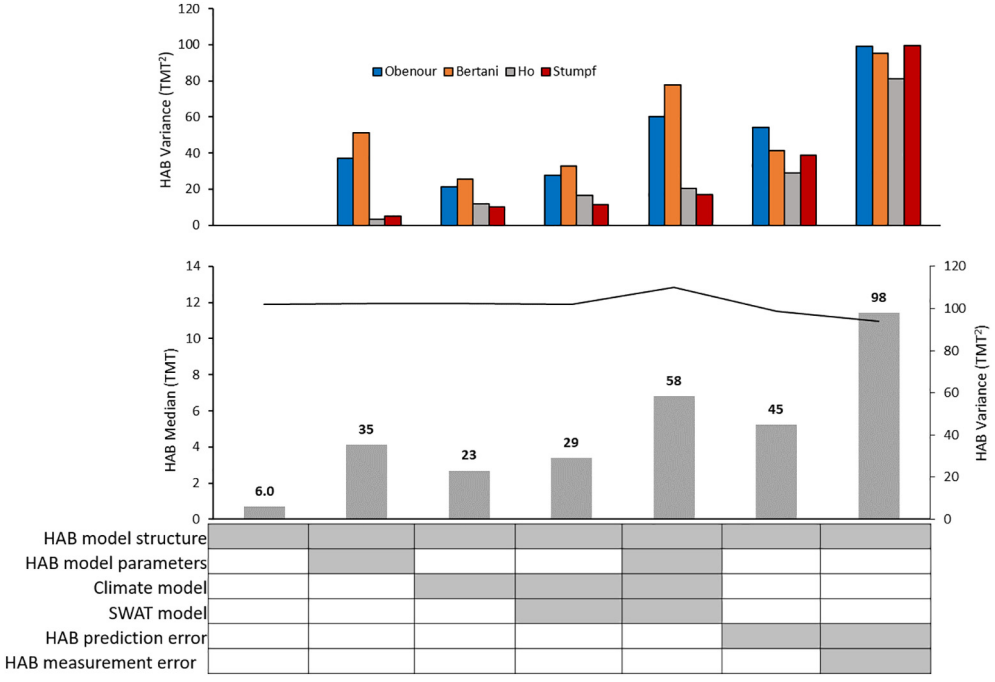


Fig. 8. a) Predictive variance of each individual HAB model and (b) combined HAB model medians (line) and variances (bars) from Monte Carlo simulations using different combinations of uncertainty sources. X-axis indicates whether the model property is fixed (open) or uncertain (shaded). Numbers above the bars are variances.

4.2. Climate impacts on HABs

The four HAB models performed well when calibrated against the three independent sets of HAB extent observations for 2002–2017, with r^2 values comparable to those reported for the original models when one apparent outlier year was removed. HAB-nutrient response curves based on these calibrations were used to simulate historical (1985–1999), current (2002–2017), and future (2051–2065) HAB extents based on both modeled and measured loads. Historical simulations (Fig. 6) were consistent with anecdotal information that HABs were rare prior to the mid-1990s (Vanderploeg et al., 2001; Conroy et al., 2005), and comparable to observations from the period of record prior to the expansion in 2008 (Fig. 3a). Current-period simulations aligned with both average (Fig. 5) and interannual (Fig. 3) observations, and our climate-based future forecasts suggest that HABs would decrease by 44% from the current period (Fig. 7, Table 1).

While our forecasts for mid-century HAB extent indicate an encouraging reduction from the current state, it is important to note that other research projecting mid-century load decreases also project increases in loads toward the end of the century (Verma et al., 2015; Cousino et al., 2015). In addition, we only evaluated the effect of climate change on nutrient loads, not on temperature-dependent HAB growth. As a species favored at warmer temperatures, *Microcystis* would likely have additional advantages in a warmer future (Paerl and Huisman, 2008; Mullin et al., 2020). For example, using a screening model of the combined effects of increased temperature and climate-induced changes in nutrient loads to assess potential changes in cyanobacteria HABs across U.S. lakes and reservoirs, Chapra et al. (2017) showed that HAB duration is likely to increase, primarily due to water temperature increases, from the current average 7 days per year to 16–23 days in 2050 and 18–39 days in 2090. Using a mechanistic HAB model, Del Giudice et al. (in press) also showed that a 2 degree increase in Lake Erie water temperature could result in roughly 20% more intense blooms initiating 10 days earlier. There is also the potential that blooms could become more toxic at higher temperatures (e.g., Davis et al., 2009), offsetting gains resulting from reduced nutrient loading.

4.3. Uncertainty propagation

We used Monte Carlo analysis to evaluate the relative contributions of climate, SWAT, and HAB model uncertainties to HAB projection uncertainty. Kujawa et al. (2020) characterized flow and nutrient load uncertainty from six climate models and five SWAT models for the Maumee River watershed and reported that climate model uncertainty was the dominant source of uncertainty for predicting total discharge, evapotranspiration, tile discharge, and total nitrogen loading; while SWAT uncertainty was the main source for surface runoff and P loadings. In our simulations that extend the analysis to HAB uncertainties, we found that SWAT model variance (6 TMT^2) was comparable to the HAB model structural uncertainty (6 TMT^2), but they each contributes less than half of the variance attributed to the climate models (17 TMT^2, Fig. 9). Combining our estimate of HAB model parameter uncertainty to HAB model structure uncertainty results in a variance contribution of 35 TMT^2, higher than the combined climate and SWAT model contributions (23 TMT^2). These results used the MBC-N climate bias correction method. The relative contributions were similar when using other methods, although Delta provided the lowest overall variance (Fig. S5).

This large contribution from HAB model parameter uncertainty, relative to the climate and SWAT models, may be because the Bayesian estimation provided formal estimates for them in addition to the structural uncertainties estimated by averaging outputs from the HAB models with fixed loads and parameter values. In contrast, parameter and structural uncertainties in the climate and SWAT models were estimated only by averaging across those models. Variability among climate models may also have been reduced by bias correcting their outputs before using them in SWAT (Yuan et al., 2020). For example, the climate

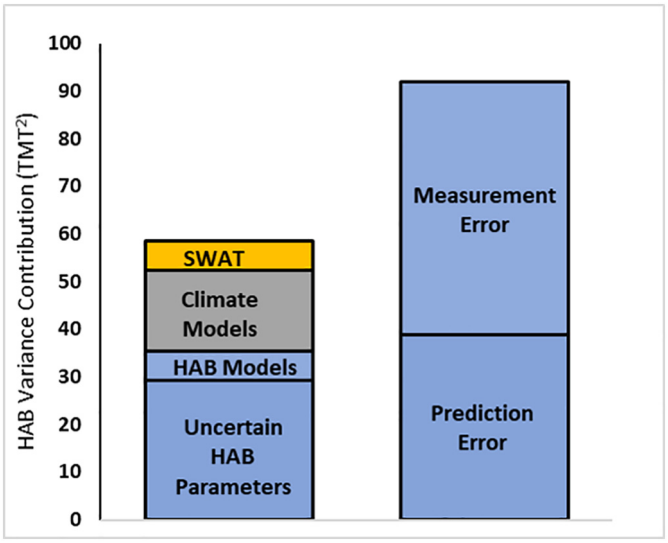


Fig. 9. Contributions to overall HAB prediction variance. The column on the left shows variances associated with the prediction of the future mean condition, while the right column shows additional variance associated with predicting bloom size in a particular year.

model contribution to overall HAB variance increased dramatically from 17 to 92 TMT^2 if model output was not bias-corrected, increasing the overall HAB extent variance of over 152 TMT^2 (Fig. S5).

The relatively low variance contribution from SWAT is likely because we were not able to account for parameter uncertainty independently as we did with the HAB models. In addition, while these SWAT implementations were quite different (Table S2), they were all calibrated and validated with a common set of data (Table S3). While this relatively low uncertainty is encouraging, Apostel et al. (in review) showed that even though each model performed well at the watershed outlet, their ability to match observations at the field scale were different, and often lacking.

4.4. Reducing uncertainties

Improvements in climate-based HAB forecasts can be affected through improvements in regional climate projections and watershed models, along with enhanced monitoring to better bound HAB model parameters. A more accurate evaluation of climate model uncertainties would involve using non-bias corrected models. However, to better represent measured nutrient loads for HAB forecasts, the corrections were necessary for TP and DRP (Miralha et al., In Press). As climate projections improve to finer resolutions and improve their accounting for large water bodies, it may be possible to both reduce uncertainties and forego bias correction.

Watershed models in the Great Lakes region have been converging on a consistent set of forecasts at the watershed outlet, which is fortunate for making connections to HAB models. However, more work needs to be done to test and validate that these models are also representative of upstream and field-scale processes (Apostel et al., in review). It is likely that improvements at the field scale should also reduce uncertainties at the outlet.

HAB uncertainty from model prediction error and measurement error was larger than the combined forecast uncertainties from the climate, SWAT, and HAB models (Fig. 9). These variances represent uncertainties associated with interannual stochastic properties that are not accounted for in the models and errors associated with HAB extent measurements, and therefore do not affect the mean HAB predictions. However, reducing these errors would lead to lower model parameter uncertainties and increase confidence in those long-term forecasts. Measurement error is best illustrated by comparing three sets of independent extent estimates (Fig. 3). Improvements in temporal and spatial

coverage, for example as suggested from the geostatistical analysis by Fang et al. (2019), should help reduce that uncertainty. Prediction errors arise because, for example, these models assume that only the P load determines maximum bloom extent. Processes not in these models that can contribute to the uncertainty include: light and/or nitrogen limitation; temperature-regulated growth, meteorological controls on stratification, vertical mixing, and advection; and grazing and sinking losses. Including aspects of these processes, as some have done (e.g., Del Giudice et al., In Press; Verhamme et al., 2016), could reduce prediction error. But, because these models can introduce dozens of additional parameters, this may be at the cost of increased parameter uncertainty.

While we only used four climate models, the number of models is comparable to the number of HAB and SWAT models used, and it likely approximates at least a subset of climate model variability. The three SWAT models were considerably different implementations, but quantifying parameter uncertainty directly, as done for the HAB models, would have increased their uncertainty, as would including additional types of watershed models. A more extensive analysis with additional climate and watershed models would be helpful in prioritizing uncertainty reduction strategies.

Recognizing uncertainties as they propagate from climate models and through watershed and HAB models is important, and efforts to reduce them are needed. While environmental models will always contain some level of uncertainty, they nonetheless provide useful inputs and guidance for policy development, and as Reichert and Borsuk (2005) demonstrate, the uncertainty in the difference of model predictions corresponding to different policies may be significantly smaller than the uncertainty in the predictions themselves. Recent efforts to use multiple models to provide ensemble forecasts and scenarios (e.g., Scavia et al., 2016, 2017a, 2017b; Martin et al., 2019) potentially mediate some of the uncertainty.

CRediT authorship contribution statement

Donald Scavia: Conceptualization, Methodology, Writing - Original draft preparation, **Yu-Chen Wang**: Methodology, Software, Data curation, **Daniel Obenour:** Conceptualization, Reviewing and Editing, **Anna Apostel:** Methodology, Reviewing and Editing, **Lorrayne Miralha**: Methodology, Reviewing and Editing, **Samantha Basile:** Reviewing and Editing, **Margaret Kalcic**: Conceptualization, Reviewing and Editing, **Christine Kirchhoff**: Reviewing and Editing **Rebecca Muenich:** Conceptualization, Reviewing and Editing. **Allison Steiner:** Conceptualization, Supervision, Reviewing and Editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This project was funded by the National Science Foundation Coastal SEES grant OCE-1600012. We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups (listed in Section 2.2 of this paper) for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2020.143487.

References

Apostel, A., M. Kalcic, R. Muenich, G. Evenson, A. Dagnew, J. Kast, R. Muenich, K. King, J. Martin, D. Scavia. Validating internal watershed processes using multiple SWAT models (in review).

Baker, D.B., Confesor, R., Ewing, D.E., Johnson, L.T., Kramer, J.W., Merryfield, B.J., 2014a. Phosphorus loading to Lake Erie from the Maumee, Sandusky, and Cuyahoga rivers: the importance of bioavailability. J. Great Lakes Res. 40, 502–517.

Baker, D.B., Ewing, D.E., Johnson, L.T., Kramer, J.W., Merryfield, B.J., Confesor, R.B., Richards, R.P., Roerdink, A.A., 2014b. Lagrangian analysis of the transport and processing of agricultural runoff in the lower Maumee River and Maumee Bay. J. Great Lakes Res. 40, 479–495.

Basile, S.J., Rauscher, S.A., Steiner, A.L., 2017. Projected precipitation changes within the Great Lakes and Western Lake Erie Basin: a multi-model analysis of intensity and seasonality. Int. J. Climatol. 37, 4864–4879. https://doi.org/10.1002/joc.5128.

Bertani, I., Obenour, D.R., Steger, C.E., Stow, C.A., Gronewold, A.D., Scavia, D., 2016. Probabilistically assessing the role of nutrient loading in harmful algal bloom formation in western Lake Erie. J. Great Lakes Res. 42 (6), 1184–1192.

Bosch, N.S., Evans, M.A., Scavia, D., Allan, J.D., 2014. Interacting effects of climate change and agricultural BMPs on nutrient runoff entering Lake Erie. J. Great Lakes Res. 40 (3), 581–589.

Chapra, S.C., Boehlert, B., Fant, C., Bierman Jr., V.J., Henderson, J., Mills, D., Mas, D.M.L., Rennels, L., Jantarasami, L., Martinich, J., Strzepek, K.M., Paerl, H.W., 2017. Climate change impacts on harmful algal blooms in U.S. freshwaters: a screening-level assessment. Environ. Sci. Technol. 51, 8933–8943.

Conroy, J.D., Edwards, W.J., Pontius, R.A., Kane, D.D., Zhang, H., Shea, J.F., Richey, J.N., Culver, D.A., 2005. Soluble nitrogen and phosphorus excretion of exotic freshwater mussels (Dreissena spp.): potential impacts for nutrient remineralisation in western Lake Erie. Freshw. Biol. 50 (7), 1146–1162.

Cousino, L.K., Becker, R.H., Zmijewski, K.A., 2015. Modeling the effects of climate change on water, sediment, and nutrient yields from the Maumee River watershed. J. Hydrol. Reg. Stud. 4, 762–775.

Culbertson, A.M., Martin, J.F., Aloysius, N., Ludsin, S.A., 2016. Anticipated impacts of climate change on 21st century Maumee River discharge and nutrient loads. J. Great Lakes Res. 42 (6), 1332–1342.

Daloğlu, I., Cho, K.H., Scavia, D., 2012. Evaluating causes of trends in long-term dissolved reactive phosphorus loads to Lake Erie. Environ. Sci. Technol. 46 (19), 10660–10666.

Davis, T.W., Berry, D.L., Boyer, G.L., Gobler, C.J., 2009. 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.

Del Giudice, D., S. Fang, D. Scavia, T. W. Davis, M. A. Evans, D. R. Obenour. In Press. Elucidating controls on cyanobacteria bloom timing and intensity via Bayesian mechanistic modeling. Sci. Total Environ.

Delignette-Muller, Marie Laure, Dutang, Christophe, 2015. Fitdistrplus: an R package for fitting distributions. J. Stat. Softw. 64 (4), 1–34.

Evans, R.O., Fausey, N.R., 2015. Effects of inadequate drainage on crop growth and yield. In: Skaggs, R., Schilfgaarde, J. (Eds.), Agricultural Drainage https://doi.org/10.2134/agronmonogr38.c2.

Fang, S., Giudice, D.D., Scavia, D., Binding, C.E., Bridgeman, T.B., Chaffin, J.D., Evans, M.A., Guinness, J., Johengen, T.H., Obenour, D.R., 2019. A space-time geostatistical model for probabilistic estimation of harmful algal bloom biomass and areal extent. Sci. Total Environ. https://doi.org/10.1016/j.scitotenv.2019.133776.

Glibert, P.M., Allen, J.I., Bouwman, A.F., Brown, C.W., Flynn, K.J., Lewitus, A.J., Madden, C.J., 2010. Modeling of HABs and eutrophication: status, advances, challenges. J. Mar. Syst. 83, 262–275. https://doi.org/10.1016/j.jmarsys.2010.05.004.

GLWQA (Great Lakes Water Quality Agreement), 2016. The United States and Canada adopt phosphorus load reduction targets to combat Lake Erie algal blooms. https://binational.net/2016/02/22/finalptargets-ciblesfinalesdep.

Gobler, C.J., 2020. Climate change and harmful algal blooms: insights and perspective. Harmful Algae 91, 101731. https://doi.org/10.1016/j.hal.2019.101731.

Ho, J.C., Michalak, A.M., 2015. Challenges in tracking harmful algal blooms: a synthesis of evidence from Lake Erie. J. Great Lakes Res. 41, 317–325.

Jarvie, H.P., Johnson, L.T., Sharpley, A.N., Smith, D.R., Baker, D.B., Bruulsema, T.W., Confesor, R., 2017. Increased soluble phosphorus loads to Lake Erie: unintended consequences of conservation practices? J. Environ. Qual. 46 (1), 123–132.

Johnson, T., Butcher, J., Deb, D., Faizullabhoy, M., Hummel, P., Kittle, J., McGinnis, S., Mearns, L.O., Nover, D., Parker, A., et al., 2015. Modeling streamflow and water quality sensitivity to climate change and urban development in 20 U.S. watersheds. J. Am. Water Resour. Assoc. 51 (5), 1321–1341.

Kalcic, M.M., Kirchhoff, C., Bosch, N., Muenich, R.L., Murray, M., Griffith Gardner, J., Scavia, D., 2016. Engaging stakeholders to define feasible and desirable agricultural conservation in Western Lake Erie watersheds. Environ. Sci. Technol. 50 (15), 8135–8145.

Kalcic, M.M., Muenich, R.L., Basile, S., Steiner, A.L., Kirchhoff, C., Scavia, D., 2019. Climate change and nutrient loading: warming can counteract a wetter future. Environ. Sci. Technol. 53, 7543–7550.

Kujawa, H., Kalcic, M., Martin, J., Aloysius, N., Apostel, A., Kast, J., Murumkar, A., Evenson, G., Becker, R., Boles, C., Confesor, R., Dagnew, A., Guo, T., Muenich, R.L., Redder, T., Scavia, D., Wang, Y.-C., 2020. The hydrologic model as a source of nutrient loading uncertainty in a future climate. Sci. Total Environ. https://doi.org/10.1016/j.scitotenv.2020.138004.

Lunn, D.J., Thomas, A., Best, N., Spiegelhalter, D., 2000. WinBUGS—a Bayesian modelling framework: concepts, structure, and extensibility. Stat. Comput. 10, 325–337.

Manning, N.F., Wang, Y.C., Long, C.M., Bertani, I., Sayers, M.J., Bosse, K.R., Shuchman, R.A., Scavia, D., 2019. Extending the forecast model: predicting harmful algal blooms at multiple spatial scales. J. Great Lakes Res. 45, 587–595.

Martin, J.F., Kalcic, M.M., Aloysius, N., Apostel, A.M., Brooker, M.R., Evenson, G., Kast, J.B., Kujawa, H., Murumkar, A., Becker, R., Boles, C., Redder, T., Confesor, R., Guo, T.,

Dagnew, A., Long, C.M., Muenich, R., Scavia, D., Wang, Y., Robertson, D., 2019. Evaluating Management Options to Reduce Lake Erie Algal Blooms With Models of the Maumee River Watershed. Final Project Report. http://kx.osu.edu/project/environment/habri-multi-model.

Menne, M.J., Durre, I., Korzeniewski, B., McNeal, S., Thomas, K., Yin, X., Anthony, S., Ray, R., Vose, R.S., Gleason, B.E., Houston, T.G., 2012. Global Historical Climatology Network - Daily (GHCN-Daily), Version 3. NOAA National Climatic Data Center https://doi.org/10.7289/V5D21VHZ.

Michalak, A.M., Anderson, E.J., Beletsky, D., Boland, S., Bosch, N.S., Bridgeman, T.B., Chaffin, J.D., Cho, K., Confesor, R., Daloğlu, I., DePinto, J.V., 2013. Record-setting algal bloom in Lake Erie caused by agricultural and meteorological trends consistent with expected future conditions. Proc. Natl. Acad. Sci. 110 (16), 6448–6452.

Millie, D.F., Fahnenstiel, G.L., Dyble Bressie, J., Pigg, R.J., Rediske, R.R., Klarer, D.M., Tester, P.A., Litaker, R.W., 2009. Late-summer phytoplankton in western Lake Erie (Laurentian Great Lakes): bloom distributions, toxicity, and environmental influences. Aquat. Ecol. 43, 915–934.

Miralha, L., Muenich, R.L., Scavia, D., Wells, K., Steiner, A.L., Kalcic, M., Apostel, A., Basile, S., Kirchhoff, C. In Press. Bias correction of climate model outputs influences nutrient load predictions. Sci. Total Environ.

Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Binger, R.L., Harmel, R.D., Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. ASABE 50 (3), 885–900.

Mullin, C.A., Kirchhoff, C.J., Wang, G., Vlahos, P., 2020. Future projections of water temperature and thermal stratification in Connecticut reservoirs and possible implications for cyanobacteria. Water Resour. Res. https://doi.org/10.1029/2020WR027185.

Obenour, D., Gronewold, A., Stow, C.A., Scavia, D., 2014. Using a Bayesian hierarchical model to improve Lake Erie cyanobacteria bloom forecasts. Water Resour. Res. 50, 7847–7860.

Paerl, H.W., Huisman, J., 2008. Blooms like it hot. Science 320 (5872), 57–58.

R Core Team, 2015. R: A Language and Environment for Statistical Computing.

Ralson, D.K., Moore, S.K., 2020. Mondeling harmful algal blooms in a changing climate. Harmful Algae 91, 101729. https://doi.org/10.1016/j.hal.2019.101729.

Reichert, P., Borsuk, M.E., 2005. Does high forecast uncertainty preclude effective decision support? Environ. Model. Softw. 20, 991–1001.

Salmaso, N., Boscaini, A., Capelli, C., Cerasino, L., 2018. Ongoing ecological shifts in a large lake are driven by climate change and eutrophication: evidences from a three-decade study in Lake Garda. Hydrobiologia 824, 177–195.

Sayers, M., Fahnenstiel, G.L., Shuchman, R.A., Whitley, M., 2016. Cyanobacteria blooms in three eutrophic basins of the Great Lakes: a comparative analysis using satellite remote sensing. Int. J. Remote Sens. 37, 4148–4171.

Scavia, D., David Allan, J., Arend, K.K., Bartell, S., Beletsky, D., Bosch, N.S., Brandt, S.B., Briland, R.D., Daloğlu, I., DePinto, J.V., et al., 2014. Assessing and addressing the re-eutrophication of Lake Erie: Central Basin hypoxia. J. Great Lakes Res. 40 (2), 226–246.

Scavia, D., DePinto, J.V., Bertani, I., 2016. A multi-model approach to evaluating target phosphorus loads for Lake Erie. J. Great Lakes Res. 42 (6), 1139–1150.

Scavia, D., Kalcic, M., Muenich, R.L., Read, J., Aloysius, N., Bertani, I., Boles, C., Confesor, R., DePinto, J., Gildow, M., et al., 2017a. Multiple models guide strategies for agricultural nutrient reductions. Front. Ecol. Environ. 15 (3), 126–132.

Scavia, D., Bertani, I., Obenour, D.R., Turner, R.E., Forrest, D.R., Katin, A., 2017b. Ensemble modeling informs hypoxia management in the northern Gulf of Mexico. Proc. Natl. Acad. Sci. 114, 8823–8828.

Stumpf, R.P., Johnson, L.T., Wynne, T.T., Baker, D.B., 2016. Forecasting annual cyanobacterial bloom biomass to inform management decisions in Lake Erie. J. Great Lakes Res. 42 (6), 1174–1183.

Sturtz, S., Ligges, U., Gelman, A., 2005. R2WinBUGS: a package for running WinBUGS from R. J. Stat. Softw. 12, 1–16.

Taylor, Karl, Stouffer, Ronald, Meehl, Gerald, 2012. An overview of CMIP5 and the experiment design. Bull. Amer. Meteor. Soc. 93 (4), 485–498. https://doi.org/10.1175/BAMS-D-11-00094.1.

Vanderploeg, H.A., Liebig, J.R., Carmichael, W.W., Agy, M.A., Johengen, T.H., Fahnenstiel, G.L., Nalepa, T.F., 2001. Zebra mussel (*Dreissena polymorpha*) selective filtration promoted toxic Microcystis blooms in Saginaw Bay (Lake Huron) and Lake Erie. Can. J. Fish. Aquat. Sci. 58, 1208–1221.

Verhamme, E., Redder, T., Schlea, D., Grush, J., Bratton, J., DePinto, J., 2016. Development of the Western Lake Erie Ecosystem Model (WLEEM): application to connect phosphorus loads to cyanobacteria biomass. J. Great Lakes Res. 42 (6), 1193–1205.

Verma, S., Bhattarai, R., Bosch, N.S., Cooke, R.C., Kalita, P.K., Markus, M., 2015. Climate change impacts on flow, sediment and nutrient export in a Great Lakes watershed using SWAT. CLEAN – Soil Air Water 43 (11), 1464–1474.

Wynne, T.T., Stumpf, R.P., Tomlinson, M.C., Schwab, D.J., Watabayashi, G.Y., Christensen, J.D., 2011. Estimating cyanobacterial bloom transport by coupling remotely sensed imagery and a hydrodynamic model. Ecol. Appl. 21, 2709–2721.

Yuan, S., Quiring, S.M., Kalcic, M.M., Apostel, A.M., Evenson, G.R., Kujawa, H.A., 2020. Optimizing climate model selection for hydrological modeling: a case study in the Maumee River Basin using the SWAT. J. Hydrol. https://doi.org/10.1016/j.jhydrol.2020.125064.