Using a Multi-Institutional Ensemble of Watershed Models to Assess Agricultural Conservation Effectiveness in a Future Climate

Haley Kujawa, Margaret Kalcic, Jay Martin, Anna Apostel, Jeffrey Kast, Asmita Murumkar, Grey Evenson, Noel Aloysius, Richard Becker, Chelsie Boles, Remegio Confesor, Awoke Dagnew, Tian Guo, Rebecca Logsdon Muenich, Todd Redder, Yu-Chen Wang, and Donald Scavia

Research Impact Statement: Multiple watershed models provide robust predictions of agricultural conservation effectiveness in a future climate.

ABSTRACT: This study investigates the combined impacts of climate change and agricultural conservation on the magnitude and uncertainty of nutrient loadings in the Maumee River Watershed, the second-largest watershed of the Laurentian Great Lakes. Two scenarios — baseline agricultural management and increased agricultural conservation — were assessed using an ensemble of five Soil and Water Assessment Tools driven by six climate models. The increased conservation scenario included raising conservation adoption rates from a baseline of existing conservation practices to feasible rates in the near future based on farmer surveys. This increased adoption of winter cover crops on 6%–10% to 60% of cultivated cropland; subsurface placement of phosphorus fertilizers on 35%–60% to 68% of cultivated cropland; and buffer strips intercepting runoff from 29%–34% to 50% of cultivated cropland. Increased conservation resulted in statistically significant (p ≤ 0.05) reductions in annual loads of total phosphorus (41%), dissolved reactive phosphorus (18%), and total nitrogen (14%) under the highest emission climate scenario (RCP 8.5). While nutrient loads decreased with increased conservation relative to baseline management for all watershed models, different conclusions on the true effectiveness of conservation under climate change may be drawn if only one watershed model was used.

(KEYWORDS: climate change; hydrology; Soil and Water Assessment Tool; nutrients; scenario analysis.)

INTRODUCTION

An increase in intense phytoplankton blooms since the 1980s has been seen across the globe (Ho et al. 2019), and is expected to worsen with climate change (Paerl and Paul 2012), posing significant risks to human health and ecology (Codd 2000; Codd et al. 2005; Liu et al. 2011; Lee et al. 2017; Wituszynski et al. 2017). A variety of factors contribute to the
intense and often harmful algal blooms (HABs), with a primary driver in many regions being nonpoint source nutrient runoff (Beaver et al. 2014; Stumpf et al. 2016). Scientists often predict climate change will increase nutrient runoff to water bodies with more frequent and more intense precipitation events (Paerl and Paul 2012), but there is evidence that in some regions increased precipitation could be offset by higher evapotranspiration, resulting in decreased nutrient loading with climate change (Kalcic et al. 2019; Kujawa et al. 2020; Scavia et al. 2021).

Watershed models are often used to predict future changes in nutrient loading and assess the influence of conservation practices on water quality (Bosch et al. 2014; Johnson et al. 2015; Verma et al. 2015; Kalcic et al. 2019; Kujawa et al. 2020). However, predictions from watershed models can be highly uncertain due to variability in climate model forecasts (e.g., Kujawa et al. 2020; Miralha et al. 2021) as well as variability in how watershed models simulate load response to landscape changes (e.g., Scavia et al. 2017; Martin et al. 2021). Considerable work in apportioning prediction uncertainty to climate vs. hydrologic models has been done for water quantity in a future climate (Wilby and Harris 2006; Wilby et al. 2006; Kay et al. 2008; Addor et al. 2014). These studies show that climate model uncertainty can greatly outweigh uncertainty from the hydrologic model, or that the hydrologic and climate model uncertainties may be comparable, depending on which factors are considered (e.g., inputs, parameterizations, chosen climate models) (Kay et al. 2008; Bosshard et al. 2013; Karlsson et al. 2016; Thober et al. 2018).

There is no comprehensive body of work on predicting conservation practice effectiveness using an ensemble of climate models, watershed models, and land management scenarios. Of the studies mentioned above, only Karlsson et al. (2016) investigated changes in discharge using multiple climate and watershed models as well as multiple land-use scenarios. The study found land-use change across hydrologic models to significantly affect the extreme hydrologic response (i.e., low flow and flooding), but overall land use had a modest contribution on average discharge variation compared to the climate and watershed models. Kujawa et al. (2020) investigated the uncertainty in predictions of climate change in the Maumee River Watershed using an ensemble of watershed models and found phosphorus predictions could be highly uncertain based on decisions made in the set-up of the watershed model (e.g., subroutines, parameterizations, land management assumptions). While Kujawa et al. (2020) concluded that uncertainty from watershed models is significant for predicting nutrient discharge under a future climate, it may be even more important when the complexity of land management scenario analysis is added. Scavia et al. (2021) investigated the relative uncertainty of nutrient prediction in a series of climate, watershed, and HABs models, and found the watershed model contributed to the overall uncertainty for nutrient load predictions, albeit less than the HABs and climate models. Few studies have examined the watershed model variations’ impact on water quality under changing climate and land management.

The goal of this study was to understand the combined impacts of climate change and increased agricultural conservation (IC) on riverine nutrient loading. The two objectives were (1) to predict if IC will reduce nutrient loadings in a future climate, and (2) to assess whether the effectiveness of agricultural conservation will change between historical and future climate periods. This study was carried out in the Maumee River Watershed, the second-largest watershed of the Great Lakes. We used an ensemble of five watershed models developed using the Soil and Water Assessment Tool (SWAT) and climate data from six downscaled General Circulation Models (GCMs) under the highest emission scenario (RCP 8.5) as well as two agricultural land management scenarios (historical management and IC practices including buffer strips, subsurface placement, and cover crops).

METHODS

Study Area

The study area was the Maumee River Watershed (~17,300 km²), located in northwest Ohio, northeast Indiana, and southeast Michigan (Figure 1). The Maumee River is a major tributary to Lake Erie. Lake Erie has experienced significant eutrophication issues since the 1960s and the region has since focused on managing phosphorus to control eutrophication and HABs (Schindler 1974; De Pinto et al. 1986; Schindler et al. 2016). The 2012 Great Lakes Water Quality Agreements specified phosphorus reduction as the main strategy to control HABs in the western Lake Erie basin (GLWQA 2015; USEPA 2018).

Decreased nutrient loading to Lake Erie is necessary to lessen HABs and protect the public health of the people in the Western Lake Erie Basin. In 2014, Lake Erie bloom toxicity caused the City of Toledo to issue a three-day “Do Not Drink” advisory (Jetoo et al. 2015). While there are few studies directly linking human health and Microcystis blooms, some have shown the correlation of increases in liver disease and cancer in areas that have Microcystis blooms (Lee et al. 2019; Gorham et al. 2020).
The Maumee River Watershed’s primary land use is agriculture, specifically row crops such as corn, soy, and winter wheat (Figure 1). The Maumee River watershed is also the second-largest contributor of phosphorus loading to Lake Erie but has the highest phosphorus concentrations which are essential for algal bloom growth (Elser 1999; Michalak et al. 2013). Phosphorus from the Maumee River Watershed is largely from nonpoint sources, primarily agricultural (Maccoux et al. 2016). Therefore, increased attention has been placed on this watershed to reduce phosphorus, particularly from agricultural activities (Scavia et al. 2017; Kalcic et al. 2019; Martin et al. 2021).

Climate Models

We chose to examine mid-century climate change (2046–2065) because the projections are more certain than for the end-of-century, and agricultural and watershed managers tend to find the early period more relevant for their planning. The timescales of interest, annual and March–July, were chosen because they correspond to loading targets in the Great Lakes Water Quality Agreement as the most relevant predictors of central basin hypoxia and harmful algal bloom size, respectively (GLWQA 2015; Stumpf et al. 2016).

Six GCMs were taken from the Coupled Model Intercomparison Project Phase 5 (CMIP5) ensemble and previously downscaled to 1/8° latitude-longitude (~12 × 12 km) resolution using Bias-corrected Constructed Analogues (Reclamation 2013). The number of GCMs was chosen based on similar climate and watershed studies (Kay et al. 2008; Velazquez et al. 2013; Prudhomme et al. 2014; Giuntoli et al. 2015; Vetter et al. 2017; Thober et al. 2018). We focused on the highest-emissions scenario, RCP 8.5, and included GCMs that varied across the expected range in precipitation change. Both of these choices would be expected to produce the greatest variation in discharge and nutrient loading at the watershed scale (Michalak et al. 2013; Gao et al. 2019; Kujawa et al. 2020).

The climate model data for the watershed spanned 1.5 standard deviations of the CMIP5 ensemble mean for precipitation and remained close to the mean change (annual mean increase of 2.76°C; March–July mean increase of 2.74°C) for temperature. The changes in annual precipitation from historical (H; 1996–2015) to mid-century (MC; 2046–2065) ranged from a 5% decrease to a 12% increase, and the mean temperature increased between 2.5°C and 3.0°C. The changes in March–July precipitation ranged from a 3% decrease to a 19% increase, and the mean temperature increased between 2.5°C and 2.9°C (Table 1).
Watershed Models

The watershed model ensemble was built with SWAT. SWAT is a watershed model commonly used to assess nonpoint source pollution in watersheds, as well as climate change impacts on hydrology and nutrients (e.g., Bosch et al. 2014; Verma et al. 2015; Culbertson et al. 2016; Scavia et al. 2017; Čerkasova et al. 2018; Wang et al. 2018). The model uses inputs of elevation, land use and land cover, climate, and soils, runs on daily time scales, and is able to simulate a wide range of agricultural and land-management practices (Neitsch et al. 2009; Arnold et al. 2012). SWAT has been shown to be a suitable model for the Maumee River Watershed given its ability to represent a range of agricultural management practices and achieve a good calibration (Gebremariam et al. 2014).

Five modeling groups from different institutions built unique SWAT model configurations of the Maumee River Watershed. Some of the inputs, such as point sources and percent of agricultural land with certain management practices, were similar across these models (Table 2). Each SWAT model has the same climate stations and inputs (Figure 1). However, each group independently made most modeling assumptions, such as most land management operations, automatic or manual calibration choices, and specific model routine implementation.

All models were calibrated to a single station (Maumee at Waterville, USGS # 04193500) near the watershed outlet and achieved good standards of performance for discharge and nutrients (Moriasi et al. 2007, 2015; Table S1). However, variations in watershed models that all perform well at the outlet can have significant differences in process representation upstream (Apostel et al. 2021). Hence, we then consider each SWAT model as unique. For more detail on baseline model set-up guidelines and variation among model inputs, see supporting information in Kujawa et al. (2020).

Agricultural Management Scenarios

The two agricultural management scenarios included baseline management (BM) and IC. Conservation practices used in BM represented historical (2005–2015) rates of cover crops, buffer strips, and subsurface placement/incorporation (Table 2). Cover crops were implemented on 6%–10% of cultivated cropland and buffers intercepted runoff from 29%–34% of cultivated cropland. There was significant variation in subsurface placement, with implementation on 35%–60% of cultivated cropland depending on the model. Interpretation was left to each modeling group on whether to include fertilizer incorporation with tillage as subsurface placement.


<table>
<thead>
<tr>
<th>Shortened climate model name (in text)</th>
<th>Full climate model name</th>
<th>Institute</th>
<th>Reference</th>
<th>Original resolution (° long x ° lat)</th>
<th>Annual Change in temp. (°C)</th>
<th>Change in precip. (%)</th>
<th>March–July Change in temp. (°C)</th>
<th>Change in precip. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CanESM</td>
<td>CanESM2</td>
<td>Canadian Center for Climate Modeling and Analysis</td>
<td>Arora et al. (2011)</td>
<td>2.81° x 2.81°</td>
<td>2.8</td>
<td>12</td>
<td>2.9</td>
<td>12</td>
</tr>
<tr>
<td>CSIRO_r6</td>
<td>CSIRO-MK3-6-0</td>
<td>CSIRO Marine and Atmospheric Research</td>
<td>Rotstayn et al. (2010)</td>
<td>1.875° x 1.875°</td>
<td>2.5</td>
<td>5</td>
<td>2.7</td>
<td>19</td>
</tr>
<tr>
<td>CSIRO_r4</td>
<td>CSIRO-MK3-4-0</td>
<td>CSIRO Marine and Atmospheric Research</td>
<td>Rotstayn et al. (2010)</td>
<td>1.875° x 1.875°</td>
<td>2.5</td>
<td>2</td>
<td>2.5</td>
<td>4</td>
</tr>
<tr>
<td>CSIRO_r10</td>
<td>CSIRO-MK3-10-0</td>
<td>CSIRO Marine and Atmospheric Research</td>
<td>Rotstayn et al. (2010)</td>
<td>1.875° x 1.875°</td>
<td>3.0</td>
<td>13</td>
<td>2.9</td>
<td>17</td>
</tr>
<tr>
<td>MPI-ESM</td>
<td>MPI-ESM-LR</td>
<td>Max Planck Institute for Meteorology</td>
<td>Zanchettin et al. (2013)</td>
<td>1.875° x 1.875°</td>
<td>2.5</td>
<td>−5</td>
<td>2.5</td>
<td>−3</td>
</tr>
<tr>
<td>NorESM</td>
<td>NorESM1-M</td>
<td>Norwegian Climate Centre</td>
<td>Bentsen et al. (2013)</td>
<td>2.5° x 1.875°</td>
<td>2.9</td>
<td>9</td>
<td>2.7</td>
<td>17</td>
</tr>
</tbody>
</table>
The IC scenario increased adoption rates above the BM, from the rates listed in Table 2 to a total rate of cover crops on 60% of cultivated cropland, subsurface placement on 68% of cultivated cropland, and buffer strips intercepting runoff from 50% of cultivated cropland. These increases were chosen in collaboration with a stakeholder advisory group to represent feasible adoptions based on farmer surveys (Martin et al. 2021).

The five SWAT models were run with each management scenario and driven with output from the six climate models, resulting in 60 simulations. We applied downscaled precipitation and temperature outputs to the SWAT models by selecting and directly inputting the climate grid data having the closest centroid to the rain gauge stations included in each SWAT model. All model runs were continuous for 1996–2065 with a 5-year warm-up period beforehand.

<table>
<thead>
<tr>
<th>SWAT model (institution where built)</th>
<th>Cover crops</th>
<th>Subsurface placement/incorporation</th>
<th>Buffer strips</th>
</tr>
</thead>
<tbody>
<tr>
<td>UT (University of Toledo)</td>
<td>Cereal rye cover crop planted in 10% of cultivated cropland, limited to corn-soybean rotations</td>
<td>Incorporation via tillage immediately following phosphorus application on 60% of cultivated cropland</td>
<td>Buffers intercepting surface runoff from 32% of cultivated cropland</td>
</tr>
<tr>
<td>UM (University of Michigan)</td>
<td>Cereal rye cover crop planted in 8.4% of cultivated cropland</td>
<td>Incorporation via tillage three days after phosphorus application on 57% of cultivated cropland, with 21% of cropland having a mixture of broadcast and incorporation of fertilizer</td>
<td>Buffers intercepting surface runoff from 34% of cultivated cropland</td>
</tr>
<tr>
<td>OSU (Ohio State University)</td>
<td>Cereal rye cover crop planted in 8.4% of cultivated cropland and limited to corn-soybean rotations.</td>
<td>Subsurface application of phosphorus fertilizers on 35% of cultivated cropland.</td>
<td>Buffers intercepting surface runoff from 29% of cultivated cropland.</td>
</tr>
<tr>
<td>LT (LimnoTech)</td>
<td>Cereal rye cover crop planted in 7.5% of cultivated cropland</td>
<td>Subsurface placement of phosphorus fertilizers on 40% of cultivated cropland</td>
<td>Buffers intercepting surface runoff from 30% of cultivated cropland.</td>
</tr>
<tr>
<td>HU (Heidelberg University)</td>
<td>Cereal rye cover crop planted in 6% of cultivated cropland</td>
<td>Subsurface placement of phosphorus fertilizers on 43.6% of cultivated cropland</td>
<td>Buffers intercepting surface runoff from 30% of cultivated cropland</td>
</tr>
</tbody>
</table>

The IC scenario increased adoption rates above the BM, from the rates listed in Table 2 to a total rate of cover crops on 60% of cultivated cropland, subsurface placement on 68% of cultivated cropland, and buffer strips intercepting runoff from 50% of cultivated cropland. These increases were chosen in collaboration with a stakeholder advisory group to represent feasible adoptions based on farmer surveys (Martin et al. 2021).

The five SWAT models were run with each management scenario and driven with output from the six climate models, resulting in 60 simulations. We applied downscaled precipitation and temperature outputs to the SWAT models by selecting and directly inputting the climate grid data having the closest centroid to the rain gauge stations included in each SWAT model. All model runs were continuous for 1996–2065 with a 5-year warm-up period beforehand.

**Metrics for Assessing Conservation Effectiveness with the Ensemble**

Two objectives were tested through comparisons of these 60 simulations across climate and management scenarios (Figure 2).

Objective 1 assessed changes in nutrient loading due to conservation and climate change. We compared the effect of future climate under the two agricultural management scenarios for each climate and SWAT model combination. Changes in hydrology and nutrients from average historical (e.g., BM<sub>H</sub>, average of the BM scenario from 1996–2015) to average mid-century (e.g., IC<sub>MC</sub>, average of the IC scenario from 2046–2065) climates were calculated for each climate and SWAT model combination as the change under BM in a future climate,

\[ \Delta BM_{MC-H} = \frac{BM_{MC} - BM_{H}}{BM_{H}} \times 100, \]

and the change due to both IC and future climate,

\[ \Delta IC_{MC-H} = \frac{IC_{MC} - BM_{H}}{BM_{H}} \times 100. \]

Objective 2 assessed the effectiveness of IC under climate change. We calculated change due to IC in each of the time periods, and the compare them to one another:

\[ \Delta IC_{H} = \frac{IC_{H} - BM_{H}}{BM_{H}} \times 100, \]

\[ \Delta IC_{MC} = \frac{IC_{MC} - BM_{MC}}{BM_{MC}} \times 100. \]

These objectives were tested at both annual and March–July timescales.
for the entire ensemble (SWAT and GCMs) in \( \Delta I_{MC-H} \) were increased discharge (+3%), subsurface discharge (+17%), and ET (+8%), and decreased surface runoff (−15%). The signal-to-noise ratios were similar under both management scenarios (i.e., none changed from below to above one, or vice versa) demonstrating that changes in hydrology were not largely affected by agricultural management choices.

In contrast, differences between the changes in nutrient loads for \( \Delta I_{MC-H} \) and \( \Delta B_{MC-H} \) were statistically significant (\( p \leq 0.05 \)). The percent changes in annual loads were less for the \( \Delta I_{MC-H} \) scenario compared to \( \Delta B_{MC-H} \). In \( \Delta I_{MC-H} \), total phosphorus (TP) decreased by 41%, dissolved reactive phosphorus (DRP) by 18%, and total nitrogen (TN) by 14%. While DRP and TN signal-to-noise ratios remained below 1, they were greater for \( \Delta I_{MC-H} \) compared to \( \Delta B_{MC-H} \), indicating a tendency toward agreement among ensemble members. The strongest agreement (signal-to-noise > 1) was a reduction in TP with climate change under \( \Delta I_{MC-H} \) (Figure 3).

Some SWAT models predicted increases in annual nutrient loading under \( \Delta I_{MC-H} \) at mid-century. However, in all cases, there was a lesser increase under \( \Delta I_{MC-H} \) compared to \( \Delta B_{MC-H} \) scenario, indicating that IC was always helpful in reducing nutrient loads (Figure 4). While the overall ensemble predicted statistically significant nutrient load reductions between \( \Delta B_{MC-H} \) and \( \Delta I_{MC-H} \) (Figure 3), this difference was not always significant for individual SWAT models (Figure 4).

**Changes in March–July Hydrology and Nutrient Loads.** Changes in hydrologic characteristics between \( \Delta B_{MC-H} \) and \( \Delta I_{MC-H} \) scenarios for March–July, the period that governs the extent of HABs in Lake Erie’s western basin (GLWQA2015), were similar to those found at the annual timescale (Figure 5). Climate change alone (\( \Delta B_{MC-H} \)) did not result in large changes in discharge (+3%), but had greater changes in surface runoff (−10%), subsurface discharge (+9%), and ET (+12%). The addition of conservation (\( \Delta I_{MC-H} \)) resulted in similar and insignificant deviations from \( \Delta B_{MC-H} \) in discharge (+2%), surface runoff (−11%), subsurface discharge (+7%), and ET (+13%). The only hydrologic change with signal-to-noise above 1 was for ET in both \( \Delta B_{MC-H} \) and \( \Delta I_{MC-H} \) (Figure 5).

On the contrary, differences between management scenarios (\( \Delta B_{MC-H} \) and \( \Delta I_{MC-H} \)) were statistically significant (\( p \leq 0.05 \)) for TP, DRP, and TN loadings (Figure 5). DRP increased due to climate change (\( \Delta B_{MC-H} \) +11%) but decreased with the addition of conservation (\( \Delta I_{MC-H} \) −11%; Table 3). TP was virtually unchanged due to climate (\( \Delta B_{MC-H} \), −2%) and greatly reduced with IC (\( \Delta I_{MC-H} \), −34%). Only TP

**RESULTS**

**Changes in Hydrology and Nutrient Loads in a Mid-Century Climate (Objective 1)**

**Annual Changes in Hydrology and Nutrient Loads.** The combination of IC and climate change (\( \Delta I_{MC-H} \)) produced no significant differences in annual discharge, subsurface discharge, surface runoff, and evapotranspiration (ET) when compared with the impact of climate change under \( \Delta B_{MC-H} \) (\( p > 0.05 \); Figure 3). The mean differences...
and TN in the ΔICMC–H scenario had signal-to-noise ratios greater than one, signifying ensemble agreement (Figure 5). Taken together, this indicates a greater agreement in a more pronounced effect from IC and climate change than from climate change alone.

Individual SWAT models exhibited March–July load patterns similar to annual loads. All SWAT models had less increase in nutrient loading in the ΔICMC–H scenario compared to the ΔBMMC–H scenario under future climate, demonstrating consistency for ΔICMC–H in reducing nutrients. Statistically significant differences were detected between ΔBMMC–H and ΔICMC–H for TP in Heidelberg University (HU), LimnoTech (LT), and Ohio State (OSU) SWAT models, for DRP in University of Michigan (UM) and LT SWAT models, and for TN in HU, LT, and OSU SWAT models (Figure 6). Yet, the predicted reductions by the ΔICMC–H scenario for individual SWAT models were not as clear as for the ensemble. Some of the models showed clear improvements under ΔICMC–H relative to ΔBMMC–H (e.g., UM), while this change was not as apparent in other models (e.g., UT; Figure 6).

**Changing Effectiveness of Conservation under Climate Change (Objective 2)**

The SWAT ensemble showed a slight decrease in the effectiveness of IC for TP and DRP in the mid-century climate (Table 4). However, the differences between mid-century (ΔICMC) and historical (ΔICH) conservation effectiveness were not statistically significant when tested with a two-sided Wilcoxon Rank Sum test (p > 0.05; Table S6). On average, IC reduced annual TP by 40% in historical climate (ΔICH), and only 36% by the mid-century (ΔICMC; Table 4). Similarly, a 24% reduction of DRP in a historical climate dropped to 21% in the mid-century. March–July patterns were similar (Table 4). There was considerable variation among SWAT models in both direction and magnitude of change in conservation effectiveness for phosphorus loading (TP and DRP) by mid-century: HU and LT predicted little to no change (±1–2 percentage points) in conservation effectiveness in both historical (ΔICH) and mid-century (ΔICMC). OSU and UM showed decreased effectiveness of IC for phosphorus and UT showed increased effectiveness (Table 4). All SWAT models demonstrated greater conservation effectiveness for TN in the mid-century. On average, annual TN was reduced 10% in this historical period (ΔICH) and 14% at mid-century (ΔICMC). Slightly larger TN reductions were produced for March–July, with an 11% reduction in the historical period (ΔICH) and 17% in the mid-century (ΔICMC; Table 4).

**DISCUSSION**

**Effects of Climate Change and Increased Conservation on Nutrient Loading**

Our goal was to assess the combined impacts of climate change and IC on riverine nutrient loading. The first objective was to use an ensemble of watershed and climate models to assess if IC will reduce nutrient loadings in a future climate (ΔICMC–H). We
showed that $\Delta IC_{MC-H}$ would be effective in a future climate, with statistically significant ($p \leq 0.05$) decreases in nutrient loadings relative to $\Delta BM_{MC-H}$ for both annual and March–July loads of TP ($-41\%$ annually; $-34\%$, March–July), DRP ($-18\%$ annually; $-11\%$ March–July), and TN ($-14\%$ annually; $-24\%$).

FIGURE 4. Annual results for the individual SWAT models for $\Delta BM_{MC-H}$ and $\Delta IC_{MC-H}$. Rank Sum test for statistically significant changes ($^*p \leq 0.05$, $^{**}p \leq 0.01$, ns = not significant) between $\Delta BM_{MC-H}$ and $\Delta IC_{MC-H}$ denoted above boxplots. Signal to noise values for individual SWAT models can be found in Supporting Information (Table S3).

FIGURE 5. March–July changes for $\Delta BM_{MC-H}$ and $\Delta IC_{MC-H}$. Signal-to-noise ratio is directly above each boxplot. Rank Sum test for statistically significant changes ($^{***}p \leq 0.001$, $^{****}p \leq 0.0001$, ns = not significant) between $\Delta BM_{MC-H}$ and $\Delta IC_{MC-H}$ denoted above boxplots.
March–July; Table 3; Figures 3 and 5). The reductions in annual dissolved and TP loads were larger because there were considerable decreases from December to February (Figure S1). TN, however, exhibited greater percentage decreases during the March–July period because the largest reductions were in March and April (Figure S1).

The potential mid-century reductions in phosphorus and nitrogen due to IC (ΔICMC-H compared to ΔBMMC-H) demonstrate the effectiveness of conservation despite uncertainty associated with climate change. DRP was predicted to have an average, albeit small, increase due to climate change alone (ΔBMMC-H); however, introducing IC (ΔICMC-H) yielded a significant reduction in DRP (Figure 5). Thus, IC should help reduce future HABs because DRP is primarily bioavailable fuel for algal growth (Scavia et al. 2014).

The TP reductions between historical and mid-century were small and highly uncertain under ΔBMMC-H, but clear and consistent in ΔICMC-H (Figures 3 and 5). Phosphorus is often prioritized in developing freshwater mitigation strategies because there remains more evidence of phosphorus as the limiting factor in freshwater bloom initiation (Schindler 1974; Schindler et al. 2008, 2016; Stumpf et al. 2016; Schindler et al. 2016).

### TABLE 3. Average changes for ΔBMMC-H and ΔICMC-H between historical (1996–2015) and MC (2046–2065) climate as a percent change ± standard deviation (SD).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ΔBMMC-H</td>
<td>−7 ± 13</td>
<td>−2 ± 21</td>
<td>+1 ± 20</td>
<td>+11 ± 26</td>
<td>−1 ± 12</td>
<td>−9 ± 15</td>
</tr>
<tr>
<td>ΔICMC-H</td>
<td>−41 ± 15</td>
<td>−34 ± 21</td>
<td>−18 ± 14</td>
<td>−11 ± 22</td>
<td>−14 ± 15</td>
<td>−24 ± 19</td>
</tr>
</tbody>
</table>

**FIGURE 6.** March–July results for the individual SWAT models for ΔBMMC-H and ΔICMC-H. Rank sum test for statistically significant changes (*p ≤ 0.05, **p ≤ 0.01, ns = not significant) between ΔBMMC-H and ΔICMC-H denoted above boxplots. Signal to noise values for individual SWAT models can be found in Supporting Information (Table S3).
TABLE 4. Changes in nutrient loading (± SD) due to the IC scenario in both historical (∆IC\textsubscript{H}) and mid-century (∆IC\textsubscript{MC}) time periods averaged within each SWAT model. For the full list of climate and SWAT models (see Table S3).

<table>
<thead>
<tr>
<th>Variable</th>
<th>SWAT</th>
<th>∆IC\textsubscript{H} (1996–2015)</th>
<th>∆IC\textsubscript{MC} (2046–2065)</th>
<th>∆IC\textsubscript{H} (1996–2015)</th>
<th>∆IC\textsubscript{MC} (2046–2065)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TP</td>
<td>HU</td>
<td>−42 ± 1</td>
<td>−39 ± 2</td>
<td>−37 ± 2</td>
<td>−36 ± 1</td>
</tr>
<tr>
<td></td>
<td>LT</td>
<td>−57 ± 2</td>
<td>−56 ± 2</td>
<td>−52 ± 2</td>
<td>−53 ± 2</td>
</tr>
<tr>
<td></td>
<td>OSU</td>
<td>−35 ± 2</td>
<td>−25 ± 3</td>
<td>−29 ± 2</td>
<td>−21 ± 3</td>
</tr>
<tr>
<td></td>
<td>UM</td>
<td>−45 ± 1</td>
<td>−36 ± 1</td>
<td>−45 ± 2</td>
<td>−35 ± 1</td>
</tr>
<tr>
<td></td>
<td>UT</td>
<td>−21 ± 0</td>
<td>−26 ± 2</td>
<td>−18 ± 2</td>
<td>−24 ± 4</td>
</tr>
<tr>
<td>DRP</td>
<td>Average ± SD</td>
<td>−40 ± 13</td>
<td>−36 ± 13</td>
<td>−36 ± 13</td>
<td>−34 ± 12</td>
</tr>
<tr>
<td></td>
<td>HU</td>
<td>−19 ± 1</td>
<td>−18 ± 1</td>
<td>−17 ± 1</td>
<td>−16 ± 1</td>
</tr>
<tr>
<td></td>
<td>LT</td>
<td>−25 ± 1</td>
<td>−26 ± 1</td>
<td>−24 ± 1</td>
<td>−26 ± 1</td>
</tr>
<tr>
<td></td>
<td>OSU</td>
<td>−24 ± 2</td>
<td>−17 ± 1</td>
<td>−16 ± 1</td>
<td>−11 ± 1</td>
</tr>
<tr>
<td></td>
<td>UM</td>
<td>−38 ± 0</td>
<td>−30 ± 1</td>
<td>−39 ± 0</td>
<td>−30 ± 1</td>
</tr>
<tr>
<td></td>
<td>UT</td>
<td>−13 ± 1</td>
<td>−16 ± 3</td>
<td>−10 ± 1</td>
<td>−14 ± 2</td>
</tr>
<tr>
<td>TN</td>
<td>Average ± SD</td>
<td>−24 ± 9</td>
<td>−21 ± 6</td>
<td>−21 ± 11</td>
<td>−19 ± 8</td>
</tr>
<tr>
<td></td>
<td>HU</td>
<td>−13 ± 1</td>
<td>−17 ± 2</td>
<td>−11 ± 1</td>
<td>−13 ± 2</td>
</tr>
<tr>
<td></td>
<td>LT</td>
<td>−17 ± 1</td>
<td>−26 ± 2</td>
<td>−19 ± 2</td>
<td>−30 ± 4</td>
</tr>
<tr>
<td></td>
<td>OSU</td>
<td>−12 ± 0</td>
<td>−16 ± 1</td>
<td>−17 ± 1</td>
<td>−26 ± 4</td>
</tr>
<tr>
<td></td>
<td>UM</td>
<td>−7 ± 1</td>
<td>−7 ± 1</td>
<td>−6 ± 1</td>
<td>−8 ± 1</td>
</tr>
<tr>
<td></td>
<td>UT</td>
<td>−2 ± 0</td>
<td>−2 ± 4</td>
<td>−3 ± 6</td>
<td>−7 ± 5</td>
</tr>
<tr>
<td></td>
<td>Average ± SD</td>
<td>−10 ± 6</td>
<td>−14 ± 9</td>
<td>−11 ± 7</td>
<td>−17 ± 11</td>
</tr>
</tbody>
</table>

USEPA 2018). The predicted reductions of both TP and DRP indicate the conservation practices can decrease nutrient runoff and reduce the extent of Lake Erie HABs now and in the mid-century.

The ensemble predicted a reduction in TN in a mid-century climate under ∆BM\textsubscript{MC–H}, and predicted a further reduction due to ∆IC\textsubscript{MC–H}. Some evidence suggests nitrogen reductions could be critical in limiting freshwater algal bloom size, duration, and toxicity (Chaffin et al. 2014; Gobler et al. 2016; Paerl et al. 2016; Newell et al. 2019). While nitrogen is not at this time considered the top priority nutrient for this watershed (USEPA 2018), it is still important for Microcystis bloom duration (Chaffin et al. 2014) and toxicity (Gobler et al. 2016). It is then promising that TN is predicted to be further reduced with IC.

Our second objective was to assess whether the effectiveness of IC will change in a future climate. We found that on average IC was slightly less effective in reducing phosphorus in the mid-century (∆IC\textsubscript{MC} < ∆IC\textsubscript{H}), but the difference between these two scenarios was not statistically significant (Table 4; Table S2). Nitrogen was more clearly reduced with conservation in the mid-century (∆IC\textsubscript{MC}) because the effects of climate alone reduced TN without additional conservation (Kujawa et al. 2020). Therefore, the combined effects of climate change and IC decreased TN loadings further in the mid-century (Table 4).

Across the two objectives, we found that the combined effects of climate change and IC will likely lead to reductions in nitrogen and phosphorus loading to Lake Erie. Other studies on the climate change in the Maumee River Watershed found reductions in sediment and nutrients with increased conservation practices. Cousino et al. (2015) simulated 100% no-till on agricultural areas and found this lowered sediment yields by 16% compared to corresponding climate scenarios under historical management with conventional tillage (nutrient data not included in findings). Bosch et al. (2014) evaluated modest adoption rates of agricultural conservation practices: 25% of cropland with cover crops and no-till and 20% with filter strips. They found this scenario of agricultural conservation practices with modest changes in climate showed annual loading reductions of 6% TP, 4% DRP, and 4% TN. Similar to this study, Bosch et al. (2014) also found conservation practices (no-till, cover crops, filter strips) to be less effective in a future climate.

Bosch et al. (2014) and Cousino et al. (2015) chose to focus on variability caused by the different climate scenarios and only include one watershed model. However, our results agree with Bosch et al. (2014) that average nutrient loading in the Maumee can be reduced in a future climate with additional conservation despite increased precipitation, and that a slight decrease in conservation effectiveness may occur with climate change. Furthermore, novel results from this work demonstrated that greater but feasible adoption rates of agricultural conservation practices show greater potential for nutrient reduction. It is important to note while this study demonstrated a need for agricultural conservation to reduce nutrient loadings, some climate change studies in the Maumee found reductions in nutrients under historical management due to climate alone (Kalcic et al. 2019; Scavia et al. 2021).
While most climate scenarios show overall increases in precipitation, it may be important to study the effects of seasonal drought in the Maumee on the impacts of changes in water management. Byun et al. (2019) found decreasing soil moisture in the Great Lakes and Midwest and suggest the combination of drought and temperature stresses may lead to greater irrigation. While the research is lacking, Paul et al. (2020) simulated introducing irrigation to a rain-fed watershed increased surface runoff and suggests it may subsequently increase nutrient loss.

**Uncertainty and Variability within the Ensemble of Climate and Watershed Models**

Understanding the extent of uncertainty and variability in watershed modeling and scenario analysis helps to effectively communicate climate change uncertainty to stakeholder groups and inform the scientific development of models (Korfmaccher 1998; Gregory and Dieckmann 2013). In this study, there was greater uncertainty in the direction of change in hydrology and water quality when considering climate change alone (under $\Delta BM_{MC-H}$) than in the combination of climate change and IC. Depending on the watershed and climate model used, BM results showed increasing or decreasing phosphorus loads in a future climate (Kujawa et al. 2020). However, including increased conservation practices practices ($\Delta IC_{MC-H}$) consistently reduced phosphorus loading as compared to BM ($\Delta BM_{MC-H}$). Much of the variability stemmed from differences in setting up and calibrating the models under BM (e.g., management assumptions, spatial discretization of models, and parameters; Evenson et al. 2021; Kujawa et al. 2020) and differences in implementation of the IC scenario (Table 2). Allowing modeling groups to independently develop separate baseline SWAT models allowed for a more holistic accounting of differences in watershed models than if one team was to develop several SWAT models with varying inputs and parameterizations.

While the ensemble demonstrated promising results for greater adoption of conservation practices to reduce nutrients in the future, the confidence in nutrient reduction becomes less apparent if an individual watershed model is chosen. The effectiveness of the $\Delta IC_{MC-H}$ scenario for historical and mid-century climates varied across models, and different conclusions may be reached regarding the impacts of climate change and nutrient reduction potential. For example, the UM model predicts increased phosphorus loadings under $\Delta BM_{MC-H}$ and decreased loadings under $\Delta IC_{MC-H}$. This creates a clear message that IC is essential to counteract a negative consequence of climate change. However, the UT model showed DRP reductions under both $\Delta BM_{MC-H}$ and $\Delta IC_{MC-H}$ (Figure 4), suggesting that DRP loadings will decrease with or without additional conservation. While these models send conflicting messages about the value of conservation in future nutrient loading, the ensemble avoids drawing conclusions based on a single watershed model and still captures this variability in watershed model response. We attribute much of the variation of changes in phosphorus loading to the setup and parametrization of the SWAT models (Kujawa et al. 2020). Parameterizations and submodels used in SWAT can affect the dominant transport pathway of phosphorus and the subsequent effects of climate change and agricultural conservation on phosphorus loss.

Communicating model uncertainty and variability to stakeholders is challenging. This study includes many facets of uncertainty in models, such as impacts of parameterization, scenario analysis, and spatial discretization on nutrient predictions in climate analysis. Future research could explicitly investigate each facet of uncertainty and their interactions, as well as other factors, such as emissions scenario, downscaling techniques, watershed model calibration methods (e.g., multi-site calibration), and multiple regions of interest (e.g., Wilby et al. 2006; Kay et al. 2008; Velazquez et al. 2013). Information on the dominant causes of uncertainty in nutrient load prediction can inform subsequent studies on creating a watershed model ensemble that provides a holistic accounting of uncertainty and variability in climate change analysis, as well as inform opportunities for focused watershed model development. Phosphorus management is becoming a critical issue to address concerning eutrophication worldwide (Jeppesen et al. 2009; Bol et al. 2018; Ho et al. 2019), and this study has demonstrated the value of using multiple watershed models to capture uncertainty and variability in scenario results. While the analysis presented herein is limited to one watershed and two land management scenarios, it contributes to a growing body of information on the subject of nutrient modeling and climate analysis, given there will always be some inherent uncertainty regardless of how advanced models become (Beven 2016). One goal is to continuously improve and better integrate climate models, watershed nutrient models, and stakeholder interests to address nutrient runoff and HABs so results can better inform environmental management and adaptation (e.g., Scavia et al. 2021).

Scientists may need guidance to determine what knowledge on model uncertainty is required to further agricultural sustainability in the face of climate change before heading down the “refine-experiment-refine” pathway (Miller et al. 2011). Research has been conducted over the past several decades (e.g.,
in the Maumee River Watershed, the second largest tributary to Lake Erie, will be effective in reducing nutrients in the mid-century (2046–2065) under the highest emission scenario (RCP 8.5). The IC scenario showed significant reductions in nutrients compared to BM for this period. The combined effects of IC and climate change, on average, predicted annual (March–July) decreases of 41% (34%) for TP, 18% (11%) for DRP, and 14% (24%) for TN. The IC scenario was slightly more effective in reducing phosphorus in the historical period than in the mid-century. In addition, watershed models varied considerably in their assessment of to what degree IC, combined with climate change, will produce phosphorus load reductions. The ensemble consistently predicted nitrogen load reductions due to IC and climate change, in part because climate alone reduced nitrogen loads in the mid-century.

We suggest predictions of hydrology and water quality in a future climate should more frequently employ an ensemble of watershed models. Further study on the effects of individual sources of watershed model uncertainty on predictions of water quality can be used to improve watershed model development and inform future climate impact studies. Interdisciplinary stakeholder engagement should accompany defining further research on model uncertainty and nutrient prediction. In this way, scientists can create a body of literature on model uncertainty more suited to realistically address agricultural sustainability in the face of climate change.

**DATA AVAILABILITY STATEMENT**

The data used in this study are available upon reasonable request. Please contact the corresponding author.

**SUPPORTING INFORMATION**

Additional supporting information may be found online under the Supporting Information tab for this article: Supplemental data and analysis.

**ACKNOWLEDGMENTS**

This project was funded by the Ohio Department of Higher Education Harmful Algal Bloom Research Initiative (R/HAB-5-ODHE). We thank our stakeholder advisory group for their active engagement in designing the project and their review of the work. We also acknowledge the World Climate Research Programme’s Working Group on Coupled Modeling, which is responsible for CMIP, the
United States Department of Energy’s Program for Climate Model Diagnosis and Intercomparison, and the Global Organization for Earth System Science Portals. We are very appreciative of all research groups involved in the production and distribution of the climate data used in this project (Table 1). Open access funding enabled and organized by Projekt DEAL.

AUTHOR CONTRIBUTIONS

Haley Kujawa: Data curation; formal analysis; visualization; writing – original draft. Margaret Kalcic: Conceptualization; funding acquisition; methodology; resources; software; supervision; validation; writing – review and editing. Jay Martin: Conceptualization; funding acquisition; methodology; supervision; writing – review and editing. Anna Apostel: Data curation; visualization. Jeffrey Kast: Data curation; writing – review and editing. Asmita Murumkar: Visualization; writing – review and editing. Grey Evenson: Writing – review and editing. Noel Aloysius: Conceptualization; methodology; resources; software; validation. Richard Becker: Software; validation. Chelsie Boles: Software; validation; writing – review and editing. Remegio Con-fesor: Data curation; formal analysis; writing – review and editing. Awoke Dagnew: Software; validation. Tian Guo: Software; validation; writing – review and editing. Rebecca Logsdon Muenich: Software; validation. Todd Redder: Software; validation. Yu-Chen Wang: Resources; software; validation. Donald Scavia: Writing – review and editing.

LITERATURE CITED


