ARTICLE IN PRESS

Journal of Environmental Management xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of Environmental Management

journal homepage: http://www.elsevier.com/locate/jenvman

Research article

Evaluating management options to reduce Lake Erie algal blooms using an ensemble of watershed models

Jay F. Martin [a,b], Margaret M. Kalcic [a], Noel Aloysius [c], Anna M. Apostel [a], Michael R. Brooker [a,*], Grey Evenson [a], Jeffrey B. Kast [a], Haley Kujawa [a], Asmita Murumkar [a], Richard Becker [d], Chelsie Boles [e], Remegio Confesor [f], Awoke Dagnew [g], Tian Guo [f], Colleen M. Long [h], Rebecca L. Muenich [i], Donald Scavia [j], Todd Redder [e], Dale M. Robertson [k], Yu-Chen Wang [h]

[a] Department of Food, Agricultural, and Biological Engineering, The Ohio State University, Columbus, OH, 43210, United States
[b] Sustainability Institute, The Ohio State University, Columbus, OH, 43210, United States
[c] Department of Biomedical, Biological & Chemical Engineering, and School of Natural Resources, University of Missouri, Columbia, MO, 65211, United States
[d] Department of Environmental Sciences, University of Toledo, Toledo, OH, 43606, United States
[e] LimnoTech, Ann Arbor, MI, 48108, United States
[f] National Center for Water Quality Research, Heidelberg University, Tiffin, OH, 44883, United States
[g] Environmental Consulting and Technology, Inc., Ann Arbor, MI, 48105, United States
[h] Water Center, Graham Sustainability Institute, University of Michigan, Ann Arbor, MI, 48104, United States
[i] School of Sustainable Engineering and the Built Environment, Arizona State University, Tempe, AZ, 85281, United States
[j] School for Environment and Sustainability, University of Michigan, Ann Arbor, MI, 48104, United States
[k] Upper Midwest Water Science Center, US Geological Survey (USGS), Middleton, WI, 53562, United States

ARTICLE INFO

Keywords:
Eutrophication
Phosphorus
Best management practices
SWAT

ABSTRACT

Reducing harmful algal blooms in Lake Erie, situated between the United States and Canada, requires implementing best management practices to decrease nutrient loading from upstream sources. Bi-national water quality targets have been set for total and dissolved phosphorus loads, with the ultimate goal of reaching these targets in 9-out-of-10 years. Row crop agriculture dominates the land use in the Western Lake Erie Basin thus requiring efforts to mitigate nutrient loads from agricultural systems. To determine the types and extent of agricultural management practices needed to reach the water quality goals, we used five independently developed Soil and Water Assessment Tool models to evaluate the effects of 18 management scenarios over a 10-year period on nutrient export. Guidance from a stakeholder group was provided throughout the project, and resulted in improved data, development of realistic scenarios, and expanded outreach. Subsurface placement of phosphorus fertilizers, cover crops, riparian buffers, and wetlands were among the most effective management options. But, only in one realistic scenario did a majority (3/5) of the models predict that the total phosphorus loading target would be met in 9-out-of-10 years. Further, the dissolved phosphorus loading target was predicted to meet the 9-out-of-10-year goal by only one model and only in three scenarios. In all scenarios evaluated, the 9-out-of-10-year goal was not met based on the average of model predictions. Ensemble modeling revealed general agreement about the effects of several practices although some scenarios resulted in a wide range of uncertainty. Overall, our results demonstrate that there are multiple pathways to approach the established water quality goals, but greater adoption rates of practices than those tested here will likely be needed to attain the management targets.

1. Introduction

Setting water quality goals is a necessary step to combat the global issue of coastal eutrophication (Paerl and Otten, 2013; Schindler et al., 2016), but the necessary strategies for achieving these goals are often difficult to define. The symptoms of eutrophication include degraded water quality (Watson et al., 2016); reduced and contaminated fisheries (Bukaveckas et al., 2017; Wituszynski et al., 2017); threats to irrigation

* Corresponding author.
E-mail address: brooker.26@osu.edu (M.R. Brooker).

https://doi.org/10.1016/j.jenvman.2020.111710
Received 15 May 2020; Received in revised form 20 November 2020; Accepted 21 November 2020

and potable water supplies (Lee, J. et al., 2017); and decreased tourism, cultural activities, and coastal economies (Watson et al., 2016; Wolf et al., 2017). These effects have been observed in and around Lake Erie since the mid-1990s due to a resurgence in the excessive growth of algae and cyanobacteria, some of which produce toxins. Harmful algal blooms (HABs) in the lake are thought to be primarily growth-limited by phosphorus; however, the species of algae and their toxicity may be affected by the amount of nitrogen (Davis et al., 2015; Horst et al., 2014; Michalak et al., 2013; Newell et al., 2019; Scavia et al., 2014; Stow et al., 2015). Reducing the effects of eutrophication in Lake Erie, as well as other systems globally, requires improved land management to reduce anthropogenic nutrient loading from its watershed (Conley et al., 2009).

Prior to the 1990s, Lake Erie experienced eutrophication symptoms that were addressed largely through banning phosphorus (P) in detergents and regulating point sources contributing to total phosphorus (TP) loads (Watson et al., 2016). Despite these efforts, TP loading has remained high while there has been an upward trend in dissolved reactive phosphorus (DRP) that has led to the current manifestation of HABs (Baker et al., 2014). The negative effects of both TP and DRP loading have been recognized in a binational agreement between the United States (U.S.) and Canada. Pursuant to the Great Lakes Water Quality Agreement (GLWQA) of 2012 and subsequent U.S. and State of Ohio Action Plans of 2018, new targets for TP and DRP loading have been set for Lake Erie.

Algal blooms in the Western Lake Erie Basin (WLEB) are strongly correlated to P loads discharged from the beginning of March through the end of July (Stumpf et al., 2016). Across WLEB tributaries, the greatest load of P is delivered by the Maumee River because of its large size and agricultural land use (Maccoux et al., 2016). The targets for the March–July loads from the Maumee River watershed have been set to 860 metric tons of TP and 186 metric tons of DRP (as quantified at the Waterville, Ohio gage station) (IJC 2012; US EPA 2018; OLEC 2018). Because nonpoint source pollution may exceed the targets in years with unusually high rainfall (Michalak et al., 2013; Stow et al., 2015), the GLWQA specified that these reduction targets should be met in 9-out-of-10 years. Evaluating the effectiveness of management in wetter than normal years can also be done by comparing to flow-weighted mean concentrations (FWMCs), with targets set at 0.23 mg L^{-1} for TP and 0.05 mg L^{-1} for DRP (US EPA 2018). About 88% of annual TP loads from this basin are derived from nonpoint sources, with agriculture being the dominant (79%) land use (Ohio EPA 2016). Because of this land use composition, reaching the March–July loading targets will require a strategy focused on implementing in-field and edge-of-field agricultural management practices.

Watershed simulation models offer the opportunity to evaluate watershed-scale effects of agricultural management practices under varying levels of adoption (Wilson et al., 2018). While considerable knowledge has been gained about nutrient flux and management from edge-of-field studies in the Maumee River watershed (Hanrahan et al., 2020; Pease et al., 2018a, 2018b; Williams et al., 2018), watershed models are needed to extrapolate these field-level results across the basin (Watson et al., 2016). In the WLEB, models had shown it was likely that multiple conservation practices would be needed in combination to reach the proposed water quality targets (Bosch et al., 2013; Muenich et al., 2016). Specifically, reduced fertilizer application or converting from conventional crops to biofuel grasses approached TP and DRP loading targets on their own, both cover crops and riparian buffers were moderately effective at reducing TP but not DRP loads (Muenich et al., 2016). An ensemble modeling approach was previously used in this watershed to evaluate combinations of practices, revealing multiple management strategies that could reach the P loading targets over an average of several years (Scavia et al., 2017). However, this previous study did not include some contemporary management practices in baseline models and full agricultural inputs (for example, most models did not include manure inputs); nor did it determine whether the loading targets were likely to be met in 9-out-of-10 years, or focus on effects of total nitrogen (TN) loading (Scavia et al., 2017). Watershed models have helped to spatially optimize the placement of conservation practices in order to ensure cost-effective management (Gaddis et al., 2014; Kalcic et al., 2015). Physical variations across the watershed (*e.g.* topography, land use, and soil type) have been shown to affect the efficiency of different practices while also influencing the location of critical source areas in models (Dai et al., 2018; Gaddis et al., 2014; Geng et al., 2015; Kalcic et al., 2015). Past modeling analyses in the WLEB have shown that the efficiency of practices can be improved by targeting them to the critical source areas rather than using a randomized placement approach (Muenich et al., 2016; Scavia et al., 2017).

Similar to the past study, we used ensemble modeling to better identify areas of consistency and variability in the model results. Ensemble modeling has been used for establishing target loads for the Great Lakes (Scavia et al., 2016a) and other water bodies (Bierman, 1979), but it has rarely been applied to assess land management scenarios. Viewing water quality issues from the different perspectives of multiple models reduces the decision risk by providing multiple lines of evidence (Scavia et al., 2017). To further enhance this study, we engaged a broad stakeholder group to increase the accuracy of the inputs used in the watershed models and guide the development and interpretation of realistic management scenarios. The goal of this study was to use this expertise to refine the types and extent of agricultural management practices (individual and bundles) needed to reach P load and FWMC targets using an ensemble modeling approach.

2. Methods

2.1. Study area and model development

The 17,000 km^2 Maumee River watershed area spans the States of Indiana, Michigan, and Ohio, with its outlet in the WLEB (Fig. 1). Nutrient loads have been monitored and estimated at the U.S. Geological Survey (USGS) gage #04193500 (USGS, 2020) since the 1970s (Baker et al., 2014). We collaborated with a stakeholder group (SI Table 1) to guide model development, and selection and interpretation of management scenarios. Similar to other projects (Kalcic et al., 2016; Scavia et al., 2017), a broad stakeholder group from a variety of organizations was developed for this project to ensure the results were relevant and informative for policy (Garb et al., 2008). The resulting stakeholder group included representation from agricultural groups, regulatory and non-regulatory government agencies, non-governmental organizations, and environmental groups. Identifying limitations of previous modeling work in the Maumee River watershed with these stakeholders led to a better representation of (1) stormwater discharges, (2) manure and synthetic fertilizer applications, (3) inclusion of existing best management practices (BMPs) in the models (SI Table 2), and (4) a more accurate comparison to the GLWQA goals.

Five Soil and Water Assessment Tool (SWAT; USDA, 2020) models were independently developed for the basin by Ohio State University (OSU), LimnoTech, Inc. (LT), University of Michigan (UM), Heidelberg University (HU), and University of Toledo (UT). All of the models in the current study are also described in a public report (Martin et al., 2019). SWAT has been widely used in agricultural watersheds because of its detailed representation of many agricultural management practices and simulation of nutrient runoff and riverine loading (Gebremariam et al., 2014; Gildow et al., 2016). The five models were selected based on their availability and their inclusion in previous modeling efforts of the Maumee River watershed. For example, previous versions of four of the models (OSU, LT, UM, and HU) had been used and described in a previous ensemble modeling study (Scavia et al., 2016b, 2017). Here, the UT model was added to the ensemble to replace fifth model from the Texas A&M University. Details about the UT model have also been described (Cousino et al., 2015). The SWAT models allowed predictions of all variables of interest (streamflow (Q), TP, DRP, and TN) as well as the implementation of practices of interest. A sixth model – the USGS

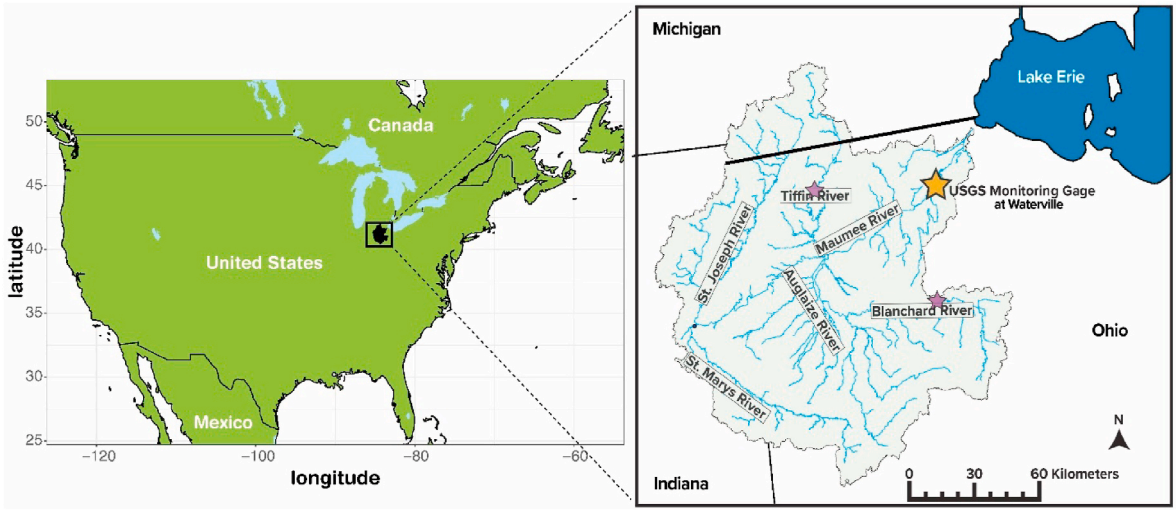

Fig. 1. This study was located in the Western Lake Erie Basin situated between the U.S. and Canada. A USGS gage (#04193500; yellow star) has measured discharge and Heidelberg University has measured TP, DRP, and TN concentrations at Waterville, Ohio for over 40 years. All models were calibrated and validated using the TP, DRP, and TN loads calculated at this gage. The OSU model was also calibrated to these nutrient loads measured in the Tiffin and Blanchard Rivers (USGS gages #04185000 and #04189000, respectively; purple stars). USGS discharge data are available at https://doi.org/10.5066/F7P55KJN. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

SPARROW model – was considered in the modeling work (Evenson et al., 2020; Scavia et al., 2016b), but this model was not able to estimate DRP loads or simulate the BMP processes.

While all five models use the same SWAT framework, they were distinct because of the independent decisions made by each modeling group regarding spatial resolution, data sources, model subroutines, land management operations, and model parameterization and calibration (SI Tables 2–4). A range of model structures resulted from choosing different algorithms related to tile drainage routines, water table routines, soil P models, and evapotranspiration methods (SI Table 2). SWAT model versions ranged from 627 to 664 (https://swat.tamu.edu/). Only the UT model used the old tile drain routine; the UT, HU, and LT models used the old water table and soil P modeling routines; and the LT model used the Hargreaves evapotranspiration method (Hargreaves and Samani, 1985), while all others used the Penman-Monteith equation (Monteith 1965). The models also used different sources of input for land use, elevation, and soils data. For example, the OSU and HU models used the Cropland Data Layer (CDL) 2007–2012 database for land use (https://nassgeodata.gmu.edu/CropScape), while the UT and LT models used the 2001 National Land Cover database (NLCD; https://www.mrlc.gov/), and the UM model used a combination of the CDL 2007–2012 and NLCD 2006 databases. The UM model used the National Aerodynamics and Space Administration/Infrared Processing and Analysis Center Extragalactic database (NED) 10m elevation model whereas all other models used NED 30m (https://ned.ipac.caltech.edu/). The LT model used State Soil Geographic dataset (STATSGO; https://data.nal.usda.gov/) for soils data, while all other models used the Soil Survey Geographic database (SSURGO; https://data.nal.usda.gov/). More structural uncertainty was added by the spatial discretization methods, which led to average hydrologic response unit (HRU) areas of 1.1–127 km^2 across the five models. The LT and UM models had the finest resolution within HRU areas (1.1–1.7 km^2); the OSU model was in between the others (11.3 km^2); and the UT and HU models were coarsest resolution (77–127 km^2). During calibration, as many as 75 parameters were altered leading to many differences between the five models (SI Table 3). The effects this variability had on modeling outputs have been discussed in further detail (Evenson et al., 2020). Together these differences create an ensemble of models that are representative of the variation in processes occurring in the watersheds.

General guidelines were provided to each modeling team on the simulation of practices, the inclusion of manure and point sources, and meteorological inputs. However, how much area existing management practices covered varied across the baseline models. For example, cover crops were implemented in 6.0% of cropland area in the HU model; 7.5% of area in the LT model; 8.4% for the OSU and UM models; and 10% for the UT model. Similarly, the distribution and extent of subsurface P application, vegetated riparian buffers, manure application, subsurface drainage also varied across the five models. To control for some input uncertainty, all models used the same temperature and precipitation data retrieved from the NOAA National Climatic Data Center and processed as previously described (Scavia et al., 2016b).

Compared to the guidance for baseline models, even stricter guidelines were given for the simulation of the wetland and controlled drainage practices in the modeling scenarios. All wetlands were sized at 250 acres with a depth of 3 feet using the default nutrient treatment parameters. Controlled drainage modified the tile drain depth to 0.91m below the ground surface, 0.49m over the summer, and 0.3m over the winter. Thus, differences in the effectiveness of these practices across models arose due to other factors, such as where they were implemented in the watershed.

As in previous work (Scavia et al., 2017), we accounted for point sources using the U.S. Environmental Protection Agency (US EPA) Discharge Monitoring Report Water Pollution Search (Scavia et al., 2017; US EPA 2016). To improve upon this in the current study, we also included the discharge data for combined sewer overflows (CSOs) to improve representation of point sources (US EPA 2016). Linear regression models were developed between nearby precipitation amounts with the CSO discharge and used to estimate missing data. Where the regression models were considered poor ($R^2 < 0.5$), we instead used the gap filling method described for point sources (Scavia et al., 2017).

To address further limitations of the past study where only two of the five models included manure applications (Scavia et al., 2017), present simulations included manure applications to row crops in all five models. Data sources included the 2012 Agricultural Census30 (https://www.nass.usda.gov/Publications/AgCensus/2012/) in conjunction with methods described by Ruddy et al. (2006); and the Nutrient Use Geographic Information System (NuGIS; Fixen et al., 2012). However, the five models varied in: the amount of cropland which received manure (9–14%), the spatial distribution (i.e. amount per county), and the timing of application. More accurate spatial application of manure was guided by new data contributed by stakeholders. There were also

differences across the five models in the type of livestock manure applied. All models continued to apply other inorganic fertilizers based on the county-level sales data (Ruddy et al., 2006) as previously described (Scavia et al., 2016b).

2.2. Baseline models and validation

All models were calibrated and validated using observed flows, suspended solids, TP, DRP, and TN loading data (NCWQR 2016; USGS 2020) collected at the USGS gage #04193500 near Waterville, Ohio. The OSU model also used this same data from two other USGS gages along the Tiffin and Blanchard Rivers (#04185000 and #04189000; e.g., Fig. 1). All models used the same weather data and were run over the 2005–2014 period. On average, the five baseline model simulations were initialized with the following cropland coverages: cover crop implementation on 8% of the agricultural land (range: 6–10%); vegetated, riparian buffer strips on 31% (29–34%); subsurface drainage on 75% (72–77%); continuous no-tillage on 32% (22–37%); and seasonal no-tillage (no-till on at least one crop, and no crop is conventionally tilled [Soil Tillage Intensity Rating >80]) on 42% (22–63%) (SI Table 4). As in a previous study (Scavia et al., 2017), the model results were standardized by multiplying the percent change from each model's baseline by the average observed load at the Waterville gage station during 2005–2014 to better compare our predicted changes in loading to the watershed targets. Model validation was evaluated using common goodness-of-fit metrics (Moriasi et al., 2007), including Nash-Sutcliffe efficiency (NSE) and percent bias (PBIAS), for monthly discharge (of water) and loads of TP and DRP.

The locations of critical source areas (CSAs) varied significantly across the models (Evenson et al., 2020). For TP, DRP, and TN, the models had an average percent agreement of 22%, 21%, and 27%, respectively. Some of the uncertainty in the location of CSAs was related to input parameters including where greater rates of fertilizers were applied. However, there was some variation that was undescribed and likely introduced due to structural and parameterization uncertainty (Evenson et al., 2020). The fact that these locations differed between models meant that the spatial distributions of practices in this modeling study varied, especially when practices were targeted to the CSAs. While these models and scenarios were evaluated at the watershed scale, related work has also investigated the impact of management scenarios at the field scale (Apostel et al. *In Press*).

2.3. Sensitivity scenarios

Initially, 13 conservation practices were simulated to guide the selection of practices that were subsequently used in bundled management scenarios (Table 1). The sensitivity scenarios used unrealistic adoption rates (i.e. 100%) to identify the most effective practices across the five models, including source contributions (1–3) and management effects (4–13). Most management actions that were tested altered the timing and delivery method of fertilizers. We also tested the independent effects of cover crops, controlled drainage, and wetlands using standardized operations across the five models (SI Table 4). For example, controlled drainage was programmed to permit free drainage during planting and harvesting; while holding subsurface water levels at 0.5 m during summer; and 0.3 m below the surface during winter (Pease et al., 2018a). The effects of each practice were evaluated in terms of percent reduction in the March–July TP and DRP loads and FWMC concentrations compared to the baseline model results.

2.4. Bundled management scenarios

Following the preliminary analysis of the sensitivity scenarios, discussions with the stakeholder group were used to develop reasonable bundled management scenarios. Based on results from the sensitivity scenarios and previous modeling work (Scavia et al., 2016b, 2017), it was determined that subsurface P application, cover crops, and headwater wetlands should be considered in the final set of scenarios because of their potential to reduce P loads. Vegetated riparian buffers were also included based upon results of past modeling and experimental evidence of their benefits (Hoffmann et al., 2009; Scavia et al., 2017). Another practice, continuous no-tillage, was chosen in one suite of practices because of known benefits to soil health (Congreves et al., 2015; Mitchell et al., 2017). Although reduced fertilizer application resulted in decreased P loads, this option was not considered actionable by our stakeholders. In addition to these practices, stakeholders had strong interest in further examining the effects of controlled drainage, despite resulting in increased P loads in the sensitivity scenarios. Many of the selected practices are also gaining acceptance by farmers across the watershed (Wilson et al., 2018, 2019).

These selected practices were combined into five bundled management scenarios (Scenarios 14–18) that were evaluated across the Maumee River watershed (Table 1). Scenario 14 tested the random implementation of the bundled practices. Because previous studies have

Table 1
Sensitivity scenarios and bundled scenarios included in the multi-model simulations.

	Identification		Description	Purpose
Sensitivity Scenarios	1	No point source discharge	100% point source removal	source contribution
	2	No manure application	100% removal of P from manure (baseline N application maintained)	source contribution
	3	25% P rate reduction	25% less P fertilizer applied	source contribution
	4	Broadcast P	All fertilizer is broadcast without incorporation	management effect
	5	Broadcast and incorporated P	All fertilizer is broadcast with incorporation	management effect
	6	Subsurface applied P	All P fertilizer is subsurface applied (99% below top 1 cm of soil)	management effect
	7	Fall manure	All manure applied in fall	management effect
	8	Spring manure	All manure applied in spring	management effect
	9	Fall and spring manure	Manure applied half in spring and half in fall	management effect
	10	Rate, placement, and timing	Fertilizer and manure P applications decreased by 50% and subsurface applied P in the fall	management effect
	11	Cereal rye cover crop	Cover crops on 100% of cropland	management effect
	12	Controlled drainage	Controlled drainage (or drainage water management) in 100% of tile-drained areas by	management effect
	13	Headwater wetlands	Wetlands in all subbasins receive 50% of flow	management effect
Bundled Management Scenarios	14	In-field + Buffers (Random)	58% cover crops + 50% subsurface placement of P + 78% buffer strips	scenario, random implementation
	15	In-field + Buffers	58% cover crops + 50% subsurface placement of P + 78% buffer strips	scenario, targeted implementation
	16	Likely Adoption of In-field + Buffers	60% cover crops + 68% subsurface placement of P + 50% buffer strips	scenario, survey data, targeted implementation
	17	In-field + Wetlands	58% cover crops + 50% subsurface placement of P + 78% wetlands	scenario, targeted implementation
	18	In-field + Controlled drainage	50% cover crops + 60% subsurface placement of P + 50% no-tillage + 15% controlled drainage	scenario, targeted implementation

documented better performance with the targeted implementation of management practices in areas with the greatest potential for P export (Muenich et al., 2016), modelers and stakeholders agreed that only targeted options would be evaluated for the remaining bundled Scenarios (15–18). A likely-adoption scenario (Scenario 16) was guided by survey data for cover crops and subsurface placement (Wilson et al., 2018) with riparian buffer adoption set to 50%. Despite the survey showing 78% of farmers were willing to adopt a riparian buffer (Wilson et al., 2018), our stakeholder group perceived this level of adoption was unlikely because there would be resistance to converting so much arable land to buffers. The final two scenarios replaced buffer strips from Scenario 15 with wetlands (Scenario 17) or controlled drainage (Scenario 18). The stakeholder group agreed that while achieving this level of adoption would be cumbersome, these scenarios represented actionable paths forward.

3. Results and discussion

3.1. Model calibration and validation

The calibrated multi-model averages were similar to the measured March–July loads of TP and DRP. The only bias identified was an over prediction of DRP (+42%) at extremely low loads (<85 metric tons) by some models (Fig. 2). Baseline model evaluation indices at the monitored outlet location were generally within recommended levels for SWAT model validation (Table 2) (Arnold et al., 2015; Moriasi et al., 2007). In most cases, the model exceeded the satisfactory range and could be considered good to very-good according to the standards of Moriasi et al. (2007). These criteria indicated that the UT model had the lowest efficiency in predicting DRP and TN loads. While this model was within the satisfactory range for DRP loads, it performed unsatisfactorily in predicting TN loads. Because TN changes were only considered in relation to TP within this study, this model was still included in the full analysis. To further evaluate the models' performance, we completed "soft validation" (Arnold et al., 2015) and found general agreement when comparing modeled results to water budget and crop yield data from field monitoring (SI Table 5). Similar to past ensemble modeling (Scavia et al., 2017), selecting one model from this group based on better performance may be appealing but all models reasonably represented the baseline conditions.

3.2. Sensitivity scenarios

The advantage of using an ensemble modeling approach was that it showed which practices had the greatest and most consistent effect in multiple models, and it also described the agreement and variability in our predictions. For sensitivity scenarios, we focused on the response of P concentrations to changes in each management practice because we

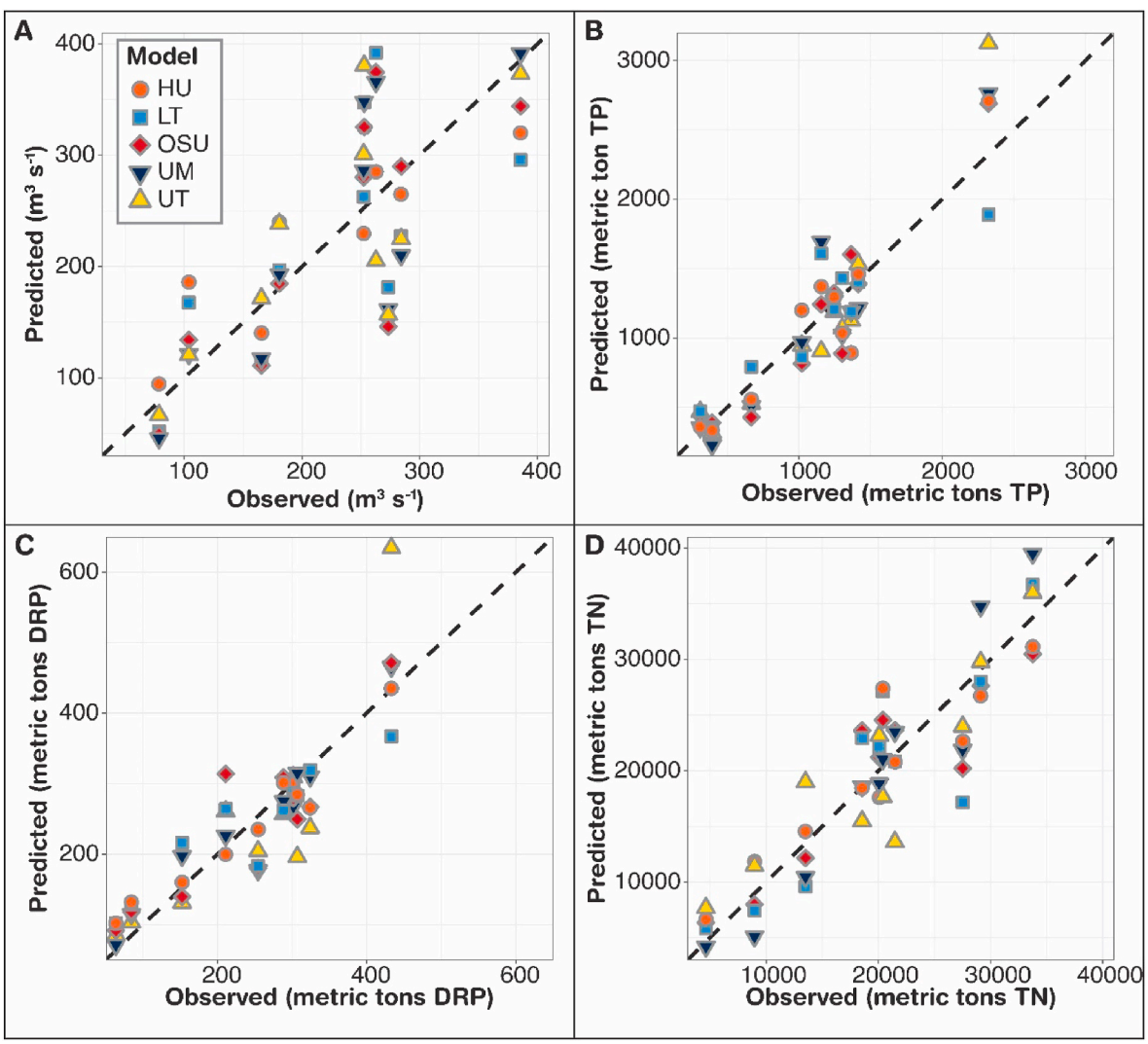

Fig. 2. Models were calibrated against the March–July flow rate and nutrient loads. The predicted values are plotted against the observed values for the five models for (**A**) average daily flow rate; (**B**) TP load; (**C**) DRP load; and (**D**) TN load. The 1:1 line is shown as the dashed line.

Table 2
The performance of the five SWAT models were evaluated over the entire 2005–2014 period and were compared to standards for satisfactory performance established by Moriasi et al. (2007)[1] for percent bias (PBIAS) and Nash-Sutcliffe Efficiency (NSE).

		Satisfactory Performance Range[1]	Multi-Model Average	Ohio State University	LimnoTech	University of Michigan	Heidelberg University	University of Toledo
PBIAS (%)	Discharge	+/- 25	2.2	−3	11	1	2	0.1
	TP	+/- 70	−2.7	19	−13	1	−7	−13
	DRP	+/- 70	5	−4	−15	7	7	32
	TN	+/- 70	−11	−11	−24	−4	−3	−12
NSE	Discharge	>0.50	0.89	0.99	0.91	0.94	0.88	0.83
	TP		0.70	0.71	0.77	0.61	0.73	0.66
	DRP		0.67	0.73	0.67	0.69	0.77	0.50
	TN		0.58	0.64	0.59	0.77	0.74	0.17

found that the relative change was similar to those observed for loads (Martin et al., 2019; SI Table 7). All five models showed that fertilizer rate reductions, subsurface placement, cover crops, and headwater wetlands (Scenarios 3, 6, 10, 11, and 13) had the largest reductions in TP concentrations and loads (Fig. 3; SI Table 6). There was also agreement across the models that DRP concentrations and loads decrease as a result of these practices, with the exception of the use of headwater wetlands. However, altering the timing of fertilizer application (Scenarios 7–9) or changing all application to broadcast and incorporation had inconsistent effects and either increased or decreased P concentrations. Two scenarios increased TP and DRP concentrations and loads across all models: broadcasting of all fertilizers (Scenario 4) and controlled drainage (Scenario 12; SI Table 7). Several other studies have shown that controlled drainage has minimal effects on TP and DRP concentrations, while load reductions have been realized because of decreases in discharge volume (Pease et al., 2017; Sunohara et al., 2015; Williams et al., 2015).

To further evaluate the effect of each management practice, as well as the variability among the five models, we computed the change in TP and DRP concentrations for each practice. Across the five models, the range of effectiveness for each of these management practices was 16% for TP and 24% for DRP on average. For example, the reduction of DRP concentrations when cover crops were implemented ranged between −1.3% and −11.4%, or a range of 10.1%. The management practices that had the greatest variability among models included headwater wetlands and subsurface application (Scenarios 6, 10, and 13). In the case of subsurface placement, one model predicted TP and DRP concentrations decreased by 44% and 67%, respectively, which deviated greatly from the average reductions of 18% and 22.8%. While there was much variability in the extent of reductions, there was consistency in the absolute and relative effect, as all models agreed that subsurface placement had the greatest potential to reduce TP and DRP concentrations and loads. Rainfall simulation studies have shown that subsurface application of fertilizer can reduce P loss by as much as 66–98% (Smith et al., 2016; Williams et al., 2018), suggesting that the large reduction predicted by the one model may be realistic. Notably, the variability in predictions may partially originate from differences in the initialization of the baseline conditions, such as the assumed initial extent of subsurface placement across the watershed, in addition to differences in processes and other parameters of the models (SI Tables 2–4; Evenson et al., 2020).

Removing P from manure and point sources yielded similar effects on March–July loads to Lake Erie. On average, TP decreased by 5.7% by removing point sources and 7.2% by removing manure. The changes to DRP were a 10.4% decrease by removing point sources and an 7.7% decrease by removing manure. Despite modeling improvements to account for CSOs, similar point source contributions of TP and DRP were found in previous work (Scavia et al., 2017). Our results provided estimates of the contribution of manure P fluxes in the Maumee River, which were not estimated in the previous work (Scavia et al., 2017). This is an important advancement because studies in the Maumee River watershed have indicated manure application to be an substantial source of P loading in this watershed (Kast et al., 2019; Robertson and Saad, 2011). Here, we showed that eliminating point sources, manure, or the combination of the two entirely from the watershed would be insufficient to reach the established P reduction goals.

3.3. Bundled-management scenarios

Evaluating the number of years that the TP loading target was met by the average of the five models revealed that the 9-out-of-10-year goal was not met by any scenario. In fact, the average number of years that the TP load target was met never exceeded 8-out-of-10 years. However, the average number of years the TP loading target was met was at least 6.8 years for three scenarios (Scenarios 15–17) and was never less than 5.8 years (Scenario 14; Fig. 4A). The average result for TP in the targeted placement of cover crops, riparian buffers, and subsurface placement (Scenario 15) was 8-out-of-10 years. This was an improvement of 2.2 years from when these same BMPs were placed randomly across the watershed (Scenario 15 vs. 14). Similar to a previous study (Scavia et al., 2017), these results highlight that even with widespread adoption rates, water quality benefits can be accelerated and resources used more efficiently when practices are placed in areas contributing the greatest amount of P.

The agreement among model results provide confidence in the finding that the bundled-management scenarios are more likely to reach the TP target than the DRP target. For every bundled scenario, the average prediction never exceeded 5.4 years in meeting the DRP load target, meaning that the targets would only be met in about 2.2 additional years compared to the baseline. Results from the five models varied by 3–6 years in terms of how frequently the TP and DRP targets were met under any scenario. The greatest variation in meeting the targets occurred in the scenario with wetlands (Scenario 17) where the range was 7.2 years for both TP and DRP. Alternatively, there was a smaller range for the "likely-adoption" scenario (Scenario 16) with variance of 2.3 and 1.7 years for TP and DRP, respectively.

Using the FWMC benchmarks provide greater clarity in the effectiveness of the various management strategies, compared to 9-out-of-10-year targets because FWMC removed the effects of the highly variable flows. For FWMC, two scenarios (Scenarios 15 and 16) on average met the TP concentration target, while DRP FWMC concentrations still exceeded the target by 20–30% in those scenarios (Fig. 4B, SI Table 6). Although no scenarios were sufficient to meet the FWMC goal of 0.05 mg L^{-1} for DRP (based on the average of model outputs), four of the bundled-management scenarios made progress toward this goal by reducing DRP concentrations from 0.08 mg L^{-1} (observed) to ≤0.065 mg L^{-1}. The largest and most consistent reductions were found in scenarios 15 and 16, both of which targeted the implementation of subsurface placement, cover crops, and riparian buffers. The range for these two scenarios were all <1% of the model averages for FWMC for both TP and DRP. Predicted percent reductions in FWMC ranged from 24% to 53% for TP (average 36%) and 22%–48% for DRP (average 35%) across the bundled scenarios among the five models. The greatest variability for decreases in TP FWMC was in Scenario 17, while the greatest

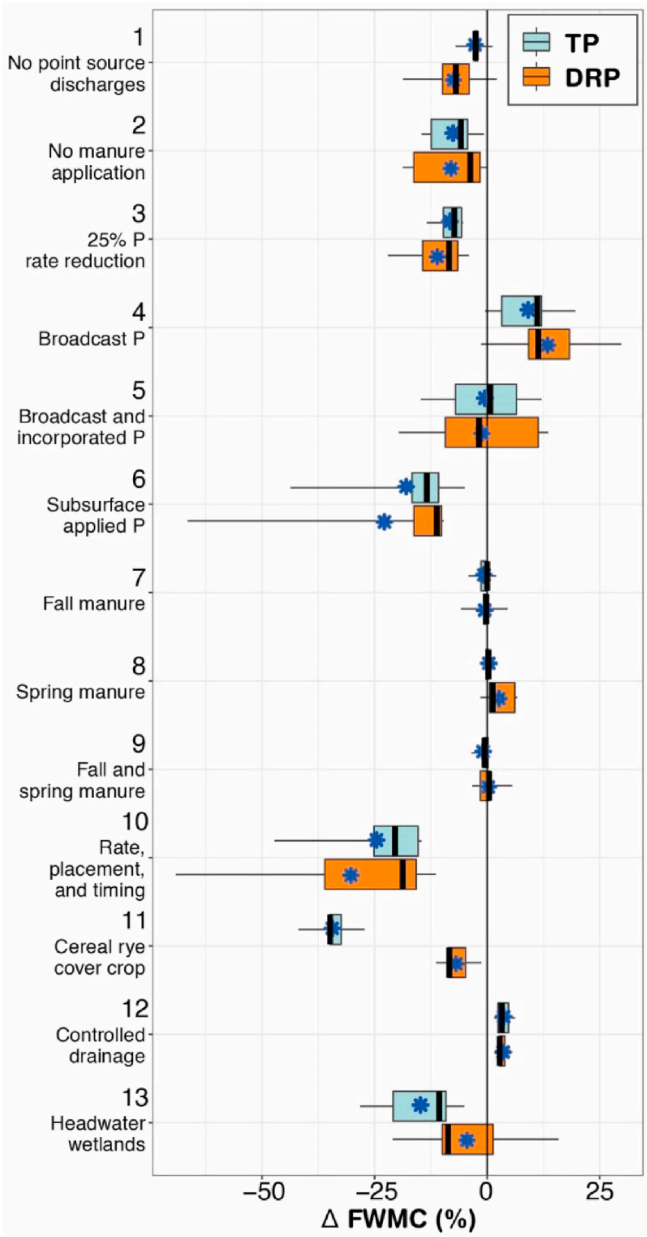

Fig. 3. Changes in March–July flow-weighted mean concentrations (FWMC) for both TP and DRP for each sensitivity scenario. Results from the five models were averaged across the 2005–2014 period (n = 5 for each box). The blue asterisks (*) represent the average, vertical line inside the box represents the median, edges of the box denote the 25th and 75th percentiles and the whiskers identify the minimum and maximum values of model results. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

variability in DRP FWMC occurred in Scenario 18. Our results highlight the need to generate more edge-of-field data for the simulated practices – particularly, subsurface drainage and wetlands – against which models can be compared.

While P reduction goals have been set to reduce the size of algal blooms, nitrogen loading also plays an important role in determining the toxicity of these blooms (Gobler et al., 2016; Newell et al., 2019). Thus, we considered the effects that the management scenarios had on TN:TP loading ratios (Fig. 4C, SI Table 6). In all cases, implementation of the scenarios increased the TN:TP ratios from the baseline scenario because the BMPs were more effective at reducing P rather than nitrogen. The uncertainty in TN:TP loading ratios generally followed the pattern observed in TP loading, with the greatest variability in Scenario 17. The magnitude of this change was smallest for the controlled drainage scenario (Scenario 18) because this practice increased anaerobic soil zones with the potential to reduce nitrogen loads through denitrification (Sunohara et al., 2015). The relation between TN:TP loading ratios and algal growth or toxin-production is complex and dynamic (Davis et al., 2015; Downing et al., 2005; Lee, S. J. et al., 2000). Therefore, it is difficult to explicitly state how the simulated management efforts would affect the toxicity of the blooms, but these results predict that focusing management action on P reduction could shift the Lake Erie stoichiometric balance towards more available nitrogen relative to P.

For the first two bundled management scenarios (Scenarios 14 and 15), the LT model showed the greatest response in TP reduction with 8 and 10 years meeting the water quality targets (SI Table 7). The greatest magnitude of change by altering to a targeted approach occurred for the UT model which went from predicting 3 years–7 years in which the target was met. Only the OSU model showed no change when the practices were targeted in terms of years that the targets were met. In fact, the OSU model predicted that the TP target would be met in 7-out-of-10 years for all of these scenarios except when wetlands (Scenario 17) were introduced leading to meeting the target in 9-out-of-10 years. For TP, the HU model never predicted that the loading target would be met in more than 5-out-of-10 years except in the case of the likely-adoption scenario (Scenario 16)

The UM model consistently yielded the greatest reduction of DRP loads, reaching the DRP reduction goals for Scenarios 15, 17, and 18, while this reduction was often less pronounced in the HU outputs. This can largely be explained by the fact that subsurface placement had greatest effects in the UM model, while there was less of an effect of subsurface placement and riparian buffers in the HU model (SI Table 6). The patterns that emerged from individual models regarding TP loads were also found with DRP loads. The UT model showed a 3-year increase in the number of years that the target would be met by altering the in-field & buffers practices to a targeted approach (Scenarios 14 and 15). Once again, the OSU model demonstrated no additional benefit from targeting in terms of the years that the target was met.

Generally, the UT and HU models predicted that meeting the TP and DRP loading targets was less likely than other models. Alternatively, the greatest effect was often observed with the UM and LT. It is worth recalling that the UM and LT models had greater resolution in terms of HRU size compared to the UT and HU models that might be associated with this greater performance. However, the UM model showed greater P load reduction with subsurface placement while the LT model had greater results with cover crops (SI Tables 5 and 6). Across the watershed, there is a lack of data to show how effective the simulated BMPs can be. This raises a further point as to why selecting the best-performing model based on validation metrics (*i.e.* PBIAS and NSE) alone is not advisable. Much of the uncertainty in the effectiveness of practices in these five models arose from how effective these practices are, and multiple models can be used to identify the range of possibilities in meeting water quality goals.

Previous modeling of the Maumee River watershed showed that management scenarios were capable of reaching the TP and DRP water quality targets based on the average load over several years (Scavia et al., 2017). Based on results from this study, we found that no scenario average was capable of reaching the targets in 9-out-of-10 years. Results of individual models, however, did predict concentrations and load reductions that did meet these goals. This results in the conclusion that to reach the established P reduction targets will require rates of adoption greater or equal to those included in these scenarios. A concerning finding was that implementation of most scenarios would make less progress in meeting the DRP loading targets compared to TP. It will be crucial to find new technologies that are more effective at reducing DRP because most of the practices examined here had a greater effect on reducing particulate P than dissolved P. Whereas only a portion of particulate P is bioavailable (~26%), DRP is entirely bioavailable to the

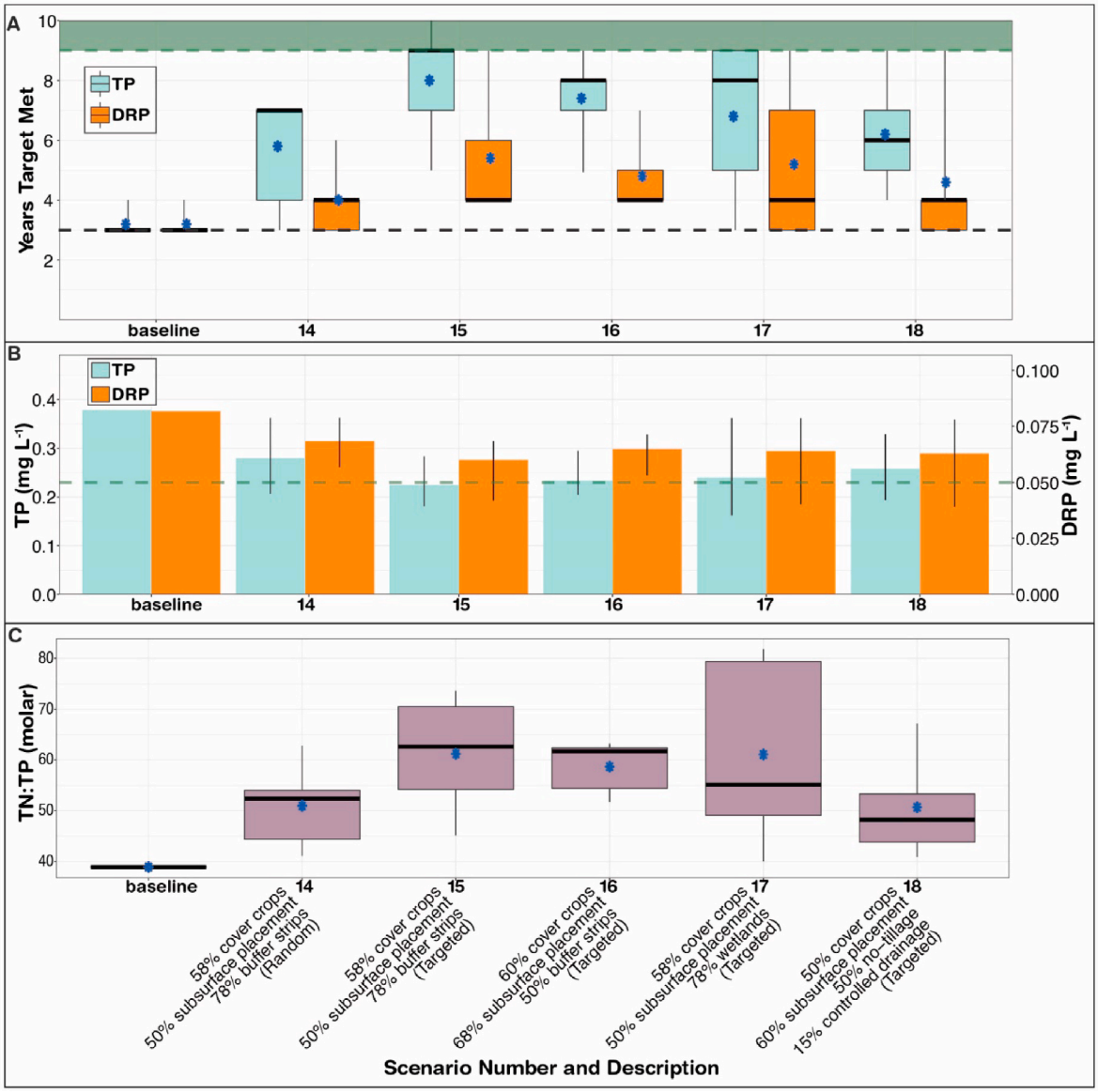

Fig. 4. Results of bundled-management scenarios simulated from 2005 to 2014. (**A**) Loads predicted within each scenario compared to the March–July loading targets for TP (860 metric tons) and DRP (186 metric tons). The blue asterisks (*) represent the average, the black line represents the median, the box is the interquartile range, and the whiskers show the range of model results. The green dashed line and shaded area represent when the 9-out-of-10 year target would have been reached. The black dashed-line represents the number of years that the TP and DRP loading targets had been met during the simulated period based upon observed data. (**B**) Effects of the management practices on March–July FWMC for TP and DRP. Bar height represents the average and whiskers show the range calculated across the five models. The green-dashed line and shaded area represent the targets set for TP and DRP concentrations are 0.23 and 0.05 mg L^{-1}, respectively. (**C**) March–July total nitrogen to total phosphorus ratio (TN:TP) calculated on a molar basis. The blue asterisks (*) represent the average, the dark black line represents the median, the box is the interquartile range, and the whiskers show the range of model results. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

algal bloom (Stumpf et al., 2016). Ideally, such practices would be capable of simultaneously reducing both forms of P and nitrogen loads.

4. Conclusions

Our results demonstrate the importance of integrated studies involving ensemble models, behavioral science and stakeholders in adaptive management approaches to address coastal eutrophication (Zedler, 2017). Advances in technology, modeling routines, and data collection have led to improved models, and these improved models could be used to simulate scenarios derived through stakeholder engagement. Ensemble modeling, possibly also including other types of watershed models, is a useful approach to gaining assurance in the types and extent of practices required to reach water quality targets. Such assurance can help alleviate some of the risks in policy-making decisions that use modeling results to guide implementation of BMPs.

There was agreement across all models that practices, such as fertilizer rate reductions, subsurface placement of P, cover crops, and headwater wetlands, would reduce P loads and concentrations. However, due to variation in P reductions found among models, there was uncertainty about the degree of reductions for some of these practices. For example, there was general agreement that subsurface placement of fertilizers and manure resulted in larger reductions in FWMC of DRP providing assurance that it would yield the greatest declines in DRP loads. While the performance of bundled management scenarios varied across models, the average output across these models were consistently

less than the 9-out-of-10-year goal indicating that rates of adoption beyond these scenarios are needed. It was also clear the targeted application was preferable to random application of management practices, that the tested practices produced greater reductions in TP than DRP, and that these practices would reduce P at greater rates than nitrogen. Variability in these modeling results is a key benefit of using an ensemble modeling approach to evaluate the potential of management practices. While this effort substantially improved upon and addressed limitations compared to previous work (Scavia et al., 2017), there was consistency in regard to the limited contributions from point sources and manure. As with the previous model, the TP FWMC target was met over the 10-year period by some scenarios, but this did not equate to the loading target being met in at least 9-out-of-10 years. This means there is no feasible path that can be followed to achieve the goals of the GLWQA.

Including a diverse set of stakeholders led to improved representation of P sources and existing management practices in our models. Further, we developed more feasible management scenarios with these partners and by using recent survey data suggesting maximum voluntary adoption rates of practices. With these improvements, model results demonstrated the challenge of meeting the 9-out-of-10-year P reduction goal. Further, stakeholders also aided in the production of a variety of outreach materials (http://kx.osu.edu/project/environment/habri-multi-model) that are available to the general public, thus developing and disseminating our findings for a wider audience.

Author statement

Jay F. Martin, Conceptualization, Methodology, Investigation, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition, Margaret M. Kalcic: Conceptualization, Methodology, Writing – review & editing, Supervision, Project administration, Funding acquisition, Noel Aloysius: Methodology, Software, Validation, Writing – review & editing, Anna M. Apostel: Methodology, Writing – review & editing, Michael R. Brooker, Writing – review & editing, Formal analysis, Data curation, Visualization. Grey Evenson: Methodology, Writing – review & editing. Jeffrey B. Kast, Methodology, Writing – review & editing. Haley Kujawa, Methodology, Data curation, Writing – review & editing. Asmita Murumkar, Methodology, Writing – review & editing. Richard Becker: Conceptualization, Methodology, Writing – review & editing. Chelsie Boles, Methodology, Writing – review & editing. Remegio Confesor: Conceptualization, Methodology, Writing – review & editing. Awoke Dagnew: Methodology, Writing – review & editing. Tian Guo: Methodology, Writing – review & editing. Colleen M. Long: Methodology, Writing – review & editing. Rebecca L. Muenich: Methodology, Writing – review & editing. Donald Scavia, Conceptualization, Writing – review & editing, Supervision, Project administration, Funding acquisition, Todd Redder: Conceptualization, Writing – review & editing, Dale M. Robertson: Conceptualization, Writing – review & editing, Project administration, Funding acquisition. Yu-Chen Wang, Methodology, Writing – review & 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

We wish to thank the individual members of the Stakeholder Advisory Group and the organizations they represented for participating in this project by providing guidance and participating in outreach activities. This work was funded by the Ohio Department of Higher Education Harmful Algal Bloom Research Initiative (R/HAB-5-ODHE) through Ohio Sea Grant. We would like to thank Dave Saad with the U.S. Geological Survey Upper Midwest Water Science Center and three anonymous reviewers of the Journal of Environmental Management for helpful comments on earlier drafts of this paper. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jenvman.2020.111710.

References

Apostel, A., Kalcic, M., Dagnew, A., Evenson, G., Kast, J., King, K., Martin, J., Muenich, R.L., Scavia, D., (in press). Simulating Internal Watershed Processes Using Multiple SWAT Models, Science of the Total Environment.

Arnold, J.G., Youssef, M.A., Yen, H., White, M.J., Sheshukov, A.Y., Sadeghi, A.M., Moriasi, D.N., Steiner, J.L., Amatya, D.M., Skaggs, R.W., 2015. Hydrological processes and model representation: impact of soft data on calibration. Transactions of the ASABE 58, 1637–1660.

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

Bierman, V.J., 1979. Comparison of models developed for phosphorus management in the Great lakes. Phosphorus management strategies for lakes 235–255.

Bosch, N.S., Allan, J.D., Selegean, J.P., Scavia, D., 2013. Scenario-testing of agricultural best management practices in Lake Erie watersheds. J. Great Lake. Res. 39, 429–436.

Bukaveckas, P.A., Lesutienė, J., Gasiūnaitė, Z.R., Ložys, L., Olenina, I., Pilkaitytė, R., Pūtys, Ž., Tassone, S., Wood, J., 2017. Microcystin in aquatic food webs of the Baltic and Chesapeake Bay regions. Estuar. Coast Shelf Sci. 191, 50–59.

Congreves, K.A., Hayes, A., Verhallen, E.A., Van Eerd, L.L., 2015. Long-term impact of tillage and crop rotation on soil health at four temperate agroecosystems. Soil Tillage Res. 152, 17–28.

Conley, D.J., Paerl, H.W., Howarth, R.W., Boesch, D.F., Seitzinger, S.P., Havens, K.E., Lancelot, C., Likens, G.E., 2009. Controlling eutrophication: nitrogen and phosphorus. Science 123, 1014–1015.

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.

Dai, C., Qin, X.S., Tan, Q., Guo, H.C., 2018. Optimizing best management practices for nutrient pollution control in a lake watershed under uncertainty. Ecol. Indicat. 92, 288–300.

Davis, T.W., Bullerjahn, G.S., Tuttle, T., McKay, R.M., Watson, S.B., 2015. Effects of increasing nitrogen and phosphorus concentrations on phytoplankton community growth and toxicity during Planktothrix blooms in Sandusky Bay, Lake Erie. Environ. Sci. Technol. 49, 7197–7207.

Downing, T.G., Sember, C.S., Gehringer, M.M., Leukes, W., 2005. Medium N: P ratios and specific growth rate comodulate microcystin and protein content in Microcystis aeruginosa PCC7806 and M. aeruginosa UV027, Microb. Ecol. 49, 468–473.

Evenson, G.R., Kalcic, M., Wang, Y., Robertson, D., Scavia, D., Martin, J., Apostel, A., Aloysius, N., Kast, J., Haley, Kujawam, Boles, C., Redder, T., Confesor, R., Guo, T., Muenich, R.L., Dagnew, A.T., Murumkar, A., Brooker, M., 2020. Uncertainty in critical source area predictions from watershed-scale hydrologic models. J. Environ. Manag. https://doi.org/10.1016/j.jenvman.2020.111506. In this issue.

Fixen, P.E., Williams, R., Rund, Q.B., 2012. NUGIS: a nutrient use geographic information system for the US. International Plant Nutrition Institute, Brookings, SD. http://www.ipni.net/ipniweb/portal.nsf/0/5D3B7DFAFC8C27688525774300 5AA07A/$FILE/1203%20FIXEN%20GPSFC%20NuGIS%20FINAL.pdf.

Gaddis, E.J.B., Voinov, A., Seppelt, R., Rizzo, D.M., 2014. Spatial optimization of best management practices to attain water quality targets. Water Resour. Manag. 28, 1485–1499.

Garb, Y., Pulver, S., VanDeveer, S.D., 2008. Scenarios in society, society in scenarios: toward a social scientific analysis of storyline-driven environmental modeling. Environ. Res. Lett. 3, 045015.

Gebremariam, S.Y., Martin, J.F., DeMarchi, C., Bosch, N.S., Confesor, R., Ludsin, S.A., 2014. A comprehensive approach to evaluating watershed models for predicting river flow regimes critical to downstream ecosystem services. Environ. Model. Software 61, 121–134.

Geng, R., Wang, X., Sharpley, A.N., Meng, F., 2015. Spatially-distributed cost–effectiveness analysis framework to control phosphorus from agricultural diffuse pollution. PloS One 10, e0130607.

Gildow, M., Aloysius, N., Gebremariam, S., Martin, J., 2016. Fertilizer placement and application timing as strategies to reduce phosphorus loading to Lake Erie. J. Great Lake. Res. 42, 1281–1288.

Gobler, C.J., Burkholder, J.M., Davis, T.W., Harke, M.J., Johengen, T., Stow, C.A., Van de Waal, Dedmer, B., 2016. The dual role of nitrogen supply in controlling the growth and toxicity of cyanobacterial blooms. Harmful Algae 54, 87–97.

Hanrahan, B.R., King, K.W., Macrae, M.L., Williams, M.R., Stinner, J.H., 2020. Among-site variability in environmental and management characteristics: effect on nutrient loss in agricultural tile drainage, J. Great Lakes Res.46, pp. 486–499.

Hargreaves, G.H., Samani, Z.A., 1985. Reference crop evapotranspiration from temperature. Appl. Eng. Agric. 1 (2), 96–99.

Hoffmann, C.C., Kjaergaard, C., Uusi-Kamppa, J., Hansen, H.C., Kronvang, B., 2009. Phosphorus retention in riparian buffers: review of their efficiency. J. Environ. Qual. 38, 1942–1955.

Horst, G.P., Sarnelle, O., White, J.D., Hamilton, S.K., Kaul, R.B., Bressie, J.D., 2014. Nitrogen availability increases the toxin quota of a harmful cyanobacterium. Microcystis aeruginosa, Water Research 54, 188–198.

International Joint Commission (IJC), 2012. Great lakes water quality agreement. https://binational.net//wp-content/uploads/2014/05/1094_Canada-USA-GLWQA-e.pdf.

Kalcic, M.M., Frankenberger, J., Chaubey, I., 2015. Spatial optimization of six conservation practices using SWAT in tile-drained agricultural watersheds. J. Am. Water Resour. Assoc. 51, 956–972.

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, 8135–8145.

Kast, J.B., Long, C.M., Muenich, R.L., Martin, J.F., Kalcic, M.M., 2019. Manure management at Ohio confined animal feeding facilities in the Maumee River watershed. J. Great Lake. Res. 45 (6), 1162–1170.

Lee, S.J., Jang, M., Kim, H., Yoon, B., Oh, H., 2000. Variation of microcystin content of Microcystis aeruginosa relative to medium N: P ratio and growth stage. J. Appl. Microbiol. 89, 323–329.

Lee, J., Lee, S., Jiang, X., 2017. Cyanobacterial toxins in freshwater and food: important sources of exposure to humans. Annual Review of Food Science and Technology 8, 281–304.

Maccoux, M.J., Dove, A., Backus, S.M., Dolan, D.M., 2016. Total and soluble reactive phosphorus loadings to Lake Erie: a detailed accounting by year, basin, country, and tributary. J. Great Lake. Res. 42, 1151–1165.

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., Meunich, R., Scavia, D., Wang, Y., Robertson, D., 2019. Evaluating management options to reduce Lake Erie algal blooms with models of the Maumee River watershed. The Ohio State University. https://kx-assets.s3.amazonaws.com/2020-02/Evaluating%20Management%20Options%20to%20Reduce%20Lake%20Erie%20Algal%20Blooms%20with%20Models%20of%20the%20Maumee%20River%20Watershed.pdf.

Michalak, A.M., Anderson, E.J., Beletsky, D., Boland, S., Bosch, N.S., Bridgeman, T.B., Chaffin, J.D., Cho, K., Confesor, R., Daloglu, I., Depinto, J.V., Evans, M.A., Fahnenstiel, G.L., He, L., Ho, J.C., Jenkins, L., Johengen, T.H., Kuo, K.C., Laporte, E., Liu, X., McWilliams, M.R., Moore, M.R., Posselt, D.J., Richards, R.P., Scavia, D., Steiner, A.L., Verhamme, E., Wright, D.M., Zagorski, M.A., 2013. Record-setting algal bloom in Lake Erie caused by agricultural and meteorological trends consistent with expected future conditions. Proc. Natl. Acad. Sci. U.S.A. 110, 6448–6452.

Mitchell, J.P., Shrestha, A., Mathesius, K., Scow, K.M., Southard, R.J., Haney, R.L., Schmidt, R., Munk, D.S., Horwath, W.R., 2017. Cover cropping and no-tillage improve soil health in an arid irrigated cropping system in California's San Joaquin Valley, USA. Soil Tillage Res. 165, 325–335.

Monteith, J.L., 1965. Evaporation and Environment. Symposia of the Society for Experimental Biology, vol. 19. Cambridge University Press (CUP, Cambridge.

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

Muenich, R.L., Kalcic, M., Scavia, D., 2016. Evaluating the impact of legacy P and agricultural conservation practices on nutrient loads from the Maumee River Watershed. Environ. Sci. Technol. 50, 8146–8154.

National Center for Water Quality Research. (NCWQR), 2016. Tributary loading program. https://ncwqr.org/monitoring/data/.

Newell, S.E., Davis, T.W., Johengen, T.H., Gossiaux, D., Burtner, A., Palladino, D., McCarthy, M.J., 2019. Reduced forms of nitrogen are a driver of non-nitrogen-fixing harmful cyanobacterial blooms and toxicity in Lake Erie. Harmful Algae 81, 86–93.

Ohio Environmental Protection Agency (Ohio EPA), 2016. Nutrient mass balance study for Ohio's major rivers. https://epa.ohio.gov/Portals/35/documents/Nutrient%20Mass%20Balance%20Study%202018_Final.pdf.

Ohio Lake Erie Commission (OLEC), 2018. State of Ohio's domestic action plan. https://lakeerie.ohio.gov/Portals/0/Ohio%20DAP/DAP%201-1%20FINAL%202018-08-27.pdf.

Paerl, H.W., Otten, T.G., 2013. Harmful cyanobacterial blooms: causes, consequences, and controls. Microb. Ecol. 65, 995–1010.

Pease, L.A., Fausey, N.R., Martin, J.F., Brown, L.C., 2017. Projected climate change effects on subsurface drainage and the performance of controlled drainage in the Western Lake Erie Basin. J. Soil Water Conserv. 72, 240–250.

Pease, L.A., Fausey, N.R., Martin, J.F., Brown, L.C., 2018a. Weather, landscape, and management effects on nitrate and soluble phosphorus concentrations in subsurface drainage in the Western Lake Erie basin. Transactions of the ASABE 61.

Pease, L.A., King, K.W., Williams, M.R., LaBarge, G.A., Duncan, E.W., Fausey, N.R., 2018b. Phosphorus export from artificially drained fields across the Eastern Corn Belt. J. Great Lake. Res. 44, 43–53.

Robertson, D.M., Saad, D.A., 2011. Nutrient inputs to the Laurentian Great Lakes by source and watershed estimated using SPARROW watershed models. J. Am. Water Resour. Assoc. 47, 1011–1033.

Ruddy, B.C., Lorenz, D.L., Mueller, D.K., 2006. County-level Estimates of Nutrient Inputs to the Land Surface of the Conterminous United States, 1982-2001.

Scavia, D., Allan, J.D., Arend, K.K., Bartell, S.M., Beletsky, D., Bosch, N.S., Brandt, S.B., Briland, R.D., Daloğlu, I., DePinto, J.V., Dolan, D.M., Evans, M.A., Farmer, T.M., Goto, D., Han, H., Höök, T.O., Knight, R., Ludsin, S.A., Mason, D., Michalak, A.M., Richards, R.P., Roberts, J.J., Rucinski, D.K., Rutherford, E., Schwab, D.J., Sesterhenn, T.M., Zhang, H., Zhou, Y., 2014. Assessing and addressing the re-eutrophication of Lake Erie: central basin hypoxia. J. Great Lake. Res. 40, 226–246.

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

Scavia, D., Kalcic, M., Muenich, R.L., Aloysius, N., Arnold, J., Boles, C., Confessor, R., De Pinto, J., Gildow, M., Martin, J., 2016b. Informing Lake Erie agriculture nutrient management via scenario evaluation. University of Michigan Water Center (UM). http://graham.umich.edu/media/pubs/InformingLakeErieAgricultureNutrientManagementviaScenarioEvaluation.pdf.

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

Schindler, D.W., Carpenter, S.R., Chapra, S.C., Hecky, R.E., Orihel, D.M., 2016. Reducing phosphorus to curb lake eutrophication is a success. Environ. Sci. Technol. 50, 8923–8929.

Smith, D.R., Harmel, R.D., Williams, M.R., Haney, R., King, K.W., 2016. Managing acute phosphorus loss with fertilizer source and placement: proof of concept. Agricultural & Environmental Letters 1.

Stow, C.A., Cha, Y., Johnson, L.T., Confesor, R., Richards, R.P., 2015. Long-term and seasonal trend decomposition of Maumee River nutrient inputs to western Lake Erie. Environ. Sci. Technol. 49, 3392–3400.

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 Lake. Res. 42, 1174–1183.

Sunohara, M.D., Gottschall, N., Wilkes, G., Craiovan, E., Topp, E., Que, Z., Seidou, O., Frey, S.K., Lapen, D.R., 2015. Long-term observations of nitrogen and phosphorus export in paired-agricultural watersheds under controlled and conventional tile drainage. J. Environ. Qual. 44, 1589–1604.

United States Environmental Protection Agency. (US EPA), 2016. Discharge monitoring report water pollution Search. https://echo.epa.gov/trends/loading-tool/water-pollution-search.

United States Environmental Protection Agency (US EPA), 2018. US action plan for Lake Erie, 2018. https://www.epa.gov/sites/production/files/2018-03/documents/us_dap_final_march_1.pdf.

U.S. Department of Agriculture Agricultural Research Service, Texas A&M AgriLife Research (USDA), 2020. Swat - soil and water assessment Tool. https://data.nal.usda.gov/dataset/swat-soil-and-water-assessment-tool.

U.S. Geological Survey (USGS, 2020. USGS water data for the nation. U.S. Geological survey national water information system database. https://doi.org/10.5066/F7P55KJN. (Accessed 5 October 2020).

Watson, S.B., Miller, C., Arhonditsis, G., Boyer, G.L., Carmichael, W., Charlton, M.N., Confesor, R., Depew, D.C., Höök, T.O., Ludsin, S.A., 2016. The re-eutrophication of Lake Erie: harmful algal blooms and hypoxia. Harmful Algae 56, 44–66.

Williams, M.R., King, K.W., Fausey, N.R., 2015. Drainage water management effects on tile discharge and water quality. Agric. Water Manag. 148, 43–51.

Williams, M.R., King, K.W., Duncan, E.W., Pease, L.A., Penn, C.J., 2018. Fertilizer placement and tillage effects on phosphorus concentration in leachate from fine-textured soils. Soil Tillage Res. 178, 130–138.

Wilson, R.S., Schlea, D.A., Boles, C.M., Redder, T.M., 2018. Using models of farmer behavior to inform eutrophication policy in the Great Lakes. Water Res. 139, 38–46.

Wilson, R.S., Beetstra, M.A., Reutter, J.M., Hesse, G., Fussell, K.M.D., Johnson, L.T., King, K.W., LaBarge, G.A., Martin, J.F., Winslow, C., 2019. Commentary: achieving phosphorus reduction targets for Lake Erie. J. Great Lake. Res. 45, 4–11.

Wituszynski, D.M., Hu, C., Zhang, F., Chaffin, J.D., Lee, J., Ludsin, S.A., Martin, J.F., 2017. Microcystin in Lake Erie fish: risk to human health and relationship to cyanobacterial blooms. J. Great Lake. Res. 43, 1084–1090.

Wolf, D., Georgic, W., Klaiber, H.A., 2017. Reeling in the damages: harmful algal blooms' impact on Lake Erie's recreational fishing industry. J. Environ. Manag. 199, 148–157.

Zedler, J.B., 2017. What's new in adaptive management and restoration of coasts and estuaries? Estuar. Coast 40, 1–21.