Quantifying the Potential Water Quality Benefits of Agricultural Conservation Practices for Stream Fish Conservation in the Western Lake Erie Basin

29 July 2016

Final Report

Western Lake Erie Basin
Conservation Effects Assessment Project – Wildlife

Submitted to
US Department of Agriculture, Natural Resources Conservation Service

Submitted by
The Nature Conservancy (Cooperative Agreement 68-7482-12-504)
and
The Ohio State University (Cooperative Agreement 68-7482-13-502)
and Project Partners, the
USDA-Agricultural Research Service and Ohio Sea Grant

Quantifying the Potential Water Quality Benefits of Agricultural Conservation Practices for Stream Fish Conservation in the Western Lake Erie Basin

S. Conor Keitzer[1], Stuart A. Ludsin[1], Scott P. Sowa[2], Anthony M. Sasson[3], Gust Annis[2], Jeff G. Arnold[4], Amy Brennan[5], Prasad Daggupati[6], August M. Froehlich[3], Matthew E. Herbert[2], Carrie Vollmer-Sanders[7], Michael J. White[4], Christopher J. Winslow[8], and Haw Yen[9]

[1] Aquatic Ecology Laboratory, Department of Evolution, Ecology, and Organismal Biology, The Ohio State University, Columbus, OH 43210, USA

[2] The Nature Conservancy, 101 E Grand River Ave., Lansing, MI 48906, USA

[3] The Nature Conservancy, 6375 Riverside Dr. Suite 100, Dublin, OH 43017, USA

[4] USDA, ARS, Grassland Soil and Water Research Laboratory, Temple, TX 76502, USA

[5] The Nature Conservancy, 20937 Parklane Drive, Fairview Park, OH 44126, USA

[6] Texas A&M AgriLife Research, Texas A&M University, 221B Spatial Science Laboratory, 1500 Research Plaza, College Station, TX 77943, USA

[7] The Nature Conservancy, 330 Intertech Pkwy #110, Angola, IN 46703, USA

[8] Ohio Sea Grant, 1314 Kinnear Rd., Columbus, OH 43212, USA

[9] Blackland Research and Extension Center, Texas A&M University, 720 East Blackland Rd., Temple, TX 76502, USA

Suggested Citation:

Keitzer, S.C., S.A. Ludsin, S.P. Sowa, A.M. Sasson, G. Annis, J.G. Arnold, A. Brennan, P. Daggupati, A.M. Froehlich, M.E. Herbert, C. Vollmer-Sanders, M.J. White, C. J. Winslow, and H. Yen. 2016. Quantifying the Potential Water Quality Benefits of Agricultural Conservation Practices for Stream Fish Conservation in the Western Lake Erie Basin. Final Report submitted to NRCS Conservation Effects Assessment Project. 64 pp.

Cover Photo: Maumee River at Roche de Boeuf, near Waterville, Ohio

TABLE OF CONTENTS

List of Tables……. iv

List of Figures…… v

Executive Summary…….. vi

Acknowledgements……... ix

Introduction……... 1

Project Plan…… 4

Study Area……. 6

Objective 1: Watershed Model Development……... 7

Objective 2: Biological Model Development & Baseline Predictions……. 11

Objective 3: Quantifying Potential Environmental Benefits…… 27

Recommendations for Improving our Approach…….. 53

Conclusions and Implications……... 55

Literature Cited…….. 57

Appendix A: Daggupati et al. 2015. Hydrological Processes 29:5307-5320

Appendix B: Yen et al. 2016. Science of the Total Environment.
http://dx.doi.org/10.1016/j.citotenv.2016.202

Appendix C: Keitzer et al. 2016. Journal of Great Lakes Research.
http://dx.doi.org/10.1016/j.jglr.2016.05.012

LIST OF TABLES

Table 1. SWAT model validation statistics……. 8

Table 2. Simulated water quality conditions for Baseline scenario…….…… 10

Table 3. Model selection statistics for fish metrics…….……. 15

Table 4. Parameter estimates and validation statistics for fish metrics…… 16

Table 5. Non-water quality variables included in species distribution models……... 22

Table 6. Validation statistics for species distribution models…….. 23

Table 7. Agricultural conservation practices included in conservation scenarios……... 28

Table 8. Variables used to develop threat index……..……..... 31

Table 9. Environmental benefits of conservation scenarios relative to Grassland scenario…….. 41

Table 10. Model selection statics for relationships between reductions in agricultural inputs and improvements in stream metrics……... 44

Table 11. Coefficient estimates from models describing relationships between reductions in agricultural inputs and improvements in stream fish metrics…... 45

LIST OF FIGURES

Figure 1. Land-use in the Western Lake Erie Basin …… 3

Figure 2. Conceptual diagram of Western Lake Erie CEAP-Wildlife project……. 5

Figure 3. Simulated water quality conditions in subwatersheds ……. 9

Figure 4. Locations of fish community data used to develop biological models……. 12

Figure 5. Modeled relationships between fish community metrics and water quality…. 17

Figure 6. Forecasted stream biological conditions based on Baseline simulated water quality…… 18

Figure 7. Comparison of stream biological conditions forecasted for the Baseline scenario and Grassland scenario……... 19

Figure 8. Relative influence of different factors on sensitive species distributions…... 24

Figure 9. Forecasted sensitive species richness in subwatersheds based on Baseline water quality conditions…… 25

Figure 10. Comparison of forecasted sensitive species richness for the Baseline scenario and Grassland Scenario……. 26

Figure 11. Estimated improvement in top predator fish populations as a result of a simulated 20% reduction in nutrients and sediments...……... 34

Figure 11. Threat index for subwatersheds……... 35

Figure 12. Potential relationships between reductions in agricultural inputs and improvements in stream fish communities…….. 36

Figure 13. Forecasted biological conditions in conservation scenarios…… 39

Figure 14. Forecasted environmental benefits of conservation scenarios……. 40

Figure 15. Return on investment of conservation scenarios…….. 42

Figure 16. Relationships between stream biological benefits and reductions in agricultural inputs…….. 46

Figure 17. Histograms of average changes in biological conditions that occurred within subwatersheds from reductions in agricultural inputs…… 48

Figure 18. Example of subwatersheds identified as potential priorities for agricultural conservation practices…… 49

Figure 19. Forecasted reductions in nutrient and sediment loading to Lake Erie in conservation scenarios……... 51

Figure 20. Comparison of environmental benefits for Lake Erie and streams …….. 52

EXECUTIVE SUMMARY

Improving the environmental sustainability of agriculture is essential to ensuring the long-term cultural, economic, and ecological well-being of the Great Lakes region. Nowhere in the region is this more apparent than in the watershed of the Western Lake Erie Basin (WLEB). An estimated 5.5 million acres of cropland exist in the WLEB, making it the most intensively farmed watershed in the Great Lakes Basin. Agriculture is fundamental to maintaining the quality of life of many people living in the WLEB and helps feed many people living outside the region. However, non-point source pollution from intensive agriculture threatens aquatic ecosystems and the essential ecosystem services that they provide. Thus, farming practices that allow for continued agricultural production without degrading the surrounding ecosystem are needed in the WLEB.

Strong support currently exists in the WLEB to increase investment in agricultural conservation practices (CPs) to improve water quality in Lake Erie and its tributaries. These practices offer the potential to improve water quality by reducing sediments and nutrients from agricultural fields, while also maintaining soil quality and farm profitability. While large investments in CPs have been made over the past several decades, the cumulative benefits of these investments for aquatic ecosystems are poorly understood at the watershed-scale and how much additional investment might be needed to achieve meaningful environmental benefits is unclear, both at present and with continued climate change. This information is vital for *strategic conservation* in the WLEB. In a large agricultural watershed, such as the WLEB, strategic conservation means getting the right CPs to the right places in the right amount to achieve realistic ecological and water quality outcomes.

The goal of our project was to provide WLEB decision-makers with the scientific information needed to make informed decisions about the use of CPs for stream conservation, primarily, and secondarily for water quality in Lake Erie proper. Specifically, we coupled a state-of-the-art hydrology model for the WLEB watershed with robust predictive biological models to quantify how CPs may improve water quality and benefit stream fish communities, using meaningful measures of stream health. Our project built upon previous work in the Saginaw Bay watershed to better understand how much conservation is enough and what the associated costs would be to achieve meaningful benefits for stream ecosystems in the WLEB. We also sought to identify areas within the watershed where CPs would provide the most benefit and to understand to what extent targeted nutrient reductions for Lake Erie (e.g., 40% reduction in total phosphorus) would benefit stream ecosystems.

Our results highlight the integral role that CPs could have for improving agricultural sustainability in the WLEB. Agricultural runoff appears to be a major contributor to the poor water quality that is widespread throughout the WLEB, potentially limiting fish community health in more than 10,000 km of streams and rivers, representing more than 50% of the watershed. While the current level of CP implementation has certainly improved water quality, a need still exists for additional structural (erosion control) and nutrient management practices. For example, our results showed that, while improvements in stream health could be made by maintaining current CP treatment levels and only further treating farm acres in high-need of CPs (~8% of the watershed), a much larger portion of the watershed (~48%) needs treatment with CPs to achieve widespread benefits for stream fishes. If the effort was made to place additional erosion control and nutrient management CPs in high- and moderate-need acres, fish

communities could be improved in more than 9,600 km of streams and water quality would no longer be limiting in more than 2,500 km of streams. Further improvements would be possible if even more farm acres were treated with CPs.

Our results also showed that multiple water quality concerns (nitrogen, phosphorus, and sediment) need to be tackled simultaneously because high levels of these pollutants often co-occurred within the same stream, each with the potential to limit stream fish communities. Based on established eutrophication thresholds for North American streams, we found that 75% and 91% of streams had phosphorus and nitrogen concentrations, respectively, that could result in eutrophic conditions. Moreover, 49% of the streams had sediment concentrations that could potentially degrade stream fish communities. Nearly 47% of streams had concentrations exceeding thresholds for all three water quality concerns. Therefore, CP implementation that addresses nitrogen, phosphorus, and sediments are needed to reduce water quality limitation and improve stream health throughout the WLEB.

Treating these multiple water quality concerns simultaneously will likely require increased investment in agricultural CPs because different farm fields will require different suites of CPs to address these multiple runoff and leaching concerns, and these CPs must be implemented across a large portion of the watershed to achieve meaningful benefits. For instance, we estimate that treating high- and moderate-needs farm acres with erosion control and nutrient management CPs would cost an additional $149 million annually. Given the apparent levels of needed economic investment, continued interaction among agencies and stakeholders regarding appropriate management and conservation targets in relation to monetary costs seems prudent. Such interaction, supplemented with information from our modeling effort and others like it, should offer the needed science-base to identify the most cost-effective next steps and associated tradeoffs.

We also found that while CPs will have an important role to play in WLEB stream conservation, they likely will not be a panacea. Water quality is expected to limit fish communities in as much as 8,513 km of streams, even if erosion control and nutrient management CPs are implemented across the majority (~80%) of farm acres in the watershed. Thus, expectations for CP benefits should be realistic. Our results showed that farmland treatment with CPs can be an integral *component* of a comprehensive watershed management strategy that considers other potential sources of water pollution (e.g., point sources, urban and exurban runoff) and non-water quality stressors (e.g., dispersal barriers, in-stream habitat, altered hydrology, and invasive species).

Lastly, we found that widespread implementation of CPs appears capable of meeting the Lake Erie total phosphorus loading target identified in the Great Lakes Water Quality Agreement. To meet this target, however, current practices will need to be maintained and high- and moderate-needs farm acres (~48% of the watershed) also will likely need to be treated with CPs. While management focused on Lake Erie should also improve stream conditions, solely focusing on Lake Erie without explicitly considering the health of its tributaries may result in hundreds to thousands of stream kilometers remaining degraded by poor water quality. For instance, treating high- and moderate-needs farm acres with only erosion control CPs would potentially meet Lake Erie nutrient loading goals, but would leave fish communities limited by water quality in more than 2,000 km of streams, as compared to more intensive CP implementation (i.e., erosion control and nutrient management). Considering the health of

streams in Lake Erie's watershed in addition to water quality in Lake Erie proper may help achieve "win-wins" for the whole Lake Erie ecosystem.

While the amount and cost of CP implementation needed to improve stream health in the WLEB may appear daunting, our modeling indicates that win-win-wins for agricultural productivity, local stream ecosystems, and downstream Lake Erie are possible. Achieving these wins in the most cost-effective manner, however, will require strategic conservation to ensure that the right practices are getting to the right places in the right amount, continued research to explore and maximize the potential benefits of CPs, and expanded water quality and biological monitoring to track progress and allow for adaptive management. Unprecedented collaboration across government agencies, conservation organizations, research universities, agribusinesses, and individual farmers also will be necessary to develop innovative, cost-effective solutions. And, because a perfect strategy likely does not exist that can meet all conservation, management, and socioeconomic goals in the WLEB, we must be aware of tradeoffs, be willing to take action with the best available information, and be willing to adapt. In so doing, we will maximize our ability to sustain the vital cultural, economic, and ecological services that the WLEB provides.

ACKNOWLEDGEMENTS

We thank Charlie Rewa for his continued support of this work through the Wildlife Component of the NRCS CEAP and Mari-Vaughn Johnson for all of her thoughtful input throughout the project. We also thank the Indiana Department of Environmental Management, Michigan Department of Environmental Quality, Michigan Department of Natural Resources, and Ohio Environmental Protection Agency for access to their historical databases. We also thank Maura O'Brien for her help with creating data layers used in these analyses. Brian Zimmerman and Marc Kibbey from the Museum of Biological Diversity at The Ohio State University volunteered their time to help secure, create, and check datasets used in these analyses, for which we are incredibly grateful. We also thank Noel Aloysius for assistance with climate change scenarios, Dave Glover for statistical advice, and the Aquatic Ecology Lab at The Ohio State University for reviewing presentations and providing input that greatly improved our approach. We would like to give special thanks Kristen Fussell, Christina Dierkes, and Jill Jentes from Ohio Sea Grant for drastically improving the project's clarity as we began presenting results and for maintaining the project website. We thank Jay Atwood for providing CEAP-Cropland conservation scenarios and Dale Robertson for providing water quality data so important to our modeling efforts. We also would like to thank Jessica D'Ambrosio, Steve Davis, Gail Hesse, Ken Krieger, Chris May, Bill Stanley, and John Stark for providing guidance and taking the time to edit numerous rough drafts of project outputs. We would also like to thank Jared Ross from The Nature Conservancy for his assistance with the integration and quality assurance of the biological datasets for this project.

INTRODUCTION

Agriculture plays an important role in the social and economic wellbeing of many people living in the Western Lake Erie Basin (WLEB). Agriculture is the most prevalent land use in the watershed, covering more than 70% of the land area (Figure 1), or about 5.5 million acres (USDA NRCS 2011). This amount of land in agricultural production makes the WLEB the most intensive agricultural watershed in the Great Lakes Basin, comprising 31% of all cropland acres in the region (USDA NRCS 2011). The economic and social value of this extensive agricultural production is immense. For example, agricultural sales in the WLEB exceeded $2.9 billion[1] during 2012 (www.agcensus.usda.gov). However, the socioeconomic benefits of agricultural production in the WLEB have come with environmental costs. Non-point source (NPS) pollution from agriculture threatens freshwater biodiversity and the vital ecosystem services provided by Lake Erie and its watershed (Scavia et al. 2014). For example, intensive agriculture has degraded stream habitat and water quality, contributing to widespread population declines of more than 40% of the fish species native to WLEB streams and rivers (Karr et al. 1985). Excess nutrients from agricultural runoff are also contributing to Lake Erie's re-eutrophication, which threatens drinking water, fisheries, tourism, and other valuable ecosystem services supported by a healthy Lake Erie (Michalak et al. 2013, Scavia et al. 2014). Thus, the adoption of environmentally sustainable agricultural practices is essential for the long-term cultural, economic, and ecological health of the WLEB.

Considerable interest exists among WLEB decision-makers in understanding the potential for agricultural conservation practices (CPs) to improve agricultural sustainability. Such practices can improve water quality by reducing NPS pollution from agriculture, while also improving soil health and maintaining farm profitability (Schnepf and Cox 2006). Indeed, growing evidence indicates that past implementation of CPs has improved stream water quality (Richards and Baker 2002, Richards et al. 2005, Richards et al. 2009) and biological conditions (Ohio EPA 2014, Miltner 2015) in the WLEB.

While CPs hold great promise for improving the environmental sustainability of agriculture, several key information gaps limit our ability to effectively incorporate CPs into watershed conservation plans for the WLEB. Critical outstanding questions include - (1) "How much additional CP implementation is needed to improve stream water quality and fish community health, both now and under a changing climate?"; (2) "Which types of CPs are most beneficial and cost-effective?"; (3) "How much financial investment is needed to see meaningful benefits in stream health?"; (4) "Where in the watershed will the implementation of additional CPs be most beneficial?"; and (5) "Do potential "win-wins" exist for Lake Erie water quality management and stream conservation?" The goal of our project was to quantify the potential environmental benefits of CPs and provide WLEB decision-makers with the information needed to make informed decisions about the use of CPs for protecting and rehabilitating stream health, primarily, and secondarily for improving Lake Erie water quality. This information is essential for *strategic conservation*, which involves getting the right types of CPs to the right places in the right amount to achieve realistic conservation outcomes (Sowa et al. 2011; Sowa et al. *In press*).

---

[1] This estimate is likely conservative because it only includes data from counties with at least 90% of their land area within the WLEB. For example, if counties with ~50% of their land area within the WLEB were included, this estimate increased to more than $4 billion.

We coupled a state-of-the-art hydrologic model (Soil Water Assessment Tool, SWAT, Arnold et al. 1999) with robust predictive biological models to quantify how additional implementation of CPs may improve water quality and benefit stream fish communities in WLEB rivers and streams. We focused on fish communities because they provide an integrative measure of stream conditions and are commonly used by federal and state agencies to assess overall stream health. We specifically explored how effective CPs might be for reducing total phosphorus (TP), total nitrogen (TN), and suspended sediment (SS) concentrations, stressors that can affect stream fishes in a variety of ways. Excess nutrients can stimulate algal growth, resulting in altered dissolved oxygen levels and changes in stream food webs that can negatively affect stream fish communities (Miltner and Rankin 1998, Wang et al. 2007, Evans-White et al. 2009, Miltner 2010, Taylor et al. 2014). Chronic levels of high nitrogen concentrations also may be toxic to some fish species (Camargo et al. 2005, Camargo and Alonso 2006). Excess SS concentrations may directly harm fish by damaging gills, altering light levels and primary production, and reducing the foraging ability of visual predators (Waters 1995, Wood and Armitage 1997). For these reasons, high levels of TN, TP, and SS are considered a major threat to stream fish communities and are a main contributor to stream impairment in the United States (US EPA 2016). While other stressors also are important (e.g., invasive species, in-stream habitat degradation, altered hydrology), by focusing explicitly on these three water quality stressors, we were able to delineate how CPs may benefit stream fishes by reducing these persistent and ubiquitous stressors.

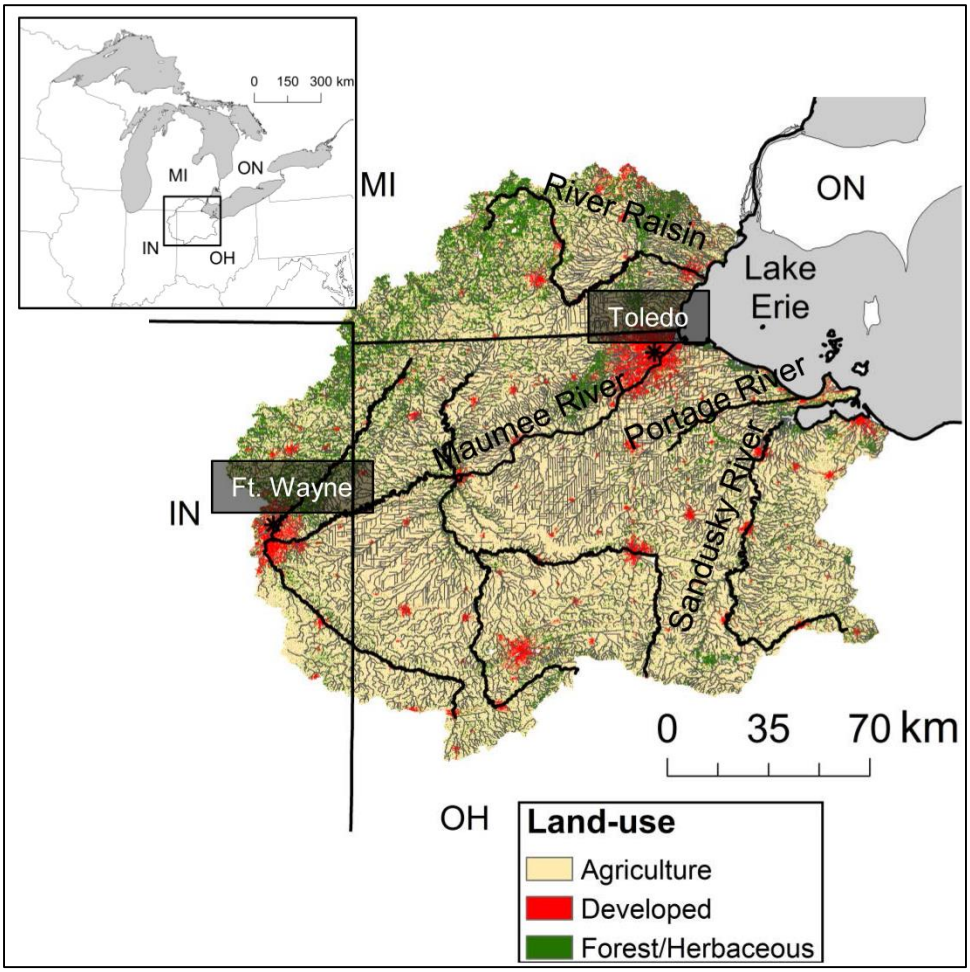


Figure 1. Land-use in the Western Lake Erie Basin (WLEB) and its extensive stream network, which contains more than 20,000 km of streams. Major land uses include agriculture (>70% of the watershed), developed land (~12% of the watershed), and forested or herbaceous lands (~12% of the watershed). Major tributaries in the WLEB include the Maumee River, Portage River, River Raisin, and Sandusky River.

PROJECT PLAN

We developed four specific objectives to complete this project (Figure 2). To simulate the effects of land-use, including farming practices, on stream water quality we developed a fine-scale watershed model using SWAT (**Objective 1**). We then linked simulated water quality conditions to observed fish community data and developed stressor-response biological models (**Objective 2**). We used these models to forecast fish community health throughout the WLEB based on simulated water quality conditions during 1990-2010. To quantify the potential benefits of additional implementation of CPs, we used conservation scenarios based on the CEAP-Cropland conservation practice adoption scenarios (USDA NRCS 2011) that consisted of different types and amounts of CP implementation (**Objective 3**). Farm-scale reductions in agricultural runoff and leaching losses in each conservation scenario were simulated as part of CEAP-Cropland using the Agricultural Policy/Environmental eXtender model (APEX; Gassman et al. 2009). These reductions were used to adjust edge-of-field nutrient and sediment inputs in our SWAT model and then routed through the WLEB stream network to simulate changes in water quality at the stream-scale. We then used our biological stressor-response models to forecast and quantify potential impacts on fish community health that resulted from simulated CP implementation. The results of this modeling effort are being shared with managers, stakeholders, and decision-makers to inform strategic conservation in the WLEB (**Objective 4**).

Detailed descriptions of how each objective was accomplished are provided in a series of progress reports and peer-reviewed manuscripts (Appendices A-C), which are also available on the project website (http://lakeerieceap.com/) or by request (Dr. Stuart Ludsin, The Ohio State University, email: ludsin.1@osu.edu). Here, we provide a general overview of the project's approach, synthesize overall findings, and discuss the broader implications of our modeling effort.

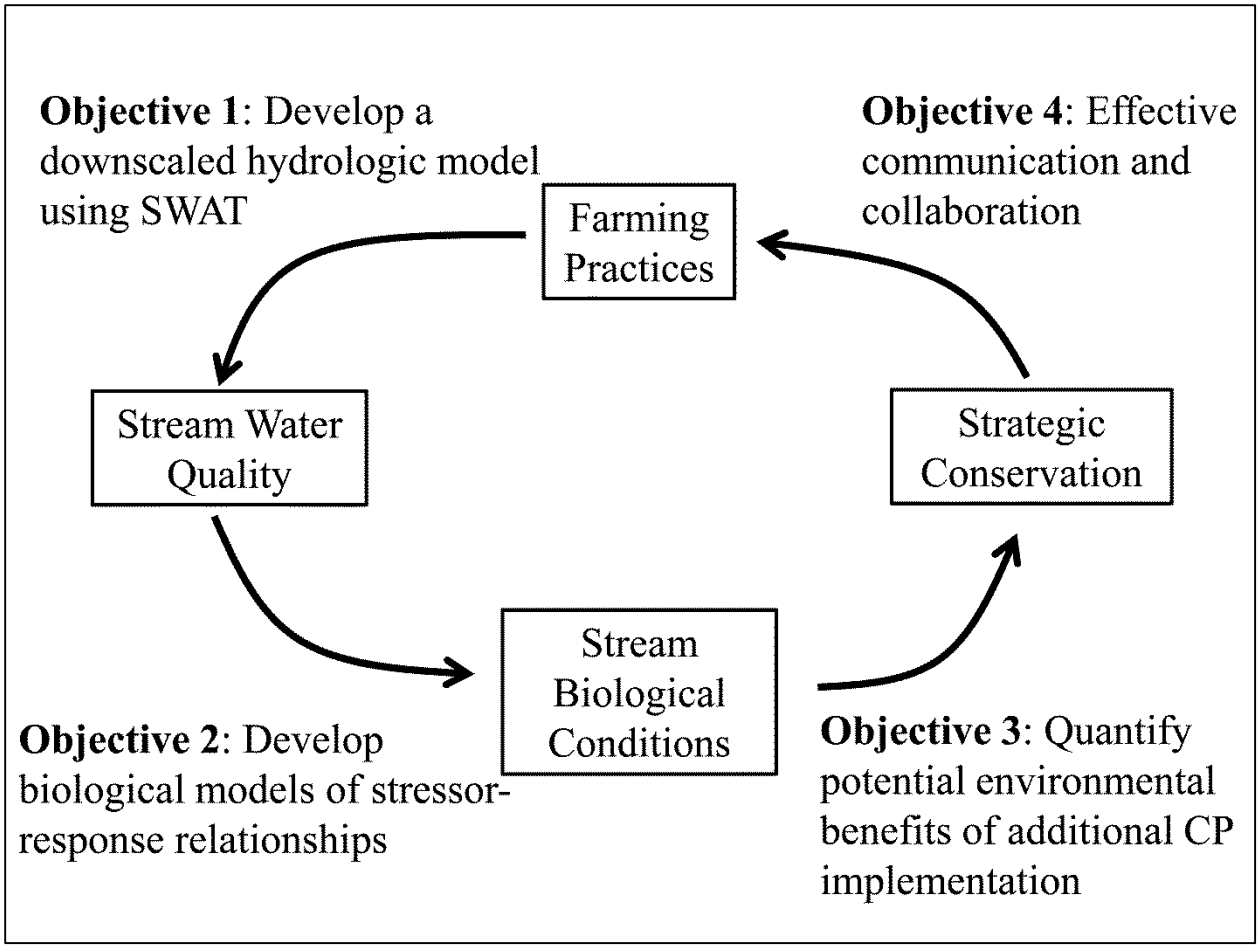


Figure 2. Conceptual diagram for the Western Lake Erie CEAP-Wildlife project. Detailed descriptions of how we accomplished each objective are available in a series of progress reports on the project website (http://lakeerieceap.com/).

STUDY AREA

The WLEB is a ~ 26,000 km^2 watershed that drains portions of southeastern Michigan, northeastern Indiana, and northwestern Ohio (Figure 1). This large watershed is predominantly in the Eastern Corn Belt Plains (~50%) or Huron/Erie Lake Plains (~49%) Level III Ecoregions, with the remaining ~1% being in the Southern Michigan/Northern Indiana Drift Plains. These ecoregions are characterized by fertile soils that historically supported a mixture of hardwood forests, wetlands, and prairies (Woods et al. 1998). Widespread land-clearing and wetland-draining began during the mid-1800s, with the fertile soils now supporting highly productive row-crop agriculture (Trautman, 1981). Although agriculture is the most prevalent land use in the WLEB (>70% of the watershed), patchily distributed urban and forested/herbaceous areas occur throughout the watershed.

Streams in the WLEB are typically low gradient (average slope < 2%) and slow-flowing. Historically, they supported a rich fish fauna, with 98 fish species native to the WLEB (Karr et al. 1985). This level of diversity makes the WLEB the most biologically diverse watersheds in the Great Lakes Basin (Trautman, 1981). Unfortunately, a long history of habitat and water quality degradation has taken its toll on WLEB stream communities. More than 40% of the fish species native to the area have experienced population declines, with as many as 17 species potentially extirpated locally (Karr et al. 1985). These effects are most pronounced for headwater species sensitive to habitat degradation, along with specialized feeding guilds such as herbivores, predators, and insectivores (Karr et al. 1985). By contrast, pollution-tolerant and omnivorous species have increased in numbers, resulting in contemporary fish communities that would be largely unrecognizable to early European settlers in the WLEB (Trautman, 1981; Karr et al. 1985).

OBJECTIVE 1: DEVELOPMENT AND VALIDATION OF THE DOWNSCALED WATERSHED MODEL

We developed a fine-scale watershed hydrological model using SWAT to simulate stream water quality in the WLEB watershed. SWAT is a widely used process-based model (Gassman et al. 2007, Arnold et al. 2015) that performs well in comparison to other watershed models in the WLEB (Gebremariam et al. 2014). Because detailed descriptions of the model and its development are available elsewhere (Model initialization: Daggupati et al. 2015a, Flow/water quality calibration/validation: Yen et al. 2016, Appendix A, Appendix B, http://lakeerieceap.com/), we only provide a brief description here. The SWAT model was calibrated from 1990 to 1999 with a three-year warm-up period (1987-1989; Yen et al. 2016). The model was validated from 2000 to 2006 using observed TN, TP, and SS loads from five gauges throughout the WLEB. In total, watershed simulations were run from 1990-2010 to simulate stream water quality.

A novel aspect of our SWAT model is that we developed it at the NHDPlus-scale, a much finer spatial resolution than previous SWAT models. For example, typical SWAT models are developed for 12-digit hydrologic unit code (HUC-12) subwatersheds, of which there are 391 in the WLEB with an average size of 72 km^2. By contrast, there are 11,128 subwatersheds with an average size of 2.6 km^2 in the WLEB at the NHDPlus-scale (Daggupati et al. 2015a). We developed the model at this fine spatial resolution because previous research in the Saginaw Bay watershed demonstrated the value of a finer resolution SWAT model for developing accurate biological models and informing stream conservation in agricultural watersheds (Sowa et al. 2011).

To calibrate the SWAT model at this fine spatial resolution, we initially calibrated model parameters at the 1:240,000 scale using predefined HUC-12 subwatersheds and then further calibrated and validated at the 1:100,000 scale using the NHDPlus stream network (http://www.horizon-systems.com). This approach of transferring parameters across spatial scales worked well, particularly at the monthly time-step (Daggupati et al. 2015a). We also used a proxy-basin spatial calibration strategy (Klemes 1986, Daggupati et al. 2015b) to reflect spatial heterogeneity that exists in large watersheds like the WLEB (Daggupati et al. 2015a). This approach greatly improved model performance relative to a non-spatial strategy (Daggupati et al. 2015a). We incorporated tile drainage into our SWAT model by assuming that any agricultural fields located in poorly drained soils with < 1% slopes contained tile drains (Daggupati et al. 2015a). We found that this approach for including tile drainage greatly improved model performance (Daggupati et al. 2015a). We also used “soft data” during model development to ensure simulations properly reflected realistic watershed behavior (Yen et al. 2014a, Yen et al. 2014b). Specifically, the denitrification rate was controlled to be less than 50 kg/ha (David et al. 2009) and the ratio of nitrate contributions from tile drainage versus surface runoff was above two-thirds of total nitrate losses (Schilling 2002). More details about the SWAT functions used for soft data and nutrient processes in the tile drain system can found elsewhere (Yen et al. 2016).

Model development focused on accurately estimating long-term average conditions at the monthly time-step, using the percent bias (PBIAS) as our measure of accuracy. Negative PBIAS values indicate that the model predictions were below observed values on average and positive values indicate that model predictions were above observed values. Values close to zero indicate

a close match of model predictions to observed values. Based on widely adopted levels of acceptable model performance (Moriasi et al., 2007), we found that our watershed model performed well for all water quality attributes except for River Raisin stream flow, which was marginally (+1%) outside the range of what is considered acceptable (Table 1).

Table 1. Percent bias (PBIAS) for watershed model validation at the monthly time-step for stream discharge and suspended sediment (SS), total phosphorus (TP), and total nitrogen loads (TN) within the WLEB during 2000 to 2006. Boldface font indicates acceptable model performance based on PBIAS (Moriasi et al. 2007).

	Stream discharge	**SS**	**TP**	**TN**
Station	**PBIAS**	**PBIAS**	**PBIAS**	**PBIAS**
Maumee River	**-14%**	**10%**	**-3%**	**-13%**
River Raisin	26%	**-35%**	**23%**	**-4%**
Sandusky River	**-7%**	**35%**	**-1%**	**12%**
St. Joseph River	**-19%**	**20%**	**-5%**	**69%**
St. Marys River	**-25%**	**-19%**	**-9%**	**-21%**

Baseline Water Quality Conditions in the WLEB

Nutrient and sediment concentrations during Baseline conditions (i.e., simulated conditions during 1990-2010 before CPs were added in our conservation scenarios) appeared to be high in streams throughout the WLEB (Figure 3). Mean (± 1 S.D.) annual concentrations (mg/L) in WLEB streams were 5.27 (± 3.22) for TN, 0.191 (± 0.148) for TP, and 142.31 (± 173.26) for SS. To place this into context, TN concentrations above 1.5 mg/L and TP concentrations above 0.075 mg/L may signify eutrophic conditions in North American streams (Dodds et al. 1998). Although less well defined, SS concentrations greater than 80 mg/L may negatively affect freshwater fisheries (http://www.in.gov/idem/nps/3484.htm). Based on average annual concentrations, we found that as much as 75% (15,342 km) and 91% (18,533 km) of streams in the WLEB were above TP or TN eutrophication thresholds, respectively, with SS concentrations in 49% (9,989 km) of streams being above 80 mg/L. Nearly 47% (9,454 km) of streams had average values above thresholds for all three water quality attributes, indicating the need to reduce inputs of all three potential stressors to improve stream health. While these exact percentages should be interpreted with caution, surveys by state agencies also consistently document high nutrient levels that are potentially harmful to aquatic life, biological indicators of eutrophication, drinking water impairment as the result of high nitrate levels in several municipalities, and potentially harmful levels of sedimentation/siltation in many streams in WLEB streams (Ohio EPA 2014). Moreover, nutrient and sediment stressors often co-occurred in biologically impaired streams from surveys during 2012-2013 (e.g., Table 3 in Ohio EPA 2014). Taken together, our results and recent observations from within the watershed suggest that water quality stressors are widespread in WLEB streams, often co-occurring. More generally, our results support the growing recognition of the need to address multiple stressors in freshwater ecosystems (Ormerod et al. 2010), which can interact to affect stream biota (Townsend et al. 2008, Matthaei et al. 2010, Piggott et al. 2015).

Agriculture appears to be a major contributor to these water quality conditions. When compared to simulations from a Grassland scenario, in which we converted all agricultural lands to native grasses and simulated resulting conditions, water quality stressors were considerably higher in the Baseline scenario. We found that average concentrations of TP, TN, and SS were, respectively, 0.15 mg/L, 4.24 mg/L, 117.25 mg/L higher for Baseline conditions than in the Grassland scenario (Table 2). Although not a precise measure of agricultural contributions, these large differences between the Baseline conditions and the Grassland scenario suggest that agriculture is contributing a substantial amount of nutrients and sediments to WLEB streams.

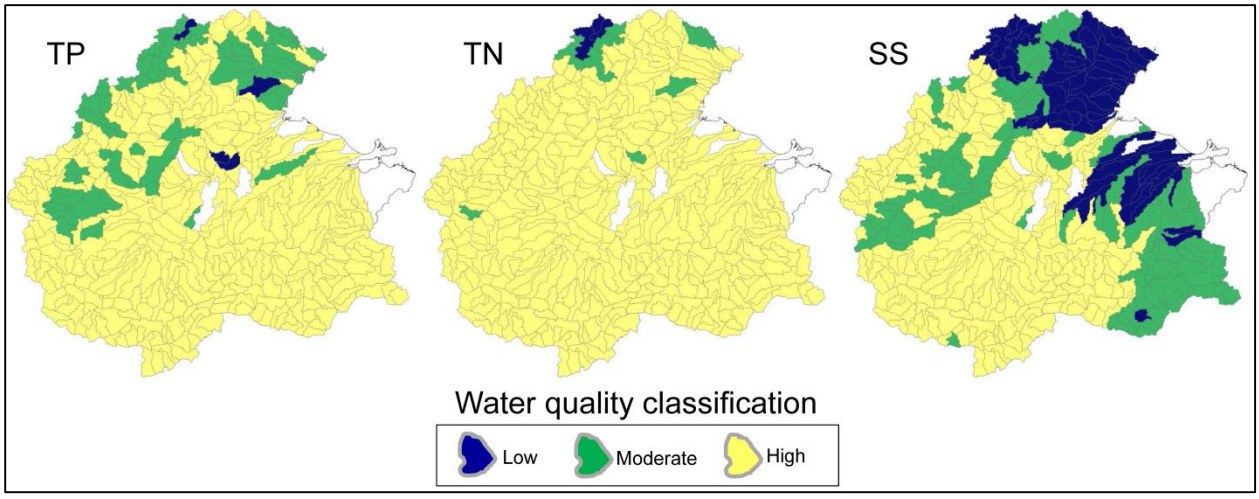


Figure 3. Simulated Baseline water quality conditions in WLEB subwatersheds based on average annual total phosphorus (TP), total nitrogen (TN), and suspended sediment (SS) concentrations simulated during 1990-2010. Concentrations below 0.025 mg/L for TP, 0.7 mg/L for TN, and 25 mg/L for SS were considered low while those above 0.075 mg/L for TP, 1.6 mg/L for TN, and 80 mg/L for SS were considered high.

Table 2. Minimum (Min), maximum (Max), and average (Mean) simulated water quality conditions for the Baseline (Baseline) and Grassland (Grassland) scenarios. Averages represent the mean concentration or flow across all stream segments. Q = Discharge; SS = suspended sediments; TP = total phosphorus; TN = total nitrogen; NI = nutrient index.

Annual Water Quality										
	Q (L/s)		**SS (mg/L)**		**TP (mg/L)**		**TN (mg/L)**		**NI**	
	Baseline	**Grassland**	**Baseline**	**Grassland**	**Baseline**	**Grassland**	**Baseline**	**Grassland**	**Baseline**	**Grassland**
Min	0	0	0	0	0.00	0.00	0	0	0.00	0.00
Max	117302	119696	1243	908	0.93	0.51	18.8	9.1	1.55	1.14
Mean	2303	2340	142	25	0.19	0.04	5.3	1.0	0.80	0.29

OBJECTIVE 2: DEVELOPMENT OF BIOLOGICAL STRESSOR-RESPONSE MODELS AND BASELINE PREDICTIONS

Compiled Biological Data

Stream fish community data were provided by Indiana Department of Environmental Management (18 stream reaches), Michigan Department of Environmental Quality and Michigan Department of Natural Resources (101 stream reaches), and the Ohio Environmental Protection Agency (722 stream reaches). Samples were collected during 1979-2012, which we reduced to only samples collected after 1990 to more closely match our hydrological simulation period. Fish samples were collected through various electro-shocking methods according to the size of stream being sampled, but were similar among agencies (Ohio EPA 1987; Michigan DEQ 1997; Indiana DEM 2007). These samples were rigorously evaluated to ensure that they represented community samples, species occurrences were correct given their known ranges, and sample locations were geographically accurate. Samples were spatially linked to the NHDPlus stream network according to their latitude, longitude, and written descriptions of their location. If multiple samples were collected in the same stream segment during the same year, we used the mean value to represent that stream segment and year combination. If stream segments were sampled during multiple years, we used the most recently available sampling year to reflect current biological conditions as closely as possible. This data-filtering process resulted in 841 unique fish samples from across the watershed that were used to develop biological models (Figure 4). The average number of samples within major HUC-8 drainages was 76, with the St. Mary's River watershed (04100004) sampled the least (N = 14) and the Sandusky River watershed (04100011) sampled the most (N = 200). The average upstream drainage area of sampled stream segments was 799 km^2, but ranged from 0.89 km^2 to 17,018 km^2. Most samples (N = 605) were from mid-sized to small streams (Strahler stream order 1-3), with the rest (N = 236) from larger rivers (Strahler stream order 4-12)

Fish community samples were used to calculate an index of biotic integrity (IBI) according to Michigan's procedure-51 (Michigan DEQ 1997). The IBI is a multimetric index of various feeding, reproductive, and pollution tolerance guilds that measures the condition of a fish community relative to what would be expected if human disturbances were not present (Karr 1981). We also calculated the relative abundance of piscivorous fishes to assess biological conditions. The presence of piscivores represents a healthy and trophically diverse fish community, with these species typically declining in abundance with increasing human disturbance (Karr 1981). Moreover, piscivores are often recreationally important (e.g., smallmouth bass, *Micropterus dolomieu*; northern pike, *Esox lucius*). Hereafter, we refer to the relative abundance of piscivorous fishes as the Piscivore Index. These fish community metrics were re-scaled as $z_i = \frac{x_i - \min(x)}{\max(x) - \min(x)}$ to have a minimum of zero and a maximum of one based on the minimum and maximum observed values from our fish community dataset.

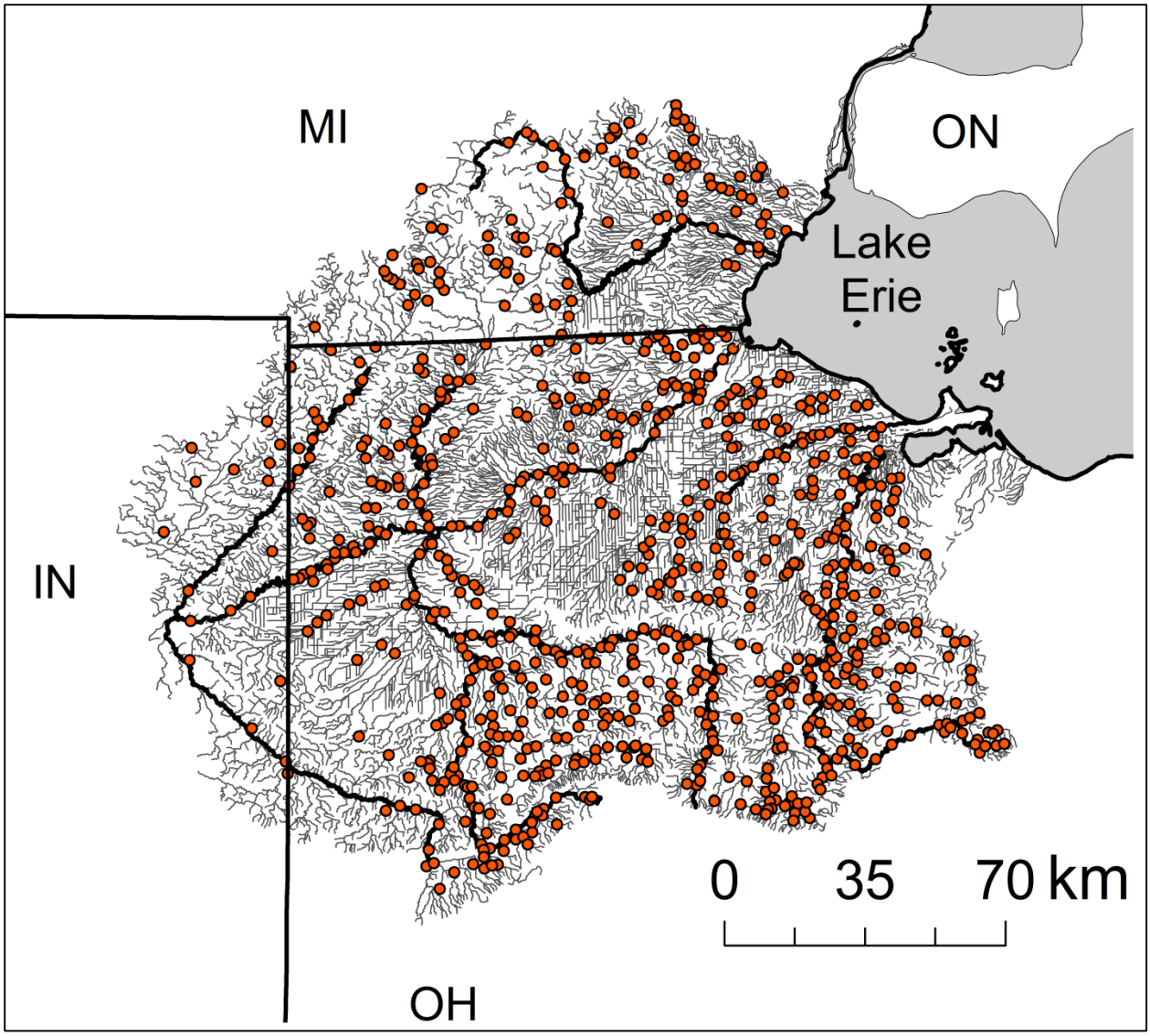


Figure 4. Locations of fish community samples used to develop biological models. Fish community samples (n=841) were collected by Indiana Department of Environmental Management, Michigan Department of Environmental Quality, Michigan Department of Natural Resources, and Ohio Environmental Protection Agency during 1990-2010 using a variety of electro-shocking methods.

Development of Community Level Stressor-Response Biological Models

We modeled limiting relationships, or ceilings, between fish community metrics and various water quality and stream flow parameters using quantile regression. We used quantile regression because it is appropriate for identifying limiting relationships between stressors and responses despite not all potentially limiting factors being measured (Cade et al. 1999, Cade and Noon 2003). A detailed description of our approach is provided in project progress reports (http://lakeerieceap.com/), as well as published manuscripts (Keitzer et al. 2016; Appendix C).

In short, we linked stream water quality and flow conditions simulated from our watershed model during 1990 to 2010 to observed fish community data. We then used an information theoretic approach to identify the most parsimonious model for explaining variation in the observed data from our candidate set of models (Burnham and Anderson 1998). This initial candidate set included models describing various additive and interactive effects of water quality and flow on stream fish communities (Table 3). We included models containing average annual, spring (1 March through 30 June), or summer (1 July through 30 September) values as covariates in candidate models to determine if seasonal or annual variables better explained fish community health. These water quality and flow variables represent long-term averages for a stream segment from the 21 years simulated by our hydrologic model.

Two important updates were made to our approach from earlier reports. The first is that we modeled the 97th percentile, as opposed to the 95th percentile, of the fish community data to model ceilings in relationships more closely. Second, we did not include TP and TN within the same models because of concerns over multicollinearity (Dormann et al. 2013). Instead, we created a nutrient index variable that was the additive combination of $\log_{10}$+1-transformed TN and $\log_{10}$+1-transformed TP concentrations. This allowed us to examine the combined effects of TN and TP without the issue of multicollinearity. A similar nutrient index, using dissolved inorganic nitrogen instead of TN, has proven useful for understanding nutrient effects on stream fauna (Niyogi et al. 2007, Townsend et al. 2008). The best[2] models were then validated using *k*-fold cross-validation. Cross-validation consisted of randomly dividing the data into 10 approximately equal-sized groups (k=10) and iteratively selecting one group for model parameterization, and validating model predictions using observed data from the remaining "test" data. We considered models acceptable if we found statistically significant and positive correlations for each *k*-fold and at least 97% of the observed values fell below predicted values (Vaz et al. 2008).

Fish community metrics were logit-transformed (Warton and Hui 2011) prior to analyses because values were bounded between zero and one. Stream flow values were $\log_{10}$+1-transformed to improve linearity. Water quality attributes (flow, nutrient index, and SS) were standardized to have a mean of zero and standard deviation of one to place all predictors on the same scale and improve interpretation of their relative effects. Quantile regressions were fit using the 'quantreg' package (Koeneker 2015) and model selection statistics were calculated using the 'MuMIn' package (Barton 2015) in the R statistical environment (R Core Team 2015).

[2] The best or most parsimonious model was identified by the bias corrected Akaike's Information Criteria (AIC_c; Burnham and Anderson 1998).

Our models accurately captured limiting relationships between fish community metrics and stream water quality and flow (Table 4). The inclusion of the nutrient index, SS, and an interaction between them in the best model for the IBI and Piscivore Index showed that both nutrients and sediments are important in explaining fish community health in the WLEB. The modeled relationships suggest that water quality stressors had a larger impact on the Piscivore Index than the IBI (Figure 5). While relationships were generally negative for water quality stressors and positive for stream flow, as expected, there did appear to be a positive relationship between SS and the IBI. The reason for this unexpected positive relationship is unclear, but the effect was relatively minor and may suggest that sediments have a small effect on the IBI when suspended in the water column. Instead, the well-documented negative effects of sediment pollution may become more pronounced when sediments are deposited in streams, where they can smother foraging and reproductive habitat and alter food webs (Waters 1995, Wood and Armitage 1997).

We used the best supported model for both fish community metrics to forecast the potential fish community health within a stream segment. This process was done for all streams in the WLEB and served as the 'Baseline' condition for stream fish communities based on average water quality conditions during 1990-2010. We considered a stream segment to be "Limited" if the forecasted fish community metric from our quantile regression models was below the 90th percentile of the observed fish community dataset (Figure 5). We also simulated a scenario in which all farmland was converted to native grasses. This "Grassland" scenario served as an unrealistic, but important "bookend scenario" by which to assess the effectiveness of CPs in our other scenarios.

Our modeling suggests that water quality, including nutrients, sediments, and stream flow, were limiting stream fish communities throughout the WLEB (Figure 6). Our forecasted biological conditions indicated 61% (12,414 km) and 54% (10,967 km) of the watershed was limited for the IBI and Piscivore Index, respectively. The amount of the watershed where water quality was not limiting stream fish communities was considerably lower in the Baseline scenario than forecasted for the Grassland scenario (Figure 7), suggesting that poor water quality from agricultural runoff is a major contributor to degraded stream fish communities in the WLEB.

Table 3. Candidate models used to forecast fish metrics and model selection statistics. Models were ranked according to the bias-corrected Akaike's Information Criterion (AIC_c). Spring = Sp; Annual = An; Summer = Su; Discharge = Q, suspended sediments = SS, nutrient index = NI. Boldface indicates the model used to forecast fish community metrics. IBI = Index of Biotic Integrity; Piscivore Index = Relative abundance of piscivorous fishes.

IBI		
Model	**AIC_c**	**Δ AIC_c**
SpQ + SpSS + SpNI + SpSS x SpNI	**2988.27**	**0.00**
SpQ + SpSS + SpNI	2990.17	1.90
AnQ + AnSS + AnNI	3035.38	47.11
AnQ + AnSS + AnNI + AnSS x AnNI	3037.31	49.05
SpQ + SpNI	3047.20	58.93
SuQ + SuSS + SuNI + SuSS x SuNI	3050.54	62.27
SuQ + SuSS + SuNI	3052.72	64.46
SuQ + SuNI	3069.91	81.64
AnQ + AnNI	3078.70	90.44
SuQ	3092.88	104.61
SuQ + SuSS	3093.07	104.80
SpQ	3099.11	110.85
SpQ + SpSS	3101.05	112.78
AnQ	3115.77	127.50
AnQ + AnSS	3117.50	129.23
Null	3271.29	283.02
Piscivore Index		
SpQ + SpSS + SpNI + SpSS x SpNI	**3681.49**	**0.00**
SpQ + SpSS + SpNI	3696.22	14.73
SpQ + SpNI	3697.17	15.68
AnQ + AnSS + AnNI + AnSS x AnNI	3697.83	16.33
AnQ + AnSS + AnNI	3698.95	17.46
AnQ + AnNI	3700.01	18.51
SuQ + SuSS + SuNI + SuSS x SuNI	3736.21	54.72
SuQ + SuSS + SuNI	3749.85	68.35
SuQ + SuNI	3749.91	68.42
SuQ + SuSS	3761.98	80.48
AnQ + AnSS	3764.98	83.49
SpQ + SpSS	3767.88	86.39
AnQ	3830.66	149.17
SuQ	3833.21	151.71
SpQ	3833.40	151.90
Null	3866.52	185.02

Table 4. Parameter estimates and model validation statistics for the best models for fish community metrics in the WLEB. Parameter estimates (± 1 S.D.) for the intercept, stream discharge (Q), suspended sediments (SS), nutrient index (Nut. Ind.), and interactions are shown on the logit-scale. Model validation statistics included the mean (± 1 S.D.) Spearman's rank correlation (r_s) between observed and predicted data and the percentage of observed data that were less than the predicted data (% below). Boldface font indicates acceptable validation statistics. IBI = Index of Biotic Integrity; Piscivore Index = Relative abundance of piscivorous fishes.

Model	**Intercept**	**Q**	**Nut. Ind.**	**SS**	**Nut. Ind. x SS**	r_s	**% below**
IBI	1.65	0.64	-0.37	0.20	0.06	**0.53 (0.08)**	**97 (2)**
Piscivore Index	-0.62	0.81	-0.73	-0.41	0.51	**0.44 (0.09)**	**97 (2)**

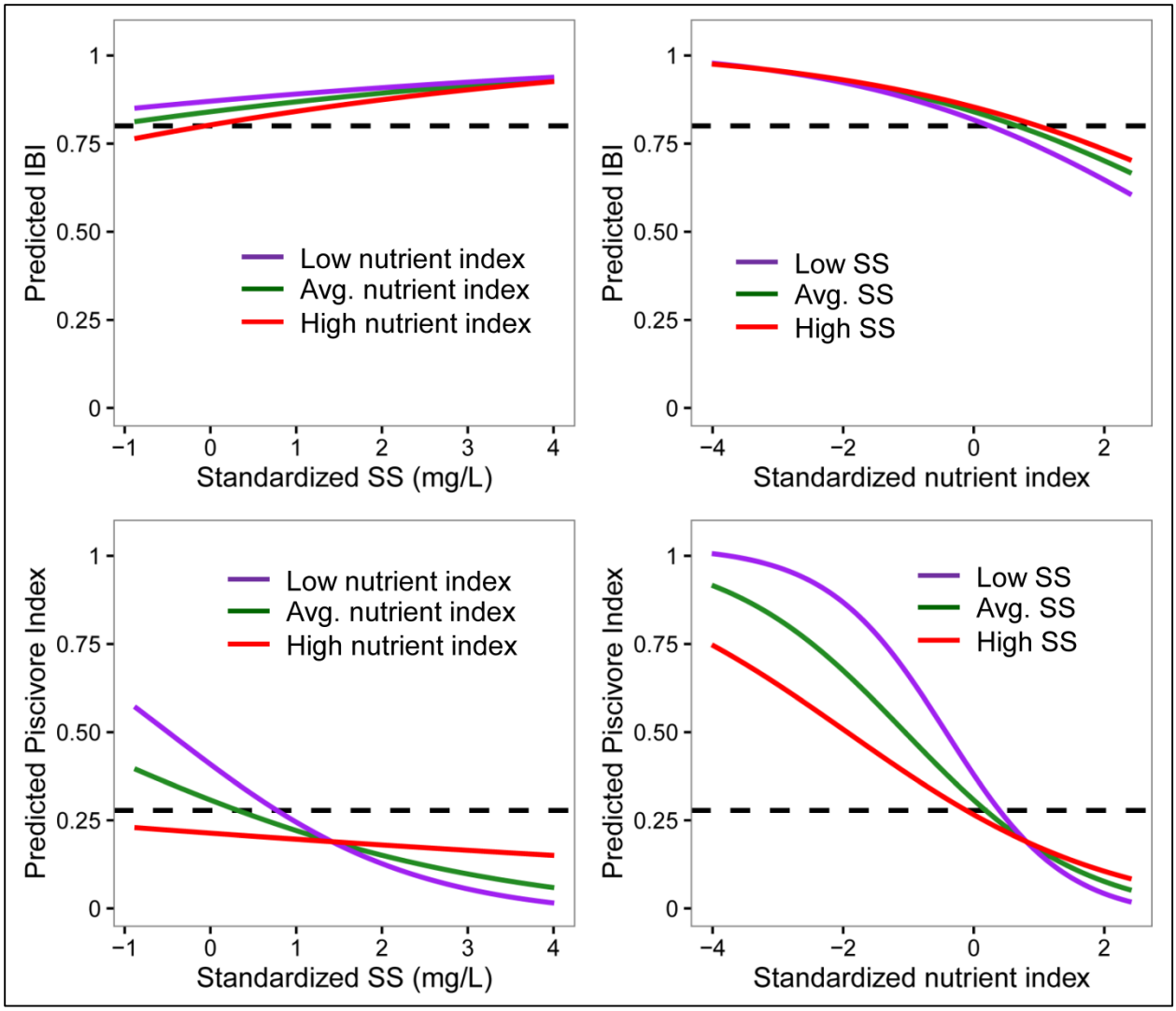


Figure 5. Modeled relationships between fish community metrics and suspended sediments (SS) or the nutrient index. Water quality parameters were the average of simulated spring (1 March to 30 June) concentrations during 1990 – 2010, and standardized to have a mean of zero and S.D. of one. The dashed black line indicates the threshold where fish metrics were no longer considered limited; this value was based on the 90th percentile of the observed fish sampling data. Solid lines indicate relationships between fish metrics and water quality stressors relative to low (purple), average (green), or high (red) levels of the other stressor. All relationships are relative to a stream with a standardized spring flow of -0.18, which represents the median flow of streams that had observed fish community data. IBI = Index of Biotic Integrity; Piscivore Index = Relative abundance of piscivorous fishes.

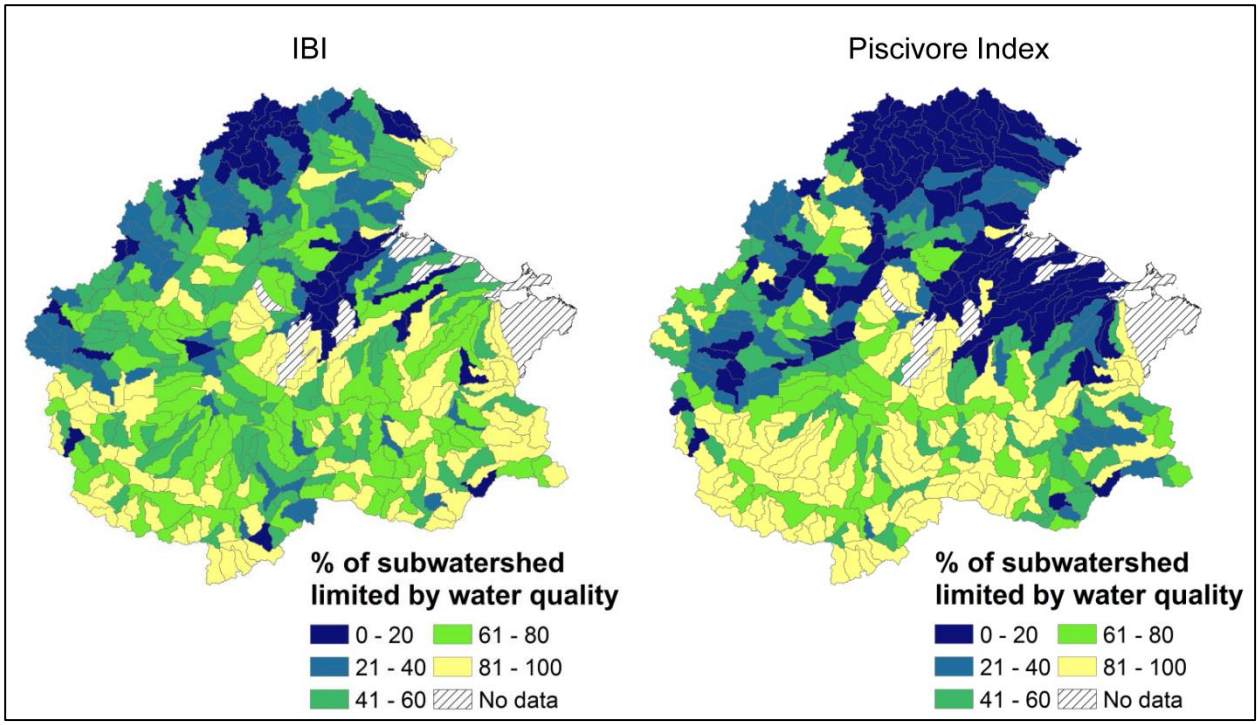


Figure 6. Forecasted stream biological conditions in the WLEB based on water quality simulated in the Baseline scenario. The percentages of streams within a subwatershed classified as "Limited" are shown. Streams were classified as Limited if the biological conditions forecasted from the quantile regression models were below the 90th percentile of the observed fish community data. IBI = Index of Biotic Integrity, Piscivore Index = Relative abundance of piscivorous fishes.

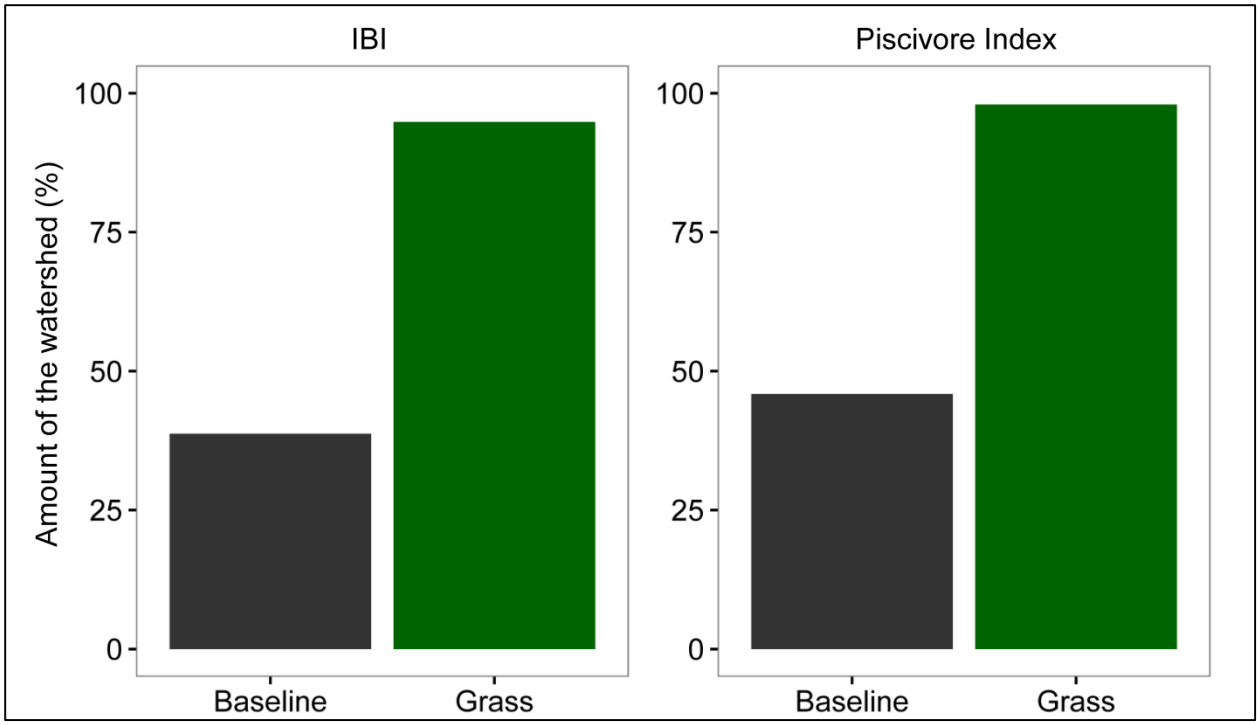


Figure 7. Comparison of the percentage of the WLEB that is not limited by nutrients, sediments, or stream flow in the Baseline and Grassland (Grass) scenarios. IBI = Index of Biotic Integrity; Piscivore Index = Relative abundance of piscivorous fishes.

Development of Species Distribution Models

In addition to modeling fish community metrics, we developed species distribution models (SDMs) to predict the presence/absence of species within a stream segment for 19 fish species that are sensitive to degraded habitat conditions. This allowed us to explore how individual species respond to water quality benefits of CP implementation. These species-level changes provided a deeper understanding of CP effectiveness and are of management concern in their own right. Species sensitivities were assigned from published literature (Angermeier and Karr 1986, Lyons et al. 1996) and Ohio Environmental Protection Agency (Ohio EPA) guidelines (Ohio EPA 2013a).

We used boosted regression trees (BRT) to develop SDMs because it is a flexible regression approach that often outperforms traditional SDM modeling approaches, such as generalized linear models (Elith et al. 2008). Another useful property of BRTs is that the relative influence of variables can be compared, with higher relative influences indicating variables more important in explaining species distributions compared to other variables included in SDMs (Elith et al. 2008). We initially fit BRT models with a tree complexity of four and a learning rate of 0.001, which was then adjusted to ensure that at least 1000 regression trees were included (Elith et al. 2008). Our BRT models were fit using the 'dismo' package (Hijmans et al. 2015) in the R statistical environment (R Core Team 2015).

We included several natural and human threat variables in addition to water quality variables in SDMs (Table 5). These additional variables were included to improve the predictive performance of SDMs and evaluate the relative importance of water quality and non-water quality variables in shaping species distributions. We only included species that were present at $\geq$ 5% of sites ($n \geq 42$).

These models produce a probability of species occurrence within a stream segment. Thus, a probability threshold must be used to assign a species as present or absent within a stream segment. We used a probability threshold that was equal to a given species presence/absence in the observed data set, an approach that works as well or better than other threshold approaches (Liu et al. 2005). We evaluated our predictions using *k*-fold cross validation ($k = 10$) and the receiver-operator curve (AUC), which ranges from 0-1; AUC values of 0.5 indicate that models performed no better than random and values approaching unity indicate increased model accuracy. We considered a cross-validated AUC of ≥ 0.7 to be acceptable. We also calculated the percentage of times a species was correctly assigned as being present or absent (% Correct).

Model validation statistics indicated that SDMs accurately predicted sensitive species distributions (Table 6). Although the relative influence differed among species, we found that stream discharge, SS concentration, and the nutrient index were the most influential variables when averaged across all species compared to non-water quality variables (Figure 8). We used these models to predict the presence/absence of each species within a stream segment based on water quality conditions simulated during 1990-2010 throughout the WLEB and non-water quality factors. We then summed sensitive species presences within a stream segment to calculate the sensitive species richness for that stream segment given Baseline and Grassland conditions. We assumed that a higher richness of sensitive species was indicative of a healthier fish community.

The mean sensitive species richness in the Baseline scenario was three, but varied across the watershed from 0 to 15 (Figure 9). By contrast, the mean species richness in the Grassland scenario was 4.5 and ranged from 0 to 17. These differences in sensitive species richness between the Grassland and Baseline scenarios suggests that agricultural NPS pollution is playing a major role in limiting stream fishes that are sensitive to habitat degradation. (Figure 10).

Table 5. Definitions of non-water quality variables included in species distribution models. We did not include other in-stream habitat variables known to influence fish species (e.g., in-stream cover, pool/riffle/run quality, bank erosion) because these variables were unavailable for the vast majority of streams within the WLEB and we would not have been able to forecast species distributions across the watershed if we had included them.

Abbreviation	Definition
Slope	The slope of a stream segment
Sinuosity	A measure of a stream segment's curviness
Nat. Rip.	% of a stream segment's upstream riparian area that was classified as forested, herbaceous, or wetland
Tmax	The average maximum temperature in July (1975-2013)
Road Cross.	The number of road crossings within a watershed
Dam Dist.	The distance from the nearest downstream dam
Imp. Surface	The amount of impervious surfaces (e.g., roads, parking lots, developed areas) within a watershed

Table 6. Model validation statistics for 19 fish species (species abbreviation) that are highly sensitive to water quality or habitat degradation. We considered models acceptable if the cross-validated area under the operating curve (AUC) was ≥ 0.70. The mean (± 1 S.D.) of the AUC and percentage of correctly assigned species presences/absences (% Correct) from cross validation are shown.

	In-sample statistics		Cross-validated statistics	
Species	**AUC**	**% Correct**	**AUC**	**% Correct**
Brook silverside (Brooksilver)	0.89	86	0.79 (± 0.12)	83 (± 2)
Golden redhorse (Goldred)	0.87	86	0.80 (± 0.07)	81 (± 5)
Greater redhorse (Greatred)	0.93	91	0.82 (± 0.13)	88 (± 3)
Greensided darter (Greendar)	0.85	85	0.74 (± 0.05)	75 (± 5)
Hornyhead chub (Hornyhead)	0.93	94	0.83 (± 0.11)	93 (± 3)
Logperch (Logperch)	0.88	86	0.80 (± 0.06)	79 (± 5)
Longear sunfish (Longear)	0.93	92	0.89 (± 0.03)	90 (± 3)
Mottled sculpin (Msculpin)	0.95	92	0.81 (± 0.09)	89 (± 3)
Northern hogsucker (Nhogsuck)	0.92	91	0.82 (± 0.06)	82 (± 4)
Northern pike (Npike)	0.93	91	0.78 (± 0.08)	85 (± 3)
Pumpkinseed sunfish (Pumpkin)	0.90	86	0.78 (± 0.05)	81 (± 4)
Rainbow darter (Rainbowdar)	0.88	87	0.78 (± 0.08)	79 (± 4)
Rock bass (Rockbass)	0.88	89	0.76 (± 0.05)	78 (± 5)
Sand shiner (Sandsh)	0.83	83	0.77 (± 0.07)	79 (± 4)
Shorthead redhorse (Shortheadred)	0.94	93	0.85 (± 0.09)	90 (± 3)
Silver redhorse (Silverred)	0.97	94	0.90 (± 0.06)	91 (± 2)
Smallmouth bass (Smallmouth)	0.97	96	0.82 (± 0.06)	84 (± 4)
Spotted sucker (Spotsuck)	0.83	84	0.77 (± 0.05)	79 (± 3)
Stonecat madtom (Stonecat)	0.79	75	0.75 (± 0.05)	73 (± 5)

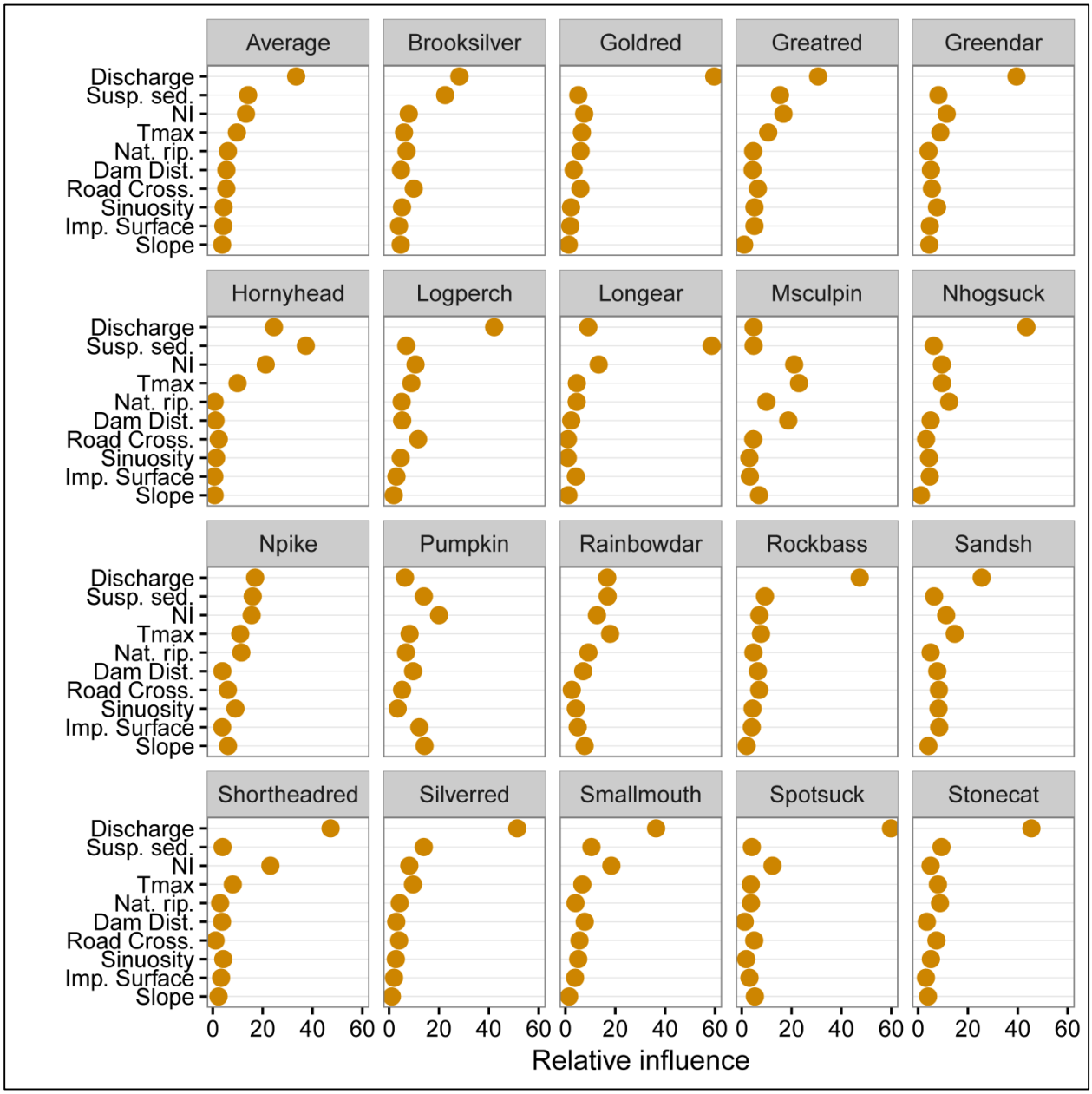


Figure 8. The relative influence of different WLEB water quality and non-water quality factors on the distribution of 19 fish species that are sensitive to habitat degradation. Water quality factors included stream discharge (Discharge), suspended sediment concentration (Susp. Sed.), and the nutrient index (NI). See Table 5 for definitions of non-water quality factors and Table 6 for fish species abbreviations. The Average panel is the average relative influence of a factor across all species. Note that the relative influence measures the absolute relative influence, whether negative or positive. In general, water quality factors had greater relative influence than other factors included in our models.

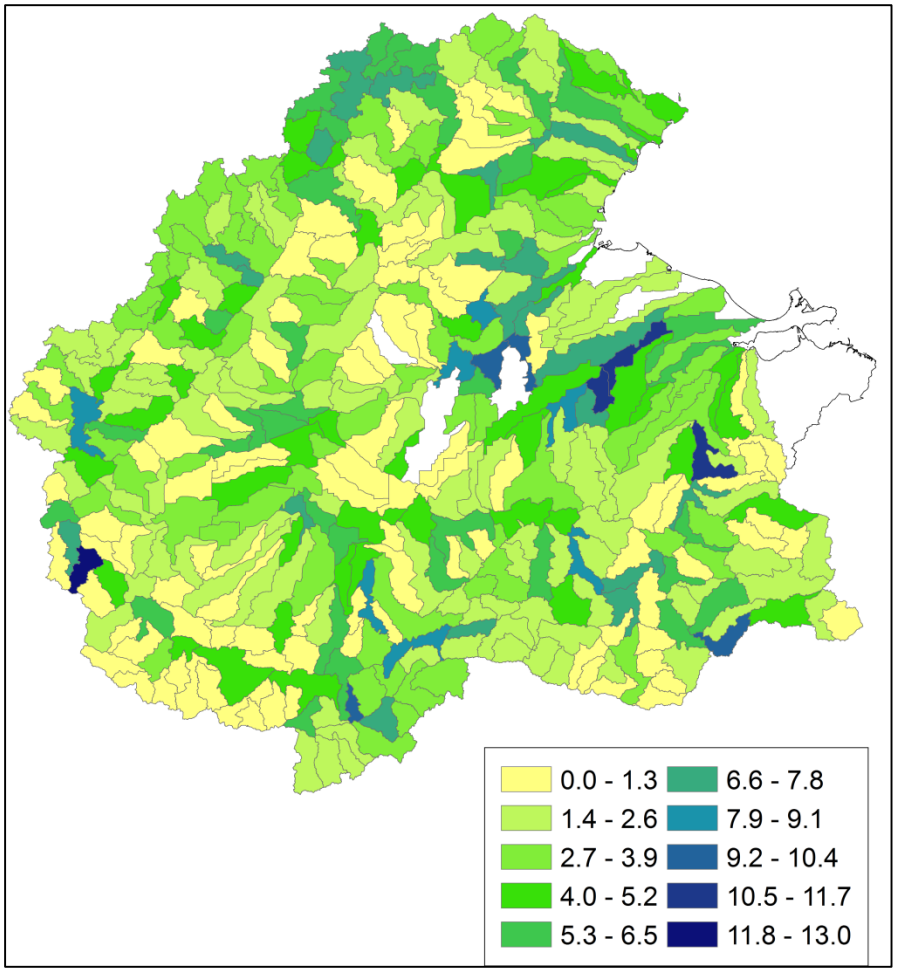


Figure 9. Forecasted sensitive species richness in the WLEB based on water quality simulated in Baseline scenario and non-water quality factors. Sensitive species richness is the average species richness within a subwatershed, based on individual species SDMs developed to forecast the presence/absence of 19 fish species sensitive to habitat degradation.

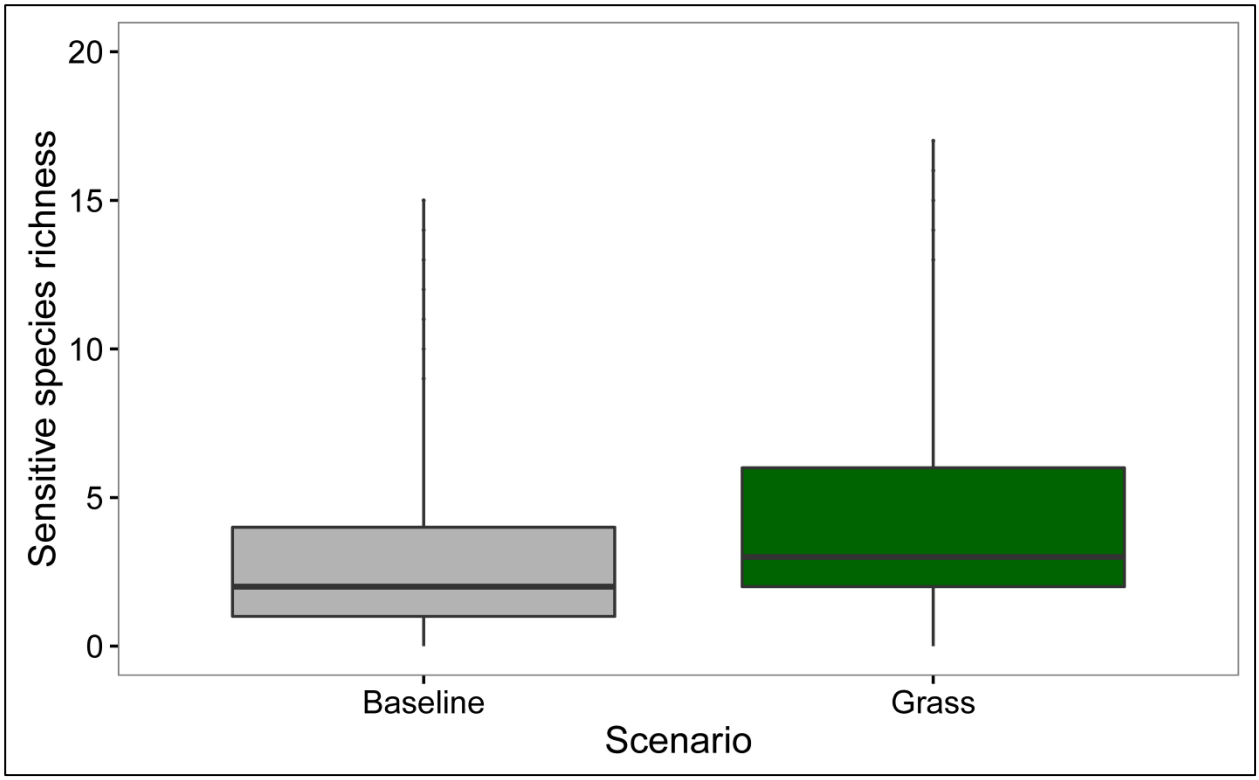


Figure 10. Forecasted sensitive species richness for the Baseline (Baseline) and Grassland (Grass) scenarios in the WLEB. Boxes show the interquartile range for forecasted sensitive species richness, with whiskers extending to the minimum and maximum of forecasted values.

OBJECTIVE 3: QUANTIFYING THE POTENTIAL ENVIRONMENTAL BENEFITS OF ADDITIONAL AGRICULTURAL CONSERVATION PRACTICE IMPLEMENTATION

Development of Conservation Scenarios

We used conservation scenarios developed by the Cropland component of the USDA NRCS Conservation Effects Assessment Project (CEAP-Cropland; USDA NRCS 2011) to simulate the potential for additional implementation of CPs to improve stream fishes by reducing water quality stressors. Our scenarios represented a gradient of increasing implementation of CPs, from treating only farm acres in high-need of CP implementation, acres in high- and moderate-need, and treating all acres. A farm acre's need for treatment was defined by the USDA NRCS according to its inherent vulnerability for nutrient and sediment loss and level of CP treatment already present (USDA NRCS 2011). The level of treatment was based on 2003-2006 farmer survey data as part of the National Resource Inventory (USDA NRCS 2011). Scenarios also consisted of either using only erosion control CPs (Table 7) or using erosion control and nutrient management (i.e., altering the amount, type, or timing of fertilizer application). We assumed an adoption rate of 80% and that the "best" option would be chosen on 75% of the treated acres to implicitly account for individual farmer behavior. This approach resulted in six conservation scenarios (3 treatment levels x 2 categories of CPs).

Conservation scenarios from CEAP-Cropland were simulated using the Agricultural Policy/Environmental eXtender (APEX) model (Gassman et al. 2009). The fractional annual reductions in edge-of-field sediment and nutrient loads associated with each CEAP-Cropland conservation scenario were used to adjust edge-of-field SWAT predictions, such that the net effect of the conservation scenario was reflected in the SWAT model. The reduced SWAT edge-of-field loads (representative of a conservation scenario) were then routed throughout the WLEB stream network using SWAT. We used our validated biological models to forecast stream fish communities and individual species based on the simulated water quality conditions in each conservation scenario to assess how much change occurred compared to the Baseline scenario (i.e., no additional CPs). The annual cost of each scenario was estimated as the total costs of planning, installation, maintenance, and forgone income on land converted to CPs.

Table 7. Potential agricultural practices implemented in WLEB conservation scenarios. Not all practices were implemented; their inclusion was determined from farmer surveys in the WLEB as part of the National Resource Inventory. Multiple practice types were potentially implemented within the same field to address both nutrient and sediment runoff concerns. All practices but nutrient management planning fall within the erosion control category.

Conservation Practices	NRCS Code
Residue management tillage	329
Contour farming	330
Cover crops	340
Wind break or shelter belt	380
Field border	386
Riparian herbaceous buffer	391
Riparian forest buffer	392
Filter strips	393
Hedgerows	422
Contour strip cropping	585
Cross wind strips, traps, etc.	589
Nutrient management planning	590
Terrace	600
Herbaceous wind barrier	603
Surface roughening	609

Development of Additional Scenarios and Prioritizing Subwatersheds for CPs

While the conservation scenarios described above allowed us to understand how much benefit CPs may be expected to provide as the result of water quality improvement and how much additional financial investment in CPs may be needed, they did not allow us to identify where in the watershed CPs might provide the most benefit. Therefore, we developed a separate set of scenarios that allowed us to identify subwatersheds within the WLEB where CPs would be most likely to provide a benefit to stream fishes. These scenarios used our simulated Baseline conditions as a starting point. We then "manually" reduced the amount of TN, TP, and sediments in agricultural runoff and leaching by 20%, 40%, 60%, 80% and 99% to create five additional scenarios. Hereafter, we use the term "reductions in agricultural inputs" to encompass these reductions in agricultural runoff and leaching. We forecasted potential changes in stream fish metrics that resulted from these reductions and compared subwatersheds to identify where in the WLEB the largest improvements in stream fishes occurred (Figure 11).

We combined estimates of stream fish improvement with an agricultural threat index to identify subwatersheds where additional CP implementation will likely provide the most benefit to stream fishes. We calculated the agricultural threat index following the approach of Fore et al. (2014). The threat index scales from -4 to 4, with positive values indicating an increasing threat of agricultural relative to non-agricultural threats and negative values indicating an increasing amount of non-agricultural threats (Table 8). Importantly, this threat index only considers the threat of row-crop agriculture; other forms of agriculture were not considered agricultural threats (e.g., cattle operations). Threat index scores ≥ 2 indicate watersheds where in-field and edge-of-field CPs have the greatest potential to improve stream health (Figure 12; Fore et al. 2014). We identified subwatersheds with an agricultural threat index ≥ 2 and where improvements in stream fish communities were in the 80th percentile as potential priority subwatersheds for CPs.

We also used the agricultural runoff reduction scenarios to examine the relationship between stream fish metrics (IBI and Piscivore Index) and the percent reduction in agricultural inputs. Understanding the shape of these relationships (i.e., the amount of effort needed to improve stream fishes) could help inform management. We speculated that the relationship between stream benefits, as measured by the percentage of stream kilometers in which water quality was not limiting biological conditions, and reductions in agricultural inputs would follow one of three general shapes (Figure 13):

(1) Linear: would suggest that benefits for stream fishes continued to increase at a constant rate with larger reductions in agricultural inputs. This relationship can be described by a linear model -

$$y = \beta_0 + \beta_1 x + \varepsilon$$

β_0 is the estimate for the model intercept, β_1 is the estimated effect of reducing agricultural inputs by a given percentage (x), and ε is an error term.

(2) Concave down: would indicate that the rate of stream fish improvements is higher at lower reductions in agricultural inputs compared to higher inputs. This would suggest that the largest benefits were provided initially. This relationship can be described by a quadratic model -

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \varepsilon$$

A negative estimate of β_2 would indicate a concave down relationship.

(3) Concave up: would indicate that improvements in stream fishes were small at lower reductions in agricultural inputs, but benefits increased with larger reductions. This would suggest that reductions in agricultural inputs must achieve some threshold before high rates of improvements in stream fish communities are possible. This relationship can also be described by a quadratic model, with a positive estimate of β_2 indicating a concave up relationship.

We examined the relationship between forecasted benefits for stream fish community metrics and percent reductions in agricultural inputs by comparing linear and quadratic models using an information theoretic approach (Burnham and Anderson 1998). We fit hierarchical models that allowed for slopes and intercepts to vary at the subwatershed scale (HUC-12) to account for autocorrelation that may result from repeatedly "sampling" subwatersheds in each reduction scenario. Percentage data were logit-transformed to meet normality assumptions of linear regression. Percent reductions in agricultural inputs were standardized to have a mean of zero and standard deviation of one to improve model convergence. Significance of coefficient estimates for fixed effects were tested with a parametric bootstrap (N = 500 replicates). We fit mixed models using the 'lme4' package (Bates et al. 2015) in the R statistical environment (R Core Team 2015). Models were fit with maximum likelihood to allow for model comparison (Bates et al. 2015).

Table 8. Variables used to develop an agricultural threat index for the WLEB. Local Variable Only refers to variables that were only available at the scale of the local catchment (i.e., NHDPlus catchment). All other variables were also calculated at the watershed-scale (i.e., upstream drainage area). Ag Threat and Non-Ag Threat refer to whether a threat variable was considered an agricultural threat (Ag. Threat) or a non-agricultural threat (Non-Ag Threat). Data sources (Data Source) of threat variables were from the Great Lakes Basin Fish Habitat Partnership (GLBFHP; http://greatlakes.fishhabitat.org/) or the Great Lakes Aquatic Gap (GLGAP; http://wi.water.usgs.gov/gap/index.htm) datasets.

Threat Variables\	Description	Local Variable Only	Ag Threat	Non-Ag Threat	Data Source
Dams	National Inventory of Dams, 2002-2004: Number of dam(s) present in catchment			X	GLBFHP
Ground-water use	USGS National Atlas of the US: Ground Water Use by COUNTY 2000: Millions gallons per day/km^2			X	GLBFHP
Surface-water use	USGS National Atlas of the US: Surface Water Use by COUNTY 2000: Millions gallons per day/km^2			X	GLBFHP
Cattle	Agricultural Census 2002, 1:2M scale, INTEGER: average number of cattle/acre farmland			X	GLBFHP
Road stream crossings	Census 2000 TIGER Roads, 1:100K scale, road crossings identified by INTERSECT, with points generated, #/km^2			X	GLBFHP
Mines	USGS Active Mines and Mineral Processing Plants, 2003, #/km^2			X	GLBFHP
Toxic release inventory sites	USEPA, 2007: #/km^2 Toxics Release Inventory Program sites			X	GLBFHP
NPDES sites	USEPA, 2007: #/km^2 National Pollutant Discharge Elimination System (NPDES) sites			X	GLBFHP

CERCLIS sites	USEPA, 2007: #/km^2 Compensation and Liability Information System (CERCLIS) sites			X	GLBFHP
Impervious surface	NLCD 2006 percent impervious, average, catchment			X	GLBFHP
Barrier between stream segment and Great Lake	100 = dam/waterfall between target arc and Great Lakes system; 1 = no dam/waterfall between target arc and Great Lakes system	X		X	GLGAP
Distance to upstream dam	Distance from closest upstream segment coded as dam (measured from downstream end of dam segment) to target segment	X		X	GLGAP
Distance to downstream dam	Distance from closest downstream segment coded as dam (measured from upstream end of dam segment) to target segment	X		X	GLGAP
Size of closest upstream impoundment	Area of closest lake/impoundment with area >= 5 acres upstream of target segment	X		X	GLGAP
Size of closest downstream impoundment	Area of closest lake/impoundment with area >= 10 acres downstream of target segment	X		X	GLGAP
Riparian non-row crop agriculture	Riparian agriculture, non-row crop (1992 NLCD)			X	GLGAP
Riparian developed lands	Riparian urban/developed classes, classes 11, 12, 13, 14 (1992 NLCD)			X	GLGAP
Riparian row crop agriculture	Riparian agriculture classes 22 and 23; row crop and Orchards/Vineyards/Other (1992 NLCD)		X		GLGAP

Pasture/Hay	NLCD 2006 percent pasture/hay, class 81 (2006 NLCD)		X	GLBFHP
Cultivated crops	NLCD 2006 percent crop land, class 82 (2006 NLCD)	X		GLBFHP

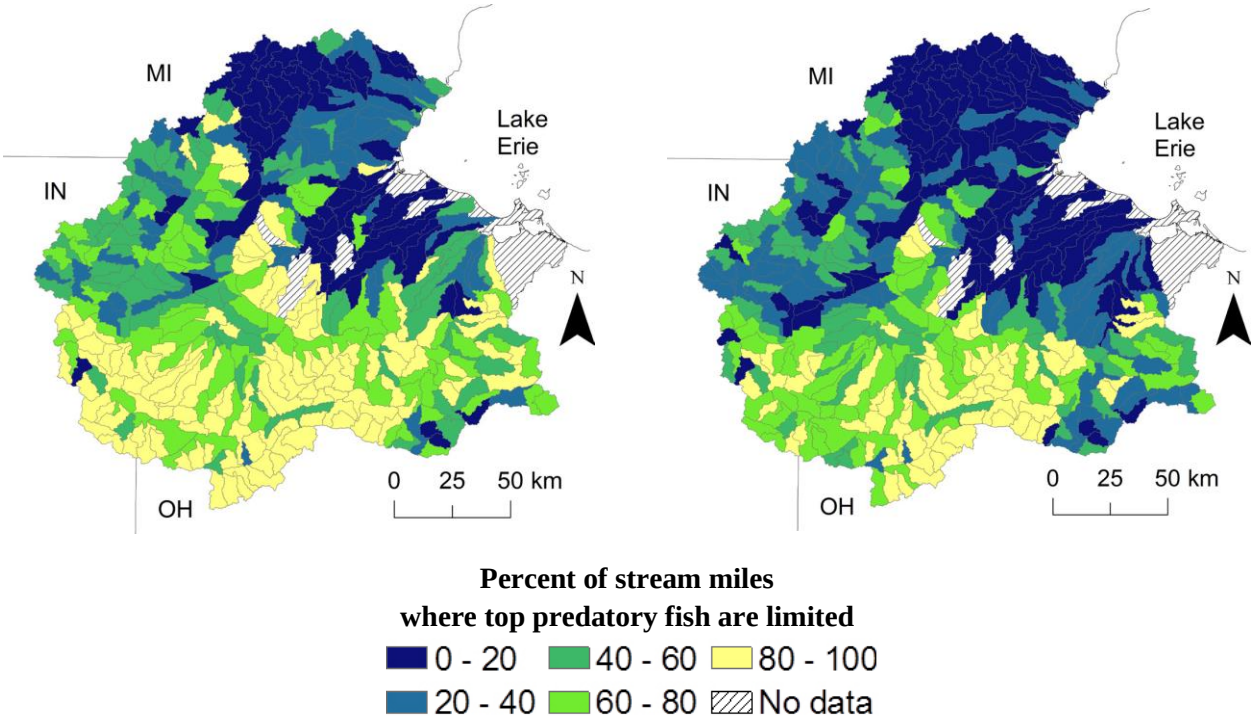


Figure 11. Maps showing the benefits even small improvements in water quality can make to fish communities in WLEB streams. Map on the left shows the percent of stream kilometers in each subwatershed where populations of top predatory fish species, like largemouth bass, are predicted to be limited by sediments and nutrients. Map on the right shows the same percentages after a simulated 20% reduction in sediments and nutrients. Note: The maps are based on model estimates and not direct observation.

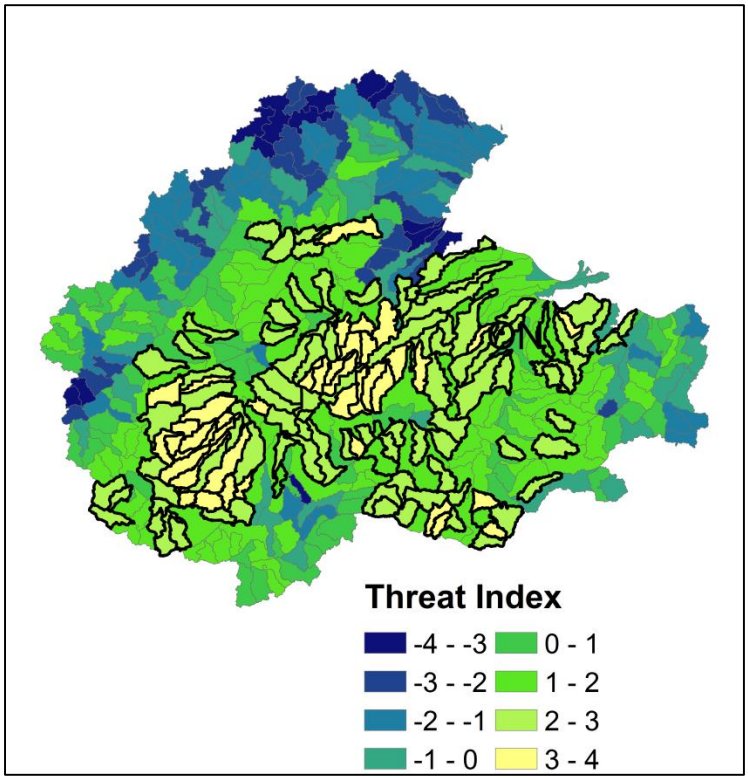


Figure 12. Threat index for the Western Lake Erie Basin. This index scales from -4 to 4, with higher positive values indicating that agricultural threats are increasing relative to non-agricultural threats (e.g., point sources of pollution, dams, urban development). Threat index scores ≥ 2 are outlined in bold and indicate subwatersheds where row-crop agricultural is likely the primary threat to stream health (Fore et al. 2014).

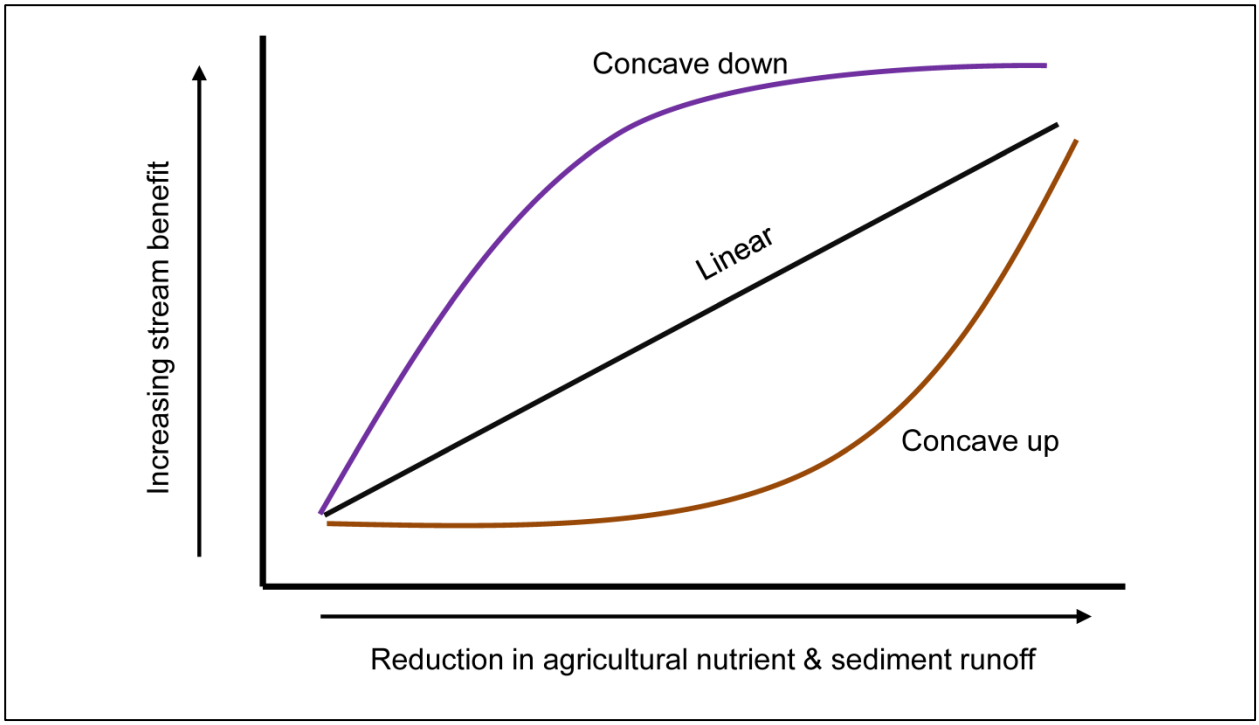


Figure 13. Generalized potential relationships between stream biological benefits and reductions in agricultural nutrient and sediment inputs. See text for descriptions of possible relationships. We tested the shape of these relationships using a model selection approach.

Potential Benefits of CP Implementation to Stream Fishes

Our simulations suggest that *widespread* implementation of CPs throughout the WLEB has the potential to improve stream fish communities by reducing water quality as a limiting factor. Treating only those farm acres in high-need of CPs provided a relatively modest benefit in forecasted biological conditions compared to more widespread CP implementation (Figure 14). Treating only high-need acres improved the average IBI score by ~1%, the Piscivore Index by as much as 17%, and sensitive species richness by as much as 1% relative to the Baseline conditions. By contrast, treating farm acres in high- and moderate-need improved the average IBI score by as much as 6%, the Piscivore Index by as much as 42%, and sensitive species richness by as much 7% relative to Baseline conditions. Further improvements in forecasted biological conditions appear possible if acres in low-need are also treated with CPs.

We also observed large increases in the percentage of the watershed where water quality was no longer limiting stream fishes with *widespread* CP implementation (Figure 15). For example, treating all (i.e., high-, moderate-, and low-need) farm acres with erosion control and nutrient management removed water quality limitation, based on the IBI, in an additional 19% of streams (3,901 km) and an additional 41% of streams (8,287 km) for the Piscivore Index compared to Baseline conditions.

Additional Benefits of Including Nutrient Management

Improvements in stream fish communities were larger when nutrient management was included in addition to erosion control CPs. Including nutrient management resulted in ~1,800 km and ~2,300 km fewer streams being limited for the IBI and Piscivore Index, respectively, compared to only including erosion control CPs. Moreover, while additional improvements in sensitive species richness were modest (~8%), in general, average sensitive species richness only increased when nutrient management was included in addition to erosion control CPs (Figure 15).

While including nutrient management resulted in overall greater potential improvements in stream fish communities, the additional annual financial investment was also considerably higher than only implementing erosion control practices (Table 9). We estimated that including nutrient management would cost about twice as much on average for treating comparable farm acre needs than only using erosion control CPs. Thus, including nutrient management may not be a more *cost-effective* strategy.

To help understand the cost-effectiveness of including nutrient management, we examined the return on investment (ROI) of our conservation scenarios. The ROI is simply the forecasted benefit of a conservation action divided by the cost of that action and is a useful metric to help maximize the benefits of limited conservation resources (Naidoo et al. 2006; Murdoch et al. 2007). The estimated ROI from our conservation scenarios highlights two important insights for stream conservation in the WLEB. First, nutrient management (plus erosion control) generally represented a more cost effective management strategy than erosion control alone (Figure 16). For the biological metrics that we considered, the addition of nutrient management provided a similar ROI for the Piscivore Index, but provided a 3-fold and 16-fold larger ROI for the IBI and sensitive species richness, respectively. Second, treating farm acres in high-need first was a more cost effective management option than also treating moderate- or

low-needs acres. Treating acres in high need resulted in a ROI that was 2.8-fold to 4-fold larger than treating acres in low- and moderate -need across all fish metrics. Treating acres in low-need resulted in further reductions in the ROI compared to less-intensive conservation scenarios. Thus, accurately identifying acres in high-need of CPs and treating those acres first may help maximize the benefits of limited conservation resources in the WLEB. Treating these high-need farm acres first may also increase the likelihood of rapid and measurable improvements in stream health, which is important for maintaining momentum for long-term conservation efforts. However, because treating high-need acres alone would only improve conditions in a relatively small portion of streams; the need exists to also treat moderate- and low-needs acres if we truly want to reduce water quality as a limiting factor of stream health throughout the WLEB.

Benefits Relative to the Grassland Scenario

Relative to the Grassland scenario, even widespread implementation of erosion control and nutrient management CPs fell short of removing all water quality limitation (Table 9). Even in the most intensive conservation scenario (i.e., erosion control and nutrient management applied to high-, moderate-, and low-need acres), water quality would still limit the IBI in more than 8,000 km of streams compared to just 1,051km in the Grassland scenario. Similarly, water quality would still limit the Piscivore Index in ~2,000 more km of streams in the most intensive conservation scenario as compared to the Grassland scenario, and sensitive species richness would be about 1.2-fold lower. Assuming the Grassland scenario represents a theoretical upper benchmark of potential conditions[3], our results suggest that expectations for CPs be realistic. While large improvements are possible with widespread CP implementation, eliminating water quality as a limiting factor to stream health across the whole watershed solely by adding CPs is unlikely.

[3] Ignoring other sources of water quality pollutants in the watershed (e.g., urban runoff, point sources, and concentrated animal feeding operations).

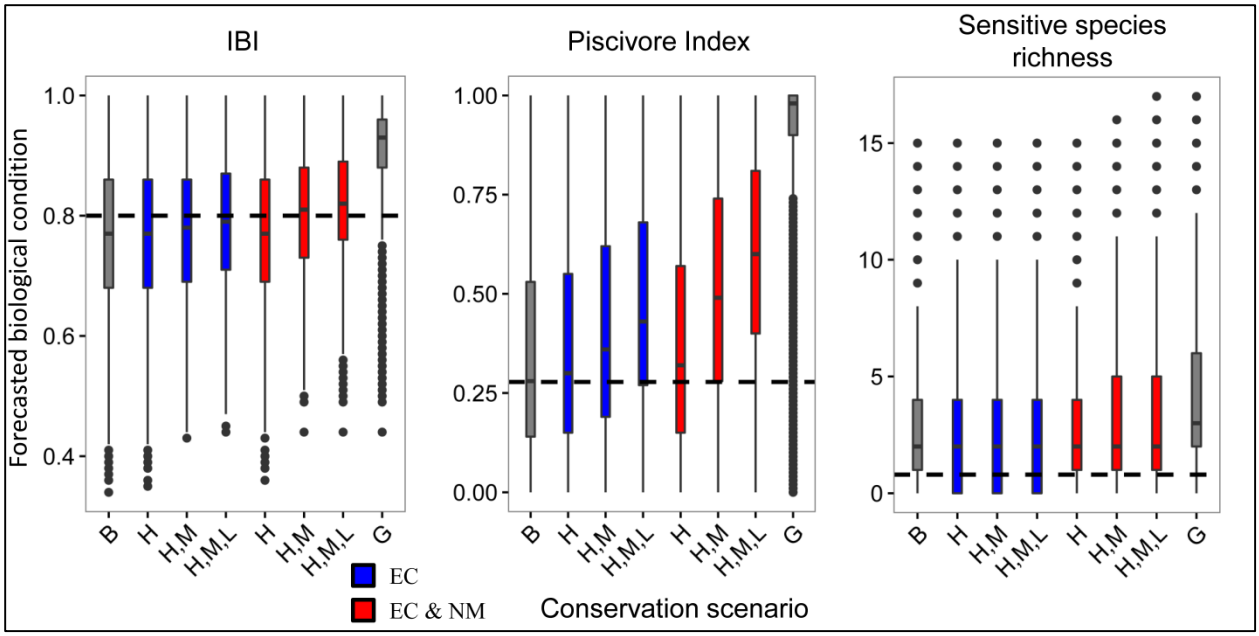


Figure 14. Forecasted biological conditions in conservation scenarios. Biological conditions included an Index of Biotic Integrity (IBI), relative abundance of piscivorous species (Piscivore Index), and sensitive species richness. EC = erosion control practices implement, EC & NM = erosion control practices and nutrient management implemented; B = Baseline conditions; G = Grassland scenario; H = treated farm acres in high-need, H,M = treated farm acres in high- and moderate-need, H,M,L = treated farm acres in high-, moderate-, and low-need.

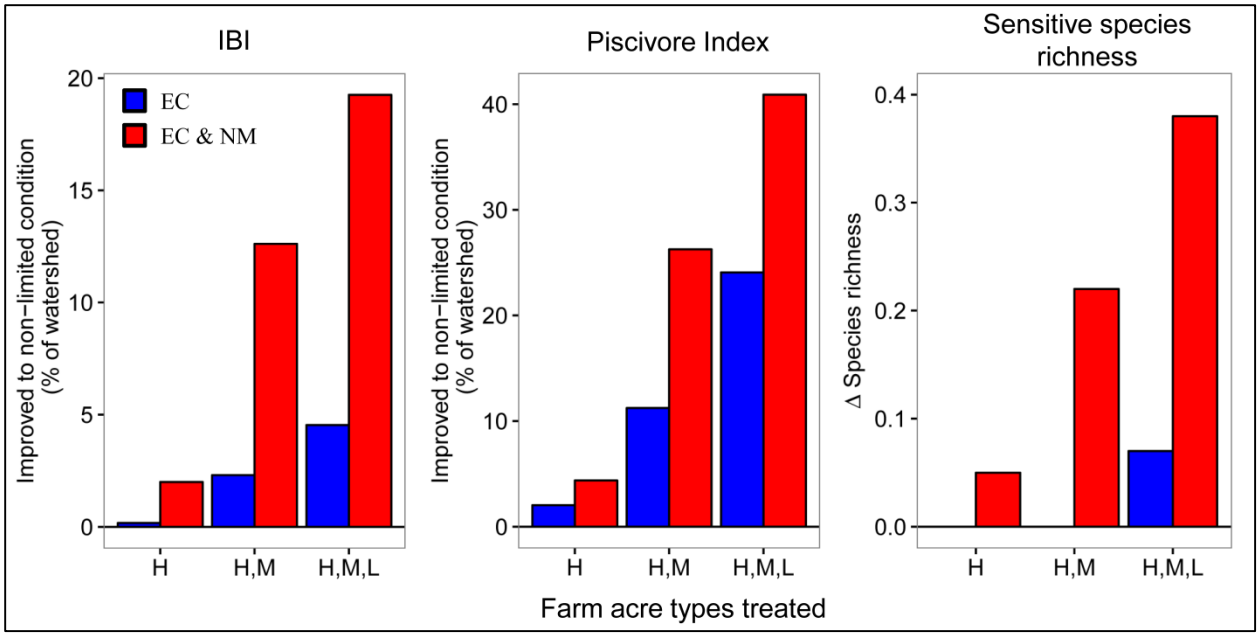
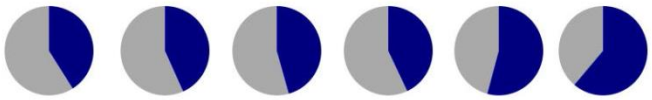
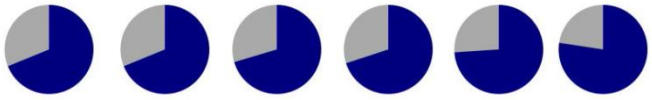


Figure 15. Forecasted benefits of additional investment in agricultural conservation practices for improving stream biological conditions, measured as the change from Baseline conditions in the stream kilometers (% of watershed) water quality was not limiting stream fish communities or the change in sensitive species richness. Biological conditions included an Index of Biotic Integrity (IBI), relative abundance of piscivorous species (Piscivore Index), and sensitive species richness. EC = erosion control practices implement, EC & NM = erosion control practices and nutrient management implemented; H = treated farm acres in high-need, H,M = treated farm acres in high- and moderate-need, H,M,L = treated farm acres in high-, moderate-, and low-need.

Table 9. Benefits of increasing investment in agricultural conservation practices relative the Grassland scenario, which represents an upper benchmark in potential stream biological conditions. EC = erosion control; EC & NM = erosion control and nutrient management; High = treated farm acres in high-need; High & Mod. = treated farm acres in high- and moderate-need; All = treated farm acres in high-, moderate-, and low-need; ● = minimal benefit; ● = maximum benefit.

Conservation Practices	Practice types	EC	EC	EC	EC & NM	EC & NM	EC & NM
	Acres treated	High	High & Mod.	All	High	High & Mod.	All
	Additional annual cost ($ million)	4.5	55.78	128.26	8.39	149.25	263.39
Stream Biological Conditions	IBI (amount of watershed in which water quality was not limiting)						
	Piscivore Index (amount of the watershed in which water quality was not limiting)						
	Sensitive species richness (average species richness for the whole watershed)						

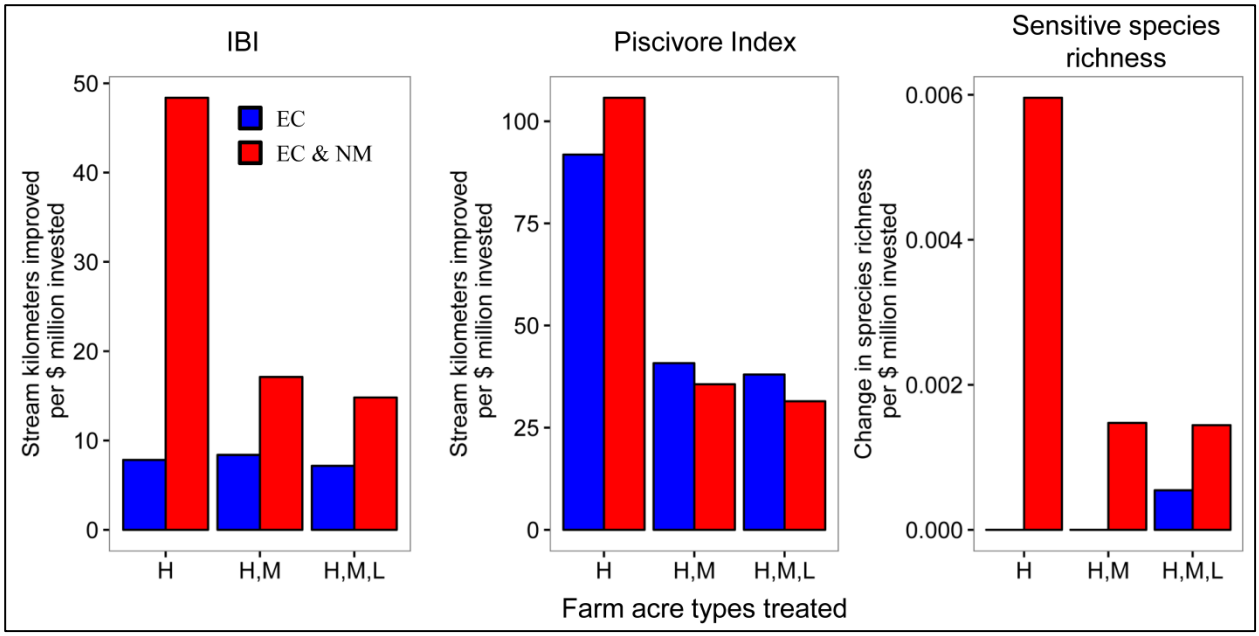


Figure 16. Return on investment of different conservation scenarios. Return on investment was estimated as the change in biological metrics from baseline conditions divided by the estimated annual cost ($ million) of a conservation scenario. Biological conditions included an Index of Biotic Integrity (IBI), relative abundance of piscivorous species (Piscivore Index), and sensitive species richness. EC = erosion control; EC & NM = erosion control and nutrient management.

Relationships between Biological Improvements and Percent Reductions in Agricultural Runoff

A quadratic relationship best described how fish community health would respond to reductions in agricultural inputs in WLEB subwatersheds (Table 10). Coefficient estimates suggest a concave-up relationship for the IBI and a concave-down relationship for the Piscivore Index (Table 11), although plots of these relationships show they only slightly differ from linearity for both biological metrics (Figure 17). These findings suggest that large reductions in agricultural inputs would be necessary maximize improvement in the IBI, whereas piscivorous fish would benefit with more moderate reductions in agricultural inputs. Based on these relationships, we would expect stream biological conditions to improve nearly linearly with increasing reductions in agricultural NPS inputs, but to slow down as reductions approach ~80% and ~40% for the IBI and Piscivore Index, respectively.

Table 10. Model selection statistics for mixed models describing the relationship between streams that were not limited by water quality and percent reductions in agricultural inputs. A lower AIC_c indicates a more parsimonious model. Models within two AIC_c units of the best model ($\Delta AIC_c \leq 2$) are generally considered equally parsimonious (Burnham and Anderson 1998). The models compared were a linear model, a quadratic model, and an intercept-only model (Null). The $R^2_{marginal}$ describes to the amount of variance explained by the fixed effect component of the model (i.e., reduction in agricultural inputs) and the $R^2_{condition}$ describes the total amount of variance explained by the model (Nakagawa and Schielzeth 2013).

IBI				
Model	**AIC_c**	**ΔAIC_c**	**$R^2_{marginal}$**	**$R^2_{conditional}$**
Linear	6094.42	382.61	0.41	0.89
Quadratic	5711.81	0.00	0.42	0.93
Null	9078.25	3366.44	0.00	0.28
Piscivore Index				
Linear	7200.67	667.53	0.30	0.86
Quadratic	6533.14	0.00	0.35	0.92
Null	9840.34	3307.20	0.00	0.28

Table 11. Coefficient estimates (± 1 S.E.) from hierarchical models examining relationships between reductions in agricultural inputs (% Reduction) and biological benefits, as measured by the percentage of a streams within a subwatershed not limited by water quality. Boldface indicates statistically significant estimates based on the parametric bootstrap. IBI = Index of biotic integrity; Piscivore Index = Relative abundance of piscivorous fishes. Note that coefficients are for standardized reductions in agricultural inputs and logit-transformed biological metrics.

Stream fish community metric	Intercept	% Reduction	%Reduction^2
IBI	**0.86 (0.07)**	**1.22 (0.03)**	**0.21 (0.02)**
Piscivore Index	**1.80 (0.09)**	**1.33 (0.04)**	**-0.13 (0.03)**

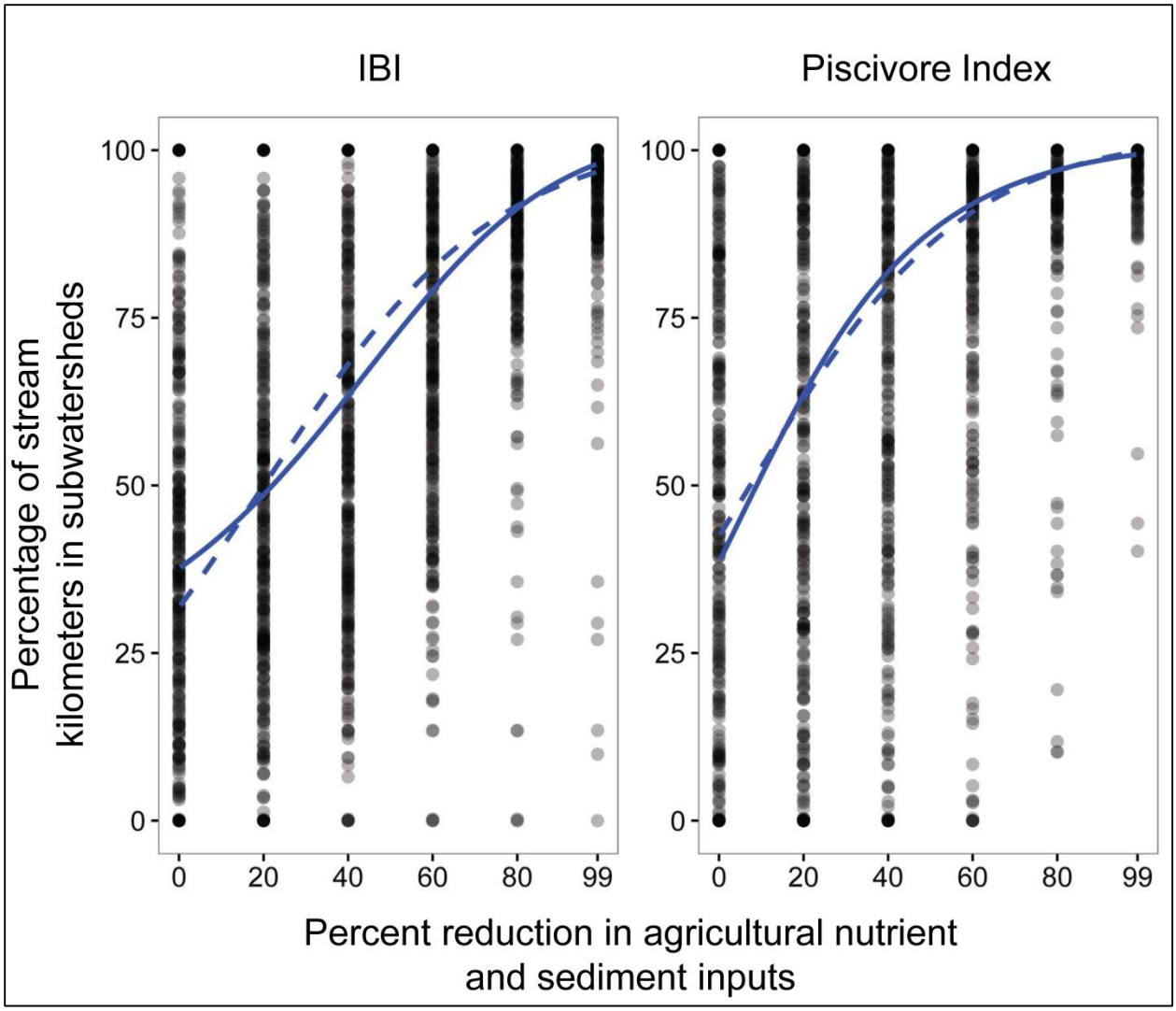


Figure 17. Relationships between stream biological benefits, measured as the percentage of stream kilometers where biological conditions were not limited by water quality within a subwatershed, and percent reduction in agricultural inputs relative to estimated conditions during 1990-2010. Dashed lines show the estimated relationships from linear models and solid lines show estimated relationships from quadratic models. Points show the forecasted percentage of each subwatershed that was not limited by water quality with reductions in agricultural input. Darker points indicate a higher number of overlapping subwatersheds. Note that the linear relationship appears non-linear because data were logit-transformed for analyses but are shown on their original scale (0-100%). The quadratic model was identified as the best model for both the IBI (Index of Biotic Integrity) and the Piscivore Index (Relative abundance of piscivorous fishes).

Example of Subwatershed Prioritization

Thus far, we have presented results that integrate responses across the watershed. However, owing to the spatially-explicit nature of our watershed model, we also had the ability to explore how different scenarios might influence subwatersheds within the WLEB. Here, we provide an example of coupling physical and biological models to prioritize subwatersheds for CP implementation. **Importantly though, this example is meant to demonstrate the potential utility of this approach to inform watershed management. The criteria we used should be modified to address specific management questions.**

In general, we found heterogeneity in the response of biological conditions within the WLEB to reductions in agricultural runoff (Figure 18). For example, with a 40% reduction in agricultural inputs, the average change in IBI scores ranged from 0 to 0.15 across subwatersheds, whereas the average change in the Piscivore Index ranged from -0.15 to 0.29. This large variation in biological responses suggests that reducing agricultural inputs, and thus, CP implementation, may have larger benefits in some subwatersheds than others.

From this analysis, we identified 79 subwatersheds based on the IBI and 70 subwatersheds based on the Piscivore Index that should be prioritized for additional CPs because the average change in biological conditions exceeded the 80th percentile (Figure 19). Of these, 30 subwatersheds were identified as a priority for both metrics and 15 of these also had an agricultural threat index ≥ 2. Using our biological and threat criteria, these 15 subwatersheds would be identified as the focus of agricultural conservation efforts because they show the highest potential for improved biological health with reduced agricultural inputs and row-crop agriculture appears to represent the dominant threat in them.

An important limitation of this approach is that we were unable to separate biological benefits that resulted from local reductions in agricultural inputs (i.e. within a subwatershed) from cumulative reductions in upstream agricultural inputs. Stream networks are connected longitudinally by stream flow (Ward 1989); with upstream inputs influencing downstream conditions (Alexander et al. 2000, Peterson et al. 2001, Dodds and Oakes 2008). Thus, priority subwatersheds located further down within drainage networks may not benefit as much as expected from local reductions in agricultural inputs if upstream sources are not also addressed. This suggest that prioritization schemes based on our approach should consider prioritizing "headwater" subwatersheds first, because we have more confidence that forecasted benefits were the result of local reductions in agricultural inputs.

Another important limitation is that while we validated our SWAT model and biological models, they are still models of reality, and some degree of error is likely in regards to the spatial distribution of agricultural inputs and expected benefits. Therefore, we recommend that our approach be combined with on-the-ground knowledge of the watershed by local stakeholders to ensure that practices are getting to the right places. We also suggest that case studies in smaller subwatersheds, using our approach to guide the placement of practices with extensive pre- and post-monitoring, would be extremely beneficial for validating our approach.

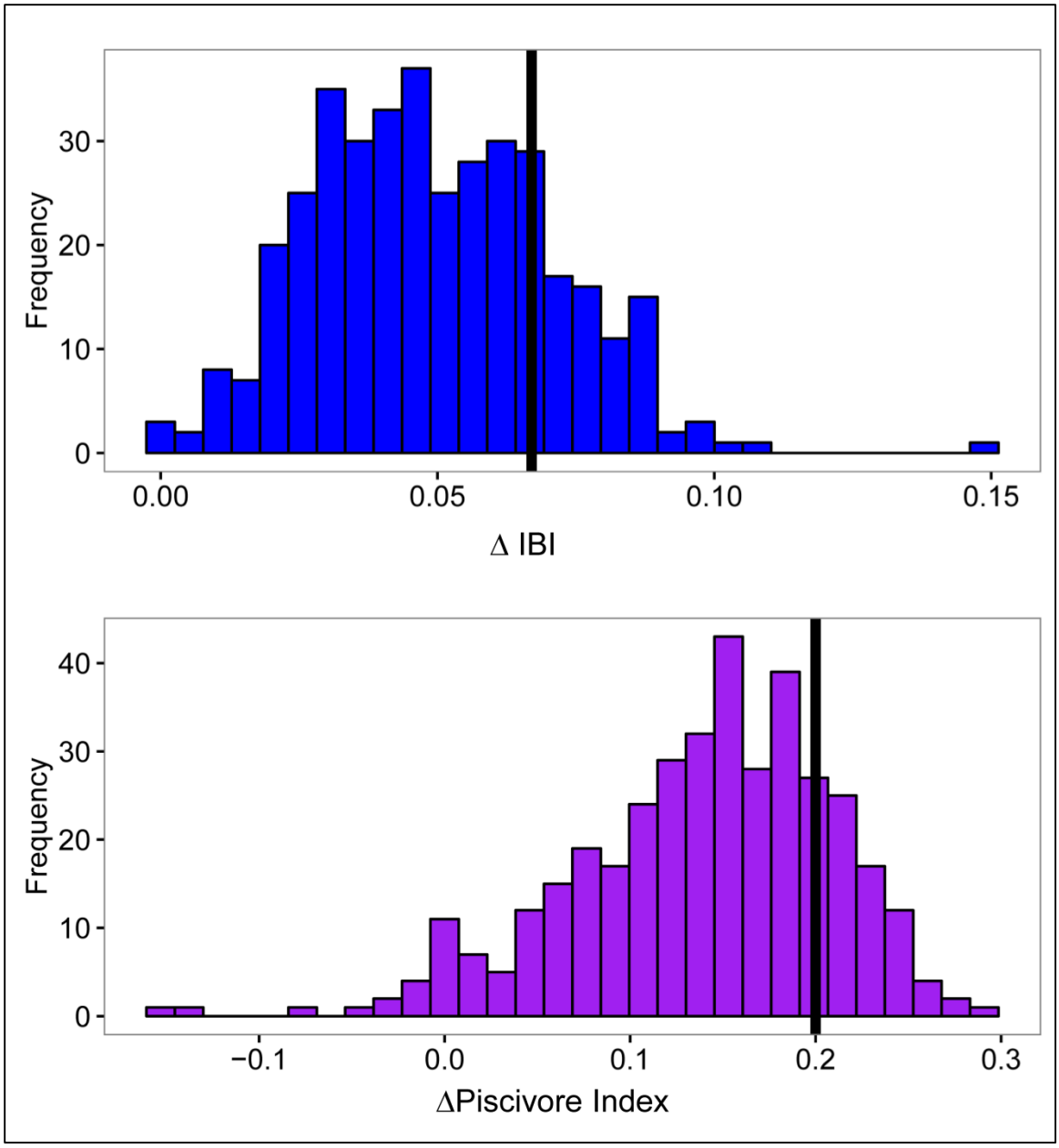


Figure 18. Histograms of the average changes in biological conditions that occurred within a subwatershed when agricultural inputs were reduced by 40% in the WLEB. This 40% reduction represents a manual reduction in nutrient and sediment inputs from agricultural land-uses. The solid black lines indicate the 80th percentile threshold for changes in biological conditions.

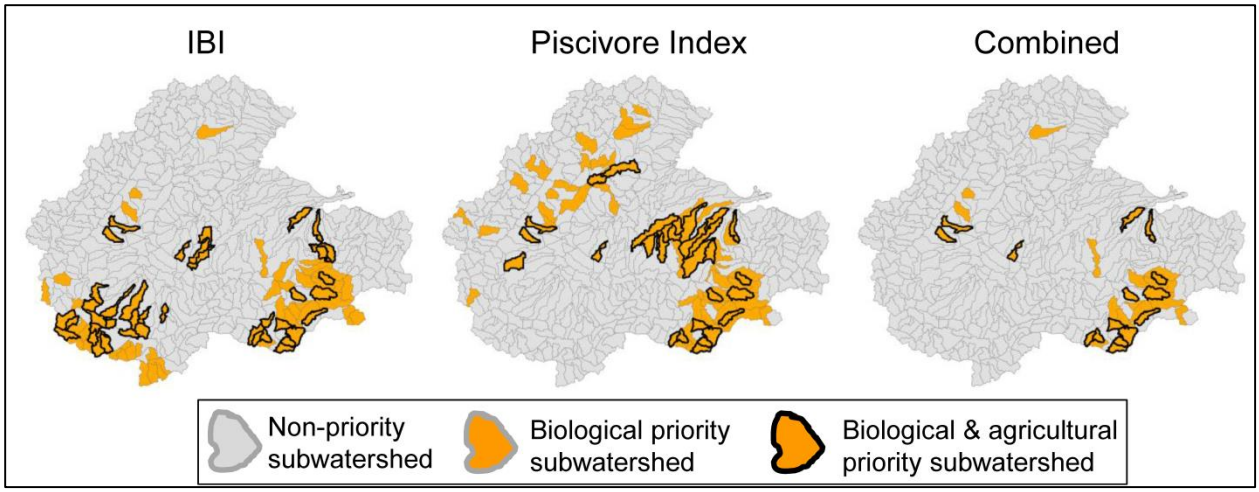


Figure 19. Subwatersheds within the WLEB that could be prioritized for conservation practices when agricultural inputs were reduced by 40%. This 40% reduction refers to both nutrients and sediments. These "biological priority" watersheds were ones in which the average change in stream biological conditions was in the 80th percentile for the index of biotic integrity (change ≥ 0.067) or the relative abundance of piscivorous species (change ≥ 0.20). The combined map shows priority subwatersheds where both biological metrics were identified as priorities. Subwatersheds outlined in bold are those where row-crop agriculture was identified as a dominant threat and agricultural conservation practices could provide benefits towards improving stream health.

Potential for "Win-Wins" for Lake Erie Water Quality Management and Stream Conservation

Water quality management in the WLEB is currently focused on reducing phosphorus loading to Lake Erie, with less consideration for improving stream water quality and biological health. As a result, the amount to which CPs may simultaneously benefit the related but unique conservation endpoints of Lake Erie and WLEB streams remains unclear. We used our conservation scenarios to quantify the potential benefits of CPs for reducing spring nutrient and sediment loading from four major WLEB tributaries (Maumee River, Portage River, River Raisin, and Sandusky River). We then compared Lake Erie TP loading reductions to stream benefits to identify potential "win-wins" for Lake Erie and stream conservation as the result of additional CP implementation.

Our modeling suggests that large reductions in spring nutrient and sediment loading to Lake Erie were possible with additional CP implementation (Figure 20). Similar to results for stream biological conditions, widespread implementation appears necessary to achieve large reductions in NPS pollution to Lake Erie. Treating only those farm acres in high-need resulted in average reductions that were 8-fold and 4-fold less than treating farm acres in high- and moderate-need for SS and nutrients, respectively. Moreover, treating only high-need acres would likely fail to meet TP loading goals for the WLEB as established by the Great Lakes Water Quality Agreement (GLWQA 2016); however, TP loading goals could potentially be met with additional CP implementation (e.g., moderate-need acres also treated with CPs; Keitzer et al. 2016; Appendix C). This result, that widespread implementation is needed to achieve water quality goals, is supported by other watershed-scale studies attempting to quantify the water quality benefits of CPs in the Great Lakes Basin (Hobbs et al., 2002; USDA NRCS 2011; Einheuser et al. 2012, Bosch et al. 2013, Scavia et al. 2016, USDA NRCS 2016). For example, a recent multi-model simulation study found that at least 25% of the WLEB watershed would need CP treatment to meet the TP loading goal, while more than 50% would need CP treatment to meet the dissolved reactive phosphorus loading goal recommended by the GLWQA (Scavia et al. 2016).

While our results suggested that implementing CPs would benefit both Lake Erie and streams in its watershed, the nature of CP implementation can greatly affect the degree to which stream health improves (Figure 21). For example, treating farm acres in high- and moderate-need with only erosion control CPs should help achieve the 40% TP loading reduction for Lake Erie (GLWQA 2016); however, the benefits for stream biological conditions were much less than was possible in more intensive conservation scenarios. To achieve a "win-win" for Lake Erie and its surrounding WLEB tributaries, wider implementation of erosion control CPs and nutrient management is necessary.

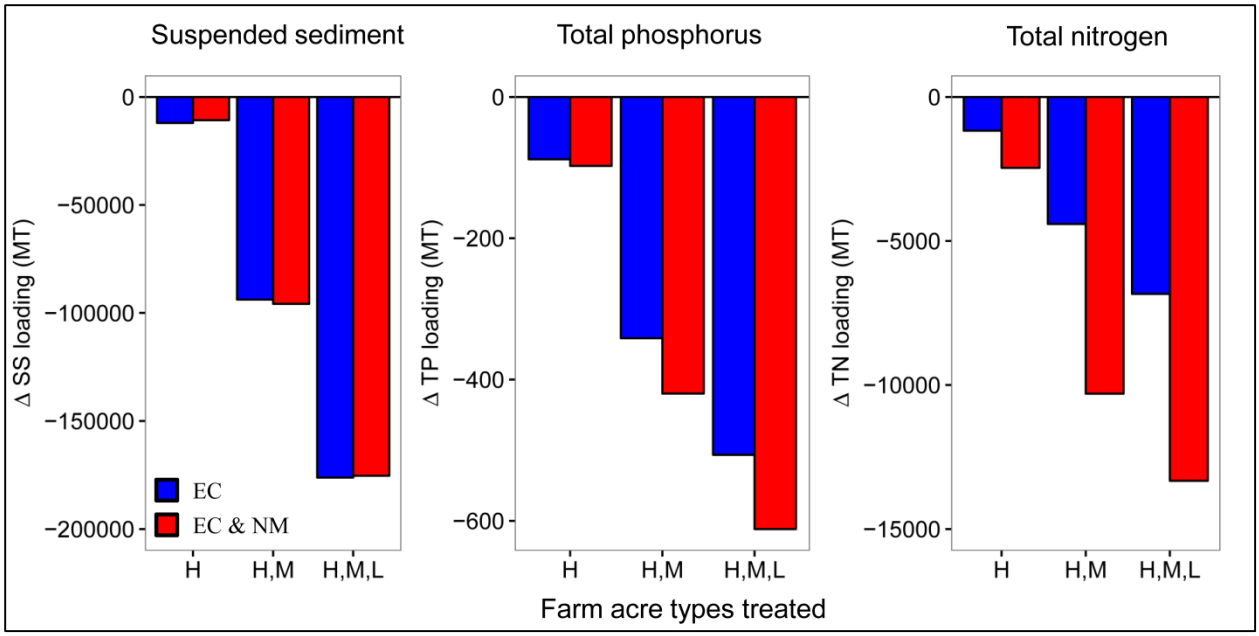


Figure 20. Forecasted benefits of additional investment in agricultural conservation practices for reducing spring nutrient and sediment loading into Lake Erie, measured as the change from baseline conditions in total spring loading from the Maumee River, Portage River, River Raisin, and Sandusky River. EC = erosion control practices implement, EC & NM = erosion control and nutrient management practices implemented; H = treated farm acres in high need, H,M = treated farm acres in high and moderate need, H,M,L = treated farm acres in high, moderate, and low need.

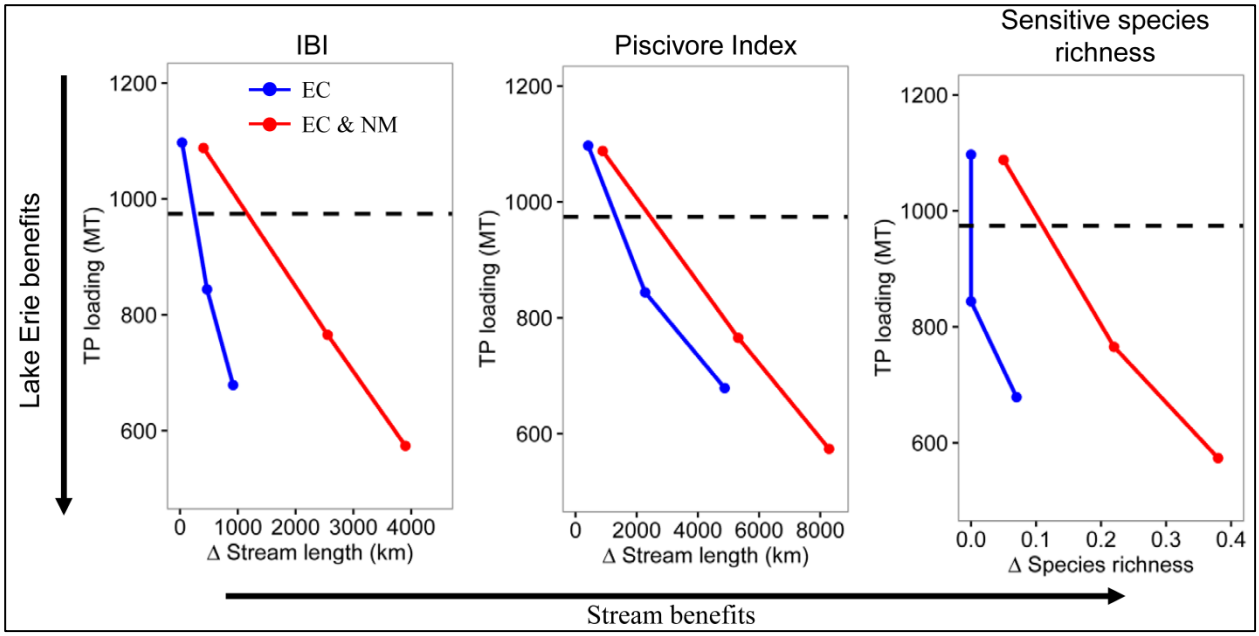


Figure 21. Comparison of potential benefits of additional investment in agricultural conservation practices for reducing spring total phosphorus (TP) loading to Lake Erie and improving stream biological conditions. Stream benefits were assessed as the increase in stream kilometers where water quality was no longer limiting fish community metrics or the change in mean sensitive species richness from baseline conditions. The dashed line indicates the target TP load for Lake Erie. IBI = Index of Biotic Integrity; Piscivore Index = Relative abundance of piscivorous fishes.

Progress Update on Climate Change Scenarios

A better understanding how climate change may alter the effectiveness of CPs is needed to inform adaptive-management in the WLEB. We are currently running simulations using outputs from 20 global climate models and two greenhouse gas emission scenarios (RCP4.5 – this scenario emphasizes climate change mitigation and RCP8.5 –assumes a business as usual scenario with emissions continuing to rise unabated into the future) to understand how projected climate changes alters biological conditions in our conservation scenarios. We only selected the conservation scenarios that included erosion control practices and nutrient management because (1) we wanted to limit the total number of simulations because of the computationally demanding nature of these simulations and (2) these were the most effective scenarios for improving stream biological conditions. To date, we have hind-cast (1986-2005) water quality conditions for all climate change scenarios and global climate models for the Baseline scenario. This was done to evaluate model performance and develop computer programs for post-processing of simulation outputs. As of July 20, 2016, these retrospective simulations are finished and we have developed the necessary computer programs to process data. We are now running simulations for future conditions (2016-2065) on the Ohio State University's Super computer high performance computing facilities. We anticipate completing all necessary simulations and data analysis by early fall (i.e., early October).

RECOMMENDATIONS FOR IMPROVING OUR APPROACH

For this project, we used a state-of-the-art, high-resolution SWAT model and integrated it with other biological, physical, economic models in a way that was both novel and comprehensive. Thus, we have confidence in modeling results. However, as with any modeling study, limitations to our approach existed. Below, we offer several ways to improve our ability to use models to better inform the use of CPs in conservation policy and program development.

- While our modeling was able to estimate how CPs could benefit stream fish community health by improving water quality, our suggested benefits of CPs may be an underestimate because our modeling did not account for likely improvement to in-stream (e.g., in-stream cover, riffle/run/pool quality, stream temperature) and riparian habitat that would be associated with reduced nutrient and sediment runoff. In-stream habitat can be an important driver of stream health (e.g., Miltner 2010, Munn et al. 2010) and studies suggest that CPs not only improve water quality, but can also improve in-stream habitat (e.g., Wang et al. 2002, Wang et al. 2007). Healthy riparian zones also offer numerous benefits to stream ecosystems. For example, increased riparian cover and nutrient sequestration could reduce in-stream nuisance algal growth even at high nutrient concentrations (Munn et al. 2010), while also reducing in-stream water temperature and increasing food subsidies to macroinvertebrates at the base of the food web. Thus, quantifying the benefits of CPs on in-stream habitat, riparian habitat, and water quality, as well as their potential interactions, would provide a more complete picture of the potential for CPs to improve stream ecosystems.
- While CPs are expected to benefit stream health, our modeling approach was unable to predict the time-course of these benefits, once CPs are implemented. While the physical models are technically capable of allowing for such assessments (i.e., daily predictions exist), our biological models are not temporally explicit. In large part, our inability to make time-based predictions stemmed from an incomplete mechanistic understanding of how water quality influences fish communities, as well as a lack of temporally explicit biological monitoring data. Further, the necessary experimental studies conducted at the realistic spatial and temporal scales do not yet exist to allow temporally explicit predictions to be made or tested. Thus, while evidence is accumulating that biological conditions in the WLEB will respond positively to habitat improvements (Miltner 2015), an understanding of how accurate our predictions are, as well as how long recovery would take, is missing. For this reason, research aimed at understanding the response-time of water quality and stream biota to improved habitat conditions would help inform realistic expectations.
- A test case in a smaller subwatershed that uses our approach to identify areas to optimally place CPs, with intensive stream monitoring before and after implementation, would help validate our approach of linking physical and biological models to inform strategic conservation.
- Comparing similar but different modeling efforts would help improve confidence in our results. Using this type of "ensemble" approach has proven effective in quantifying potential CP benefits for managing Lake Erie water quality (e.g., Scavia et al. 2016). We would expect that more studies at the watershed-scale that focus on

stream ecosystem health would similarly benefit watershed conservation in agricultural landscapes.

- Continued support of stream sampling and monitoring of water quality and biological conditions is vital. Without this data, our project would not have been possible. Moreover, this data is essential for monitoring stream health and informing adaptive management within the WLEB. It is essential that this data sampling is intensive and widespread to show improvements in water quality (Betanzo et al. 2015).
- Our conservation scenarios highlight the need for new practices, policies, or technologies that are able to more efficiently reduce agricultural inputs than we simulated. We only considered practices that were in wide-use in the WLEB and practices that we did not include, such as targeted wetland restoration or drainage water management, should be included in future conservation scenarios to quantify their potential benefits.
- The need exists for research that integrates non-social and social components of conservation policy within agricultural landscapes. For example, how likely is it that the practices and policies identified as being effective through modeling efforts will be adopted by WLEB farmers? How do attitudes about stream health improvements influence adoption rates? We strongly encourage consideration of stakeholder group and manager opinions and attitudes, as their consideration has been shown to be integral to successful CP implementation in other agricultural settings (e.g., Prokopy et al. 2008).

CONCLUSIONS AND IMPLICATIONS

Our results demonstrate that agriculture is a major contributor to water quality impairment in WLEB streams and highlight the integral role for CPs to improve agricultural sustainability in this watershed. Returning to critical outstanding questions this project sought to address, we found:

1. *How much additional CP implementation is needed to improve stream water quality and fish community health, both now and under a changing climate?*
    - Our results suggest that widespread implementation is necessary to provide meaningful benefits to stream water quality and biological health within the WLEB. While treating high-needs acres, which comprise ~8% of farm acres (USDA NRCS 2011) with erosion control and nutrient management practices would certainly provide some benefit to WLEB's vast stream network, these represented relatively small improvements compared to what may be possible with more widespread implementation. For example, treatment of high-needs acres with both CP types would result in Piscivore Index still being limited by water quality in 50% of streams, only a 4% improvement from Baseline conditions. By contrast, treatment of farm acres with moderate- and low-needs (~48% of all farm acres; USDA NRCS 2011) with both types of CPs would remove water quality limitation of the Piscivore Index in 72% of the watershed. Our results show that widespread additional implementation of CPs, ideally across all farm acres, is needed to help mitigate water quality impairment from agricultural inputs. These findings are supported by both empirical (Wang et al., 2002) and modeling (Einheuser et al. 2012) studies that have also found widespread implementation of CPs is needed to address the impacts of agricultural NPS pollution on stream ecosystems.
2. *Which types of CPs are most beneficial and cost-effective?*
    - We found that including nutrient management in addition to erosion control CPs provided the largest benefits for stream fishes and generally represented the most cost-effective management strategy. This was likely because, while erosion control CPs were effective at reducing TP and SS concentrations in streams, they were less effective than also including nutrient management at reducing TN concentrations. Our results suggest that addressing TN, TP, and SS is needed to improve stream fish communities. High levels of these stressors were present throughout the watershed, often co-occurring, each with the potential to limit stream fish communities. Addressing these multiple stressors through a variety of practices may therefore be necessary to remove water quality limitation of stream biological conditions throughout the watershed.
3. *How much financial investment is needed to achieve meaningful benefits in stream health?*
    - Because widespread implementation appears needed and multiple CPs appear necessary to reduce both nutrient and sediment inputs, additional financial investments beyond current conservation practice adoption investments appear necessary if we want to achieve a "win-win" for both Lake Erie water quality and WLEB stream biological health. Our cost estimates ranged from an additional $4.5 million annually to implement erosion control CPs on high-need acres to

$263 million annually to implement erosion control CPs and nutrient management on all farm acres. Importantly, even if an additional $263 million were spent annually, water quality still is predicted to limit stream fish communities in a large portion of WLEB streams. While these costs and benefits are *estimates*, and more cost-effective strategies could be developed, they do suggest that substantial financial investments are needed to improve water quality and fish community health throughout the WLEB's vast stream network.

4. *Where in the watershed will the implementation of CPs be most beneficial?*
    - We believe that our approach of coupling physical and biological models can be used to help identify subwatersheds where CPs will likely be most effective at improving stream fish community condition. Our modeling found considerable variation across the watershed in both the benefits of reducing agricultural inputs and the intensity of row-crop agricultural threats. While our results suggest that 100% of acres need additional CP treatment to mitigate water quality impairment from agricultural inputs, identifying areas where benefits are expected to be large and agricultural threats are high may help prioritize subwatersheds for initial conservation investments. This prioritization is important because it will help to maximize in-stream benefits of conservation investments. Moreover, identifying subwatersheds where the implementation of practices may result in rapid results can help maintain and even gain sociopolitical and financial support for continued investment in CPs.
5. *Do potential "win-wins" for Lake Erie water quality management and stream conservation exist?*
    - Our results suggest that targeting and treating only those areas of the WLEB watershed needed to achieve the Lake Erie nutrient reduction goals could leave fish communities limited by agricultural NPS pollution in thousands of miles stream. However, our results also show that if done strategically, by targeting erosion-control and nutrient management practices to areas that maximize benefits to Lake Erie and the tributaries, we can substantially increase the benefits to stream health across the WLEB watershed. Achieving such "win-win" outcomes for Lake Erie and its tributaries is vital given that both support a tremendous amount of biodiversity and provide vital ecosystems services to residents of Indiana, Ohio, and Michigan.

LITERATURE CITED

Alexander, R.B., R.A. Smith, and G.E. Schwarz. 2000. Effect of stream channel size on the delivery of nitrogen to the Gulf of Mexico. Nature 403:758-761.

Angermeier, P.L. and J. R. Karr. 1986. Applying an Index of Biotic Integrity based on stream-fish communities: considerations in sampling and interpretation. North American Journal of Fisheries Management 6:418-429.

Annex 4. 2015. Great Lakes Water Quality Agreement: Annex 4 Objectives and Targets Task Team. Recommended Phosphorus Loading Targets for Lake Erie. Final Report to the Nutrients Annex Subcommittee (70 pp.)

Arnold, J.G. R. Srinivasan, R.S. Muttiah, and P.M. Allen. 1999. Continental scale simulation of the hydrologic balance. Journal of the American Water Resources Association 35:1037-1052.

Arnold, J.G., M.A. Youssef, H. Yen, M.J. White, A. Sheshukov, A.M. Sadeghi, D.N. Moriasi, J.L. Steiner, D.M. Amatya, E.B. Haney, J. Jeong, M. Arabi, and P.H. Gowda. 2015. Hydrological processes and model representation. American Society of Agricultural and Biological Engineering 58:1637-1660.

Barton, K. 2015. MuMIn: multi-model inference. R package version 1.15.1.

Bates, D. M. Maechler, B. Bolker, and S. Walker. 2015. _lme4: linear mixed-effects models using Eigen and S4_. R package version 1.1-8.

Betanzo, E.A., A.F. Choquette, K.H. Reckhow, L. Hayes, E.R. Hagen, D.M Argue, and A.A. Cangelosi. 2015. Water data to answer urgent water policy questions: monitoring design, available data and filling data gaps for determining the effectiveness of agricultural management practices for reducing tributary loads to Lake Erie. Northeast-Midwest Institute Report, 169 pp.

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

Burnham, K.P. and D.R. Anderson. 1998. Model selection and multimodel inference: a practical information-theoretic approach 2nd ed. Springer Science. New York, NY.

Cade, B.S., J.W. Terrell, and R.L. Schroder. 1999. Estimating effects of limiting factors with regression quantiles. Ecology 80:311-323.

Cade, B.S. and B.R. Noon. 2003. A gentle introduction to quantile regression for ecologist. Frontiers in Ecology and the Environment 1:412-420.

Camargo, J.A. and A. Alonso. 2006. Ecological and toxicological effects of inorganic nitrogen pollution in aquatic ecosystems: a global assessment. Environmental International 32:831-849.

Camargo, J.A., A. Alonso, and A. Salamanca. 2005. Nitrate toxicity to aquatic animals: a review with new data for freshwater invertebrates. Chemosphere 58:1255-1267.

Daggupati, P., H. Yen, M.J. White, R. Srinivasan, J.G. Arnold, S.C. Keitzer, and S.P. Sowa. 2015a. Impact of model development decisions on hydrological processes and stream flow. Hydrological Processes 29:5307-5320.

Daggupati, P., N. Pai, S. Ale, K.R. Douglas-Mankin, R. Zeckoski, J. Jeong, P. Parajuli, D. Saraswat, and M. Youssef. 2015b. A recommended calibration and validation strategies for hydrological and water quality models. Transactions of the ASABE 58:1705-1719.

David, M.B., S.J. Del Grosso, X. Hu, E.P. Marshall, G.F. McIsaac, W.J. Parton, C. Tonitto, and M. Youssef. 2009. Modeling denitrification in a tile-drained, corn and soybean agroecosystem of Illinois, USA. Biogeochemistry 93:7-30.

Dodds, W.K., J.R. Jones, and E.B. Welch. 1998. Suggest classification of stream trophic state: distributions of temperate streams types by chlorophyll, total nitrogen, and phosphorus. Water Resources 32:1455-1462.

Dodds, W.K. and R.M. Oakes. 2008. Headwater influences on downstream water quality. Environmental Management 41:367-377.

Dormann, C.F., J. Elith, S. Bacher, C. Buchmann, G. Carl, G. Carré, J.R. García Marquéz, B. Gruber, B. Lafourcade, P. J. Leitão, T. Münkemüller, C. McClean, P.E. Osborne, B. Reineking, B. Schröder, A.K. Skidmore, D. Zurell, and S. Lautenbach. 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36:27-46.

Einheuser, M.D., A.P. Nejadhashemi, S.P. Sowa, L. Wang, Y.A. Hamaamin, and S.A. Woznicki. 2012. Modeling the effects of conservation practices on stream health. Science of the Total Environment 435-436:380-391.

Elith, J., J.R. Leathwick, and T. Hastie. 2008. A working guide to boosted regression trees. Journal of Animal Ecology 77: 802-813.

Evans-White, M.A., W.K. Dodds, D.G. Huggins, and D.S. Baker. 2009. Thresholds in macroinvertebrate biodiversity and stoichiometry across water quality gradients in Central Plains (USA) streams. Journal of the North American Benthological Society 28:855-868.

Fore, J.D., S.P. Sowa, D.L. Galat, G.M. Annis, D.D. Diamond, and C. Rewa. 2014. Riverine threat indices to assess watershed condition and identify primary management capacity of agricultural natural resource management agencies. Environmental Management 53:567-582.

Gassman, P.W., J.R. Williams, X. Wang, A. Saleh, E. Osei, L.M. Hauck, R.C. Izaurralde, and J.D. Flowers. 2009. The agricultural policy/environmental eXtender (APEX) model: an emerging tool for landscape and watershed environmental analyses. Transactions of the ASABE 53:711-740.

Gassman, P.W., M.R. Reyes, C.H. Green, and J.G. Arnold. 2007. The soil and water assessment tool: historical development, applications and future research directions. Transactions of the ASABE 50:1211-1250.

Gebremariam, S.Y., J.F. Martin, C. DeMarchi, N.S. Bosch, R. Confesor, and S.A. Ludsin. 2014. A comprehensive approach to evaluation watershed models for predicting river flow regimes critical to downstream ecosystem services. Environmental Modelling Software 61:121-134.

GLWQA 2016. Great Lakes Water Quality Agreement (GLWQA). https://www.epa.gov/glwqa. Viewed on July 18, 2016.

Hijmans, R.J., S. Phillips, J. Leathwick, and J. Elith. 2015. Dismo: species distribution modeling. R package version 1.0-12.

Hobbs, B.F., S.A. Ludsin, R.L. Knight, P.A. Ryan, J. Biberhofer, and J.J.H. Ciborowski. 2002. Fuzzy cognitive as a tool to define management objectives for complex ecosystems. Ecological Applications 12:1548-1565.

Indiana DEM. 2007. Indian Department of Environmental Management, Surface Water Quality Assessment Program. Fish Community Sampling Program. Fact Sheet B-004-OWQ-A-BS-06-0-R6 (Indianapolis, IN).

Karr, J.R. 1981. Assessment of biotic integrity using fish communities. Fisheries 6:21-27.

Karr, J.R., L. A. Toth, and D.R. Dudley. 1985. Fish communities of Midwestern rivers: a history of degradation. Bioscience 35: 90-95.

Keitzer, S.C., S.A. Ludsin, S.P. Sowa, G. Annis, J.G. Arnold, P. Daggupati, A.M. Froehlich, M.E. Herbert, M.V.V. Johnson, A.M. Sasson, H. Yen, M.J. White, and C.A. Rewa. 2016. Thinking outside of the lake: can controls on nutrient inputs into Lake Erie benefit stream conservation in its watershed? Journal of Great Lakes Research http://dx.doi.org/10.1016/j.jglr.2016.05.012.

Klemes, V. 1986. Operational testing of hydrological simulation models. Hydrological Sciences Journal 31:13-24.

Koeneker, R. 2015. Quantreg: quantile regression. R package version 5.11.

Liu, C., P.M. Berry, T.P. Dawson, and R.G. Pearson. 2005. Selecting thresholds of occurrence in the prediction of species distributions. Ecography 28: 385-393.

Lyons, J., L. Wang, and T.D. Simonson. 1996. Development and validation of an Index of Biotic Integrity for coldwater streams in Wisconsin. North American Journal of Fisheries Management 16:241-256.

Matthaei, C.D., J.J. Piggott, C.R. Townsend. 2010. Multiple stressors in agricultural streams: interactions among sediment addition, nutrient enrichment and water abstraction. Journal of Applied Ecology 47:639-649.

Michalak, A.M., E.J. Anderson, D. Beletsky, S. Boland, N.S. Bosch, T.B. Bridgeman, J.D. Chaffin, K. Cho, R. Confesor, I. Daloğlu, J. V. DePinto, MA. Evans, G.L. Fahnestiel, L. He, J.C. Ho, L. Jenkins, T.H. Johengen, K.C. Kuo, E. LaPorte, X. Liu, M.R. McWilliams, M.R. Moore, D.J. Posselt, R.P. Richards, D. Scavia, A.L. Steiner, E. Verhamme, D.M. Wright, and M.A. Zagorski. 2013. Record-setting algal bloom in Lake Erie caused by agricultural and meteorological trends consistent with expected future conditions. Proceedings of the National Academy of Sciences, USA 110:6448-6452.

Michigan DEQ, 1997. Michigan Department of Environmental Quality, Surface Water Quality Division. GLEAS Procedure #51 Survey Protocols for Wadable Rivers. Chapter 25A *in* Schneider, James C. (ed.) 2000. Manual of fisheries survey methods II: with periodic updates. Michigan Department of Natural Resources, Fisheries Special Report 25, Ann Arbor, MI, pp. 1-40.

Miltner, R.J. 2010. A method and rationale for deriving nutrient criteria for small rivers and streams in Ohio. Environmental Management 45:842-855.

Miltner, R.J. 2015. Measuring the contribution of agricultural conservation practices to observed trends and recent conditions in water quality indicators in Ohio, USA. Journal of Environmental Quality 44:1821-1831.

Miltner, R.J. and E.T. Rankin. 1998. Primary nutrients and the biotic integrity of rivers and streams. Freshwater Biology 40:145-158.

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

Munn, M., J. Frey, and A. Tesoriero. 2010. The influence of nutrients and physical habitat in regulating algal biomass in agricultural streams. Environmental Management 45:603-615.

Murdoch, W., S. Polasky, K.A. Wilson, H.P. Possingham, P. Kareiva, R. Shaw. 2007. Maximizing return on investment in conservation. Biological Conservation 139:375-388.

Naidoo, R., A. Balmford, P.J. Ferraro, S. Polasky, T.H. Ricketts, and M. Rouget. 2006. Integrating economic costs into conservation planning. Trends in Ecology and Evolution 21:681-687.

Nakagawa, S. and H. Schielzeth. 2013. A general and simple method for obtaining R^2 from generalized linear mixed-effects models. Methods in Ecology and Evolution 4:133-142.

Niyogi, D.K., M. Koren, C.J. Arbuckle, and C.R. Townsend. 2007. Stream communities along a catchment land-use gradient: subsidy-stress responses to pastoral development. Environmental Management 39:213-225.

Ohio EPA. 1987. Biological criteria for the protection of aquatic life. Standardized biological field sampling and laboratory methods for assessing fish and macroinvertebrate communities vol. III. Division of Water Quality Monitoring and Assessment, Ohio Environmental Protection Agency (61 pp.).

Ohio EPA. 2010. Ohio Lake Erie Phosphorus Task Force I. Final Report. Ohio Environmental Protection Agency (90 pp.).

Ohio EPA. 2013a. 2013 Updates to Biological Criteria for the Protection of Aquatic Life: Volume II and Volumer II Addendum. Users Manual for Biological Field Assessment of Ohio Surface Waters. Division of Water Quality Monitoring and Assessment, Ohio Environmental Protection Agency (11 pp.).

Ohio EPA. 2013b. Ohio Lake Erie Phosphorus Task Force II. Final Report. Ohio Environmental Protection Agency (96 pp.).

Ohio EPA. 2014. Biological and water quality study of the Maumee River and Auglaize River 2012-2013. Ohio EPA Technical Report EAS/2014-05-03. Division of Surface Water, Ohio Environmental Protection Agency (116 pp.).

Ormerod, S.J., M. Dobson, A.G. Hildrew, and C.R. Townsend. 2010. Multiple stressors in freshwater ecosystems. Freshwater Biology 55:1-4.

Peterson, B.J., W.M. Wollheim, P.J. Mullholland, J.R. Webster, J.L. Meyer, J.L. Tank, E. Martí, W.B. Bowden, H.M. Valett, A.E. Hershey, W.H. McDowell, W.K. Dodds, S.K. Hamilton, S. Gregory, and D.D. Morall. 2001. Control of nitrogen export from watersheds by headwater streams. Science 292:86-90.

Piggott, J.J., D.K. Niyogi, C.R. Townsend, C.D. Matthaei. 2015. Multiple stressors and stream ecosystem functioning: climate warming and agricultural stressors interact to affect processing of organic matter. Journal of Applied Ecology 52:1126-1134.

Prokopy, L.S., K. Floress, D. Klotthor-Weinkauf, and A. Baumgart-Getz. 2008. Determinants of agricultural best management practice adoption: evidence from the literature. Journal of Soil and Water Conservation 63:300-311.

R Core Team. 2015. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Richards, P.R. and D.B. Baker. 2002. Trends in water quality in LEASEQ rivers and streams (northwestern Ohio), 1975-1995. Journal of Environmental Quality 31:90-96.

Richards, R.P., D.B. Baker, J.P. Crumrine, J.W. Kramer, D.E. Ewing, and B.J. Merryfield. 2005. Thirty-year trends in suspended sediment in seven Lake Erie tributaries. Journal of Environmental Quality 37:1894-1908.

Richards, R.P., D.B. Baker, and J.P. Crumrine 2009. Improved water quality in Ohio tributaries to Lake Erie: a consequence of conservation practices. Journal of Soil and Water Conservation 64:200-211.

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

Scavia, D., M. Kalcic, R. Logsdon Muenich, N. Aloysius, J. Arnold, C. Boles, R. Confesor, J. DePinto, M. Gildow, J. Martin, J. Read, T. Redder, D. Robertson, S. Sowa, Y. Wang, M. White, and H. Yen. 2016. Informing Lake Erie agriculture nutrient management via scenario evaluation. Final Report.

Schnepf, M. and C. Cox. 2006. Environmental benefits of conservation on cropland: the status of our knowledge. Soil and Water Conservation Society. Iowa, USA.

Schilling, K.E. 2002. Chemical transport from paired agricultural and restored prairie watersheds. Journal of Environmental Quality 31:1184-1193.

Sowa, S.P., M. Herbert, K.R. Hall, L. Cole, S. Mysorekar, M. Fales, T. Bowe, G. Annis, A. Nejadhashemi, and L. Wang. 2011. Assessing the ability of conservation practices to restore biological communities in agricultural watersheds. Final Report submitted to NRCS Conservation Effects Assessment Project. 69 pp.

Taylor, J.M., R.S. King, A.A. Pease, and K.O. Winemiller. 2014. Nonlinear response of stream ecosystem structure to low-level phosphorus enrichment. Freshwater Biology 59:969-984.

Townsend, C.R., S.S. Uhlmann, and C.D. Matthaei. 2008. Individual and combined responses of stream ecosystems to multiple stressors. Journal of Applied Ecology 45:1810-1819.

Trautman, M.B. 1981. The fishes of Ohio. Ohio State University Press, Columbus, OH.

US EPA. 2016. National rivers and streams assessment 2008-2009: a collaborative survey (EPA/841/R-I6/007). Office of Water and Office of Research and Development. Washington, D.C.

USDA NRCS. 2011. Conservation effects assessment project. Assessment of the effects of conservation practices on cultivated cropland in the Great Lakes Region. United States Department of Agriculture Natural Resources Conservation Service (172 pp.).

USDA NRCS. 2016. Effects of conservation practice adoption on cultivated cropland acres in Western Lake Erie Basin, 2003-2006 and 2012. United States Department of Agriculture Natural Resources Conservation Service (120 pp.).

Vaz, S., C.S. Martin, P.D. Eastwood, B. Ernande, G.J. Meaden, and F. Coppin. 2008. Modelling species distributions using regression quantiles. Journal of Applied Ecology 45:204-217.

Wang, L., D.M. Robertson, and P.J. Garrison. 2007. Linkages between nutrients and assemblages of macroinvertebrates and fish in wadeable streams: implications to nutrient criteria development. Environmental Management 39:194-212.

Wang, L., J. Lyons, and P. Kanehl. 2002. Effects of watershed best management practices on habitat and fish in Wisconsin streams. Journal of the American Water Resources Association 38:663-680.

Ward, J.V. 1989. The four-dimensional nature of lotic ecosystems. Journal of the North American Benthological Society 8:2-8.

Warton, D.I. and F.K. Hui. 2011. The arcsine is asinine: the analysis of proportions in ecology. Ecology 92:3-10.

Waters, T.F. 1995. Sediment in Streams: Sources, Biological Effects, and Control. AFS Monographs, Bethesda, MD.

Wood, P.J. and P.D. Armitage. 1997. Biological effects of fine sediment in the lotic environment. Environmental Management 21:203-217.

Wood, S. 2006. Generalized additive models: an introduction with R. Chapman & Hall, New York, NY.

Woods, A.J., J.M. Omernik, C.S. Brockman, T.D. Gerber, W.D. Hosteter, and S.H. Azevedo. 1998. Ecoregions of Indiana and Ohio (2 sided color poster with map, descriptive text, summary tables, and photographs). USGS, Reston, VA. Scale 1:500,000.

Yen, H., R.T. Bailey, M. Arabi, M. Ahmadi, M.J. White, and J.G. Arnold. 2014a. The role of interior watershed processes in improving parameter estimation and performance of watershed models. Journal of Environmental Quality 43:1601-1613.

Yen, H., X. Wang, D.G. Fontane, R.D. Harmel, and M. Arabi. 2014b. A framework for propagation of uncertainty contributed by parameterization, input data, model structure, and calibration/validation in watershed modeling. Environmental Modelling and Software 54:211-221.

Yen, H., M.J. White, J.G. Arnold, S.C. Keitzer, M.V.V. Johnson, J.D. Atwood, P. Daggupati, M.E. Herbert, S.P. Sowa, S.A. Ludsin, D.M. Robertson, R. Srinivasan, and C.A. Rewa. 2016. Western Lake Erie Basin: Soft-data-constrained, NHDPlus resolution watershed modeling and exploration of applicable conservation scenarios. Science of the Total Environment. *In Press*. DOI: 10.1016/j.scitotenv.2016.06.202.

APPENDIX A: DAGGUPATI ET AL. 2015. HYDROLOGICAL PROCESSES 29:5307-5320.

HYDROLOGICAL PROCESSES
Hydrol. Process. **29**, 5307–5320 (2015)
Published online 7 July 2015 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/hyp.10536

Impact of model development, calibration and validation decisions on hydrological simulations in West Lake Erie Basin†

Prasad Daggupati,[1]* Haw Yen,[2,3] Michael J. White,[3] Raghavan Srinivasan,[1] Jeffrey G. Arnold,[3] Conor S. Keitzer[4] and Scott P. Sowa[5]

[1] *Spatial Science Laboratory, Texas A&M University, Suite 221, 1500 Research Plaza, College Station, TX 77843, USA*
[2] *Blackland Research and Extension Center, Texas A&M AgriLife Research, 720 East Blackland Road, Temple, TX 76504, USA*
[3] *Grassland, Soil and Water Research Laboratory, USDA-ARS, 808 East Blackland Road, Temple, Texas 76502, USA*
[4] *Aquatic Ecology Lab, Department of Evolution, Ecology and Organismal Biology, The Ohio State University, 1314 Kinnear Rd, Columbus, OH 43212, USA*
[5] *The Nature Conservancy, Michigan Field Office, 101 East Grand River Ave, Lansing, MI 48906, USA*

Abstract:

Watershed simulation models are used extensively to investigate hydrologic processes, landuse and climate change impacts, pollutant load assessments and best management practices (BMPs). Developing, calibrating and validating these models require a number of critical decisions that will influence the ability of the model to represent real world conditions. Understanding how these decisions influence model performance is crucial, especially when making science-based policy decisions. This study used the Soil and Water Assessment Tool (SWAT) model in West Lake Erie Basin (WLEB) to examine the influence of several of these decisions on hydrological processes and streamflow simulations. Specifically, this study addressed the following objectives (1) demonstrate the importance of considering intra-watershed processes during model development, (2) compare and evaluated spatial calibration *versus* calibration at outlet and (3) evaluate parameter transfers across temporal and spatial scales. A coarser resolution (HUC-12) model and a finer resolution model (NHDPlus model) were used to support the objectives. Results showed that knowledge of watershed characteristics and intra-watershed processes are critical to produced accurate and realistic hydrologic simulations. The spatial calibration strategy produced better results compared to outlet calibration strategy and provided more confidence. Transferring parameter values across spatial scales (i.e. from coarser resolution model to finer resolution model) needs additional fine tuning to produce realistic results. Transferring parameters across temporal scales (i.e. from monthly to yearly and daily time-steps) performed well with a similar spatial resolution model. Furthermore, this study shows that relying solely on quantitative statistics without considering additional information can produce good but unrealistic simulations. Copyright © 2015 John Wiley & Sons, Ltd.

KEY WORDS intra-watershed processes; spatial calibration; spatial and temporal scale parameter transfer; SWAT; NHD plus model

Received 6 October 2014; Accepted 2 May 2015

INTRODUCTION

Watershed simulation models are increasingly used to investigate policy relevant issues related to hydrologic topics, landuse changes, climate change impacts, pollutant load assessments and best management practices (BMPs) (Stone *et al*., 2001; Veith *et al*., 2003; Kannan *et al*., 2005; Benham *et al*., 2006; Tuppad *et al*., 2010; Daggupati *et al*., 2011; Knisel and Douglas-Mankin, 2012; Jha and Gassman, 2014). Model practitioners have the responsibility to make a number of critical decisions in model development, calibration and validation to ensure that the model accurately simulates real world conditions (Eckhardt *et al*., 2005; Zhang *et al*., 2009; Arnold *et al*., 2012a; Arnold *et al*., 2015).

A key step in model development (building a model for a watershed or study region) is that model practitioners should have a good understanding of the watershed characteristics and processes being simulated and should represent them appropriately in the model. Incorporating knowledge about watershed characteristics and processes from literature sources and expert opinion can ensure that models are spatially capturing the hydrological processes and water balance within reasonable limits and realistically simulating real world conditions (Seibert and McDonnell, 2002; Arnold *et al*., 2015). For example, Yen *et al*. (2014a,b) demonstrated that considering intra-watershed characteristics and processes produced accurate spatial and temporal results

*Correspondence to: Prasad Daggupati, Spatial Science Laboratory, Texas A&M University, Suite 221, 1500 Research Plaza, College Station, TX 77843, USA.
E-mail: pdaggupati@tamu.edu
†USDA is an equal opportunity employer and provider.

Copyright © 2015 John Wiley & Sons, Ltd.

that enable the model to provide the right answer for the right reasons.

In addition to considering watershed characteristics and processes during model development, decisions about how models are calibrated further influence the ability of the model to produce relatively more realistic results (Eckhardt *et al.*, 2005; White and Chaubey, 2005; Moriasi *et al.*, 2007; Zhang *et al.*, 2009; Arnold *et al.*, 2012a). Calibration of a watershed simulation model at a single site (generally the outlet of a watershed) remains a widely used calibration strategy (Cao *et al.*, 2006; Wang *et al.*, 2012). This approach is best used in small watersheds with fairly uniform characteristics (e.g. soil, slope, vegetation, meteorology). The use of a single site to calibrate large watersheds may result in calibrated parameters which, (1) represent an average of characteristics over the entire watershed, or (2) present a combination of over- or underestimated values which result in poor intra-watershed spatial accuracy. This may be undesirable for simulations of larger watersheds that are more spatially heterogeneous. In these cases, spatial calibration with additional sites is recommended because larger watersheds may contain varied, complex physical characteristics (Qi and Grunwald, 2005; Piniewski and Okruszko, 2011; Cho *et al.*, 2013; Daggupati *et al.*, 2015). This process better accounts for spatial biophysicochemical variations and reduces the problem of non-unique solutions because fewer parameter sets would satisfy calibration criteria at all sites.

After calibration, the model has to be validated to demonstrate that a given site-specific calibrated model can make sufficiently accurate simulations in a new modeling situation. Several studies have focused on transferring parameters temporally (from one time period (e.g. 1990 to 1999) to another (e.g. 2000 to 2010) (Bingner *et al.*, 1997; Van Liew and Garbrecht, 2003; Abbaspour *et al.*, 2007; Chaubey *et al.*, 2010; Sheshukov *et al.*, 2011; Douglas-Mankin *et al.*, 2013; Seo *et al.*, 2014) and spatially (from gauged to ungauged watershed) (Vandewiele and Elias, 1995; Xu, 1999; Santhi *et al.*, 2001; Merz and Blöschl, 2004; Santhi *et al.*, 2008; Parajuli *et al.*, 2009; He *et al.*, 2011; Kumar *et al.*, 2013a,b) to validate the performance of the model. However, little is known about model performance when calibrated parameters are transferred across temporal and spatial scales. For example, how would the model perform if the parameters are transferred across a temporal scale, i.e. from one time-step (e.g. monthly) to another time-step (e.g. daily)? Or how would the model perform if parameters are transferred across spatial scales such as a coarser resolution model to a finer resolution model within the same watershed? Transferring original parameters across spatial or temporal scales might be one way to save on time without sacrificing model performance. An attempt was made by Troy *et al.* (2008) to evaluate the effects of parameter transfer across spatial and temporal scales using a global land surface model known as Variable Infiltration Capacity (VIC). VIC was used to model the entire continental United States, and the results suggested that the transfer of parameters across temporal scales performed better than transfer across spatial scales. Troy *et al.* (2008) also emphasized the need for more studies using hydrologic models to determine if transferring parameters across scales is a viable validation option for producing realistic results.

This paper focuses on understanding how decisions in model development, calibration and validation influence model realism and ensure that watershed simulations provide the information needed to support science-based policy decisions. This research was motivated by the need to provide a realistic hydrologic simulation for a large watershed at a finer spatial resolution to inform policy decisions in the West Lake Erie Basin (WLEB). The major objectives of this study were to utilize the Soil and Water Assessment Tool (SWAT) model in the WLEB to (1) demonstrate the importance of considering intra-watershed processes during model development (2) Compare and evaluate spatial calibration versus calibration at an outlet and (3) evaluate parameter transfers across temporal and spatial scales. In this study, we developed a SWAT model at the 12-digit Hydrologic Unit Code (HUC-12) resolution and another at the National Hydrography Dataset (NHDPlus) resolution. Objective 1 and 2 were examined using the HUC-12 resolution model, and objective 3 was examined using both the HUC-12 and the NHDPlus resolution model.

STUDY AREA

The WLEB watershed drains 28 330 km^2 encompassing the Maumee River, Sandusky River to the south and the Raisin River in the north (Figure 1). There are over 23 000 km of natural and man-made streams in the watershed, which covers portions of Indiana (17%), Michigan (17%) and Ohio (76%). Other major rivers include the Portage, Sandusky, Blanchard, Auglaize, St. Marys, St. Joseph and Tiffin. There is little topography, with elevation ranging from 246 m to 387 m and an average slope of 2%. Average annual precipitation ranges from 838 to 940 mm.

Prior to European settlement, the watershed primarily consisted of Beech, Maple, Ash and Elm forests (Sears, 1941). The Great Black Swamp, a large wetland (>3800 km^2) located centrally in the watershed, was a major landscape feature (Kaatz, 1955). Widespread forest clearing and wetland draining began in the mid-19th century (Kaatz, 1955). The watershed is now predominantly agricultural, with more than 70% of the land in

Copyright © 2015 John Wiley & Sons, Ltd.

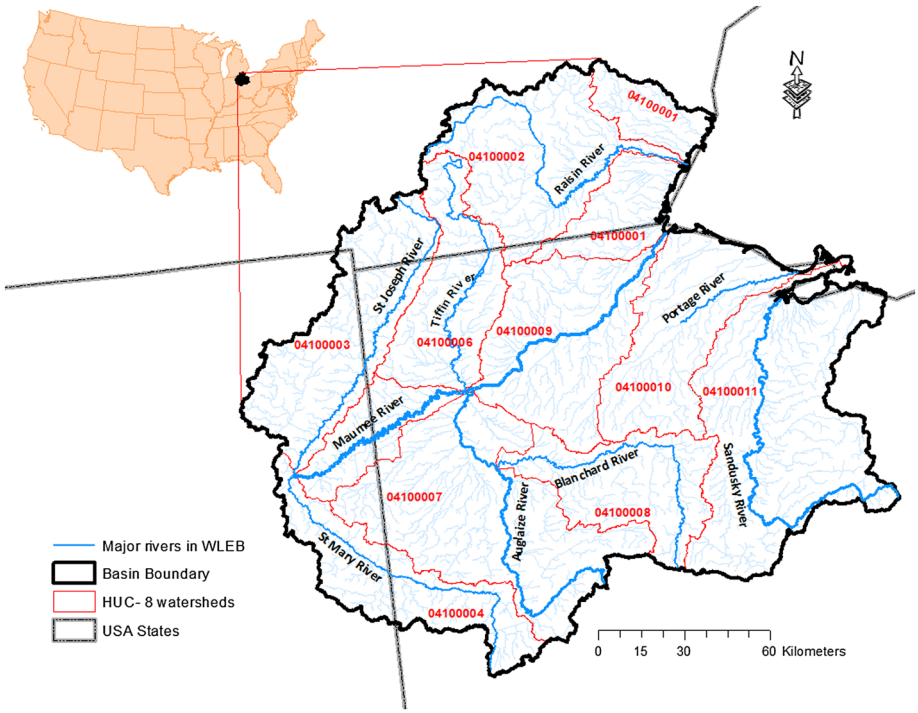


Figure 1. The study area—West Lake Erie Basin (WLEB) with major rivers and HUC-8 watershed boundaries. The insert shows continental United States with the study area shaded

cultivated cropland, the majority of which is in corn–soybean crop rotations. Tile drainage is used extensively throughout the watershed. The next most dominant land uses, forested and urban land use, each make up about 12% of the watershed.

The widespread conversion of native vegetation to agriculture and associated drainage practices (e.g. stream channelization) have degraded freshwater habitat quality and negatively affected freshwater biodiversity (Trautman, 1939; Trautman and Gartman, 1974; Karr *et al*., 1985). Additionally, the Maumee River appears to be a major contributor to eutrophication and the recent increase in harmful algal blooms in Lake Erie (Kane *et al*., 2014). These freshwater conservation and human health concerns require a finer resolution hydrologic model that realistically simulates hydrologic processes to allow policy makers to make informed decisions related to improve conditions in the WLEB.

DATA INPUTS AND MODEL SETUP

The latest version of SWAT, ArcSWAT 2012 (rev 593) for ArcGIS10.1 Geographic Information System interface, was used to set up the SWAT model. The SWAT model is a continuation of nearly 30 years of modeling efforts by the USDA Agricultural Research Service (ARS) and is widely used, watershed-scale, process-based model (Gassman *et al*., 2007; Douglas-Mankin *et al*., 2010; Arnold *et al*., 2012a). The model is supported by online documentation (Neitsch *et al*., 2011; Arnold *et al*., 2012b) which reviews all processes simulated with the model. The ArcSWAT interface allows importing pre-defined watershed boundaries and streams along with automatic delineation of streams and subwatersheds (Luo *et al*., 2011). This function was employed to develop two models, a HUC-12 model based on a predefined HUC-12 Watershed Boundary Dataset (WDB) and a more detailed NHDPlus model using the National Hydrography Dataset (NHD) and NHD-plus stream network. The HUC-12 WBD (1:240 000 scale) was downloaded from http://datagateway.nrcs.usda.gov and is a coordinated effort between the United States Department of Agriculture-Natural Resources Conservation Service (USDA-NRCS), the United States Geological Survey (USGS) and the Environmental Protection Agency (EPA). The NHDPlus data consists of NHD Plus (Version 2) streams and catchments at scale of 1:100 000 and was downloaded from http://www.horizon-systems.com/NHDPlus/index.php. The NHDPlus framework is a coordinated effort by the EPA Office of Water and the USGS. A 30-m Digital Elevation Model (DEM) was used to define the topographical characteristics for each model. A total of 391 and 11 128 subbasins, respectively, were characterized for HUC-12 and NHD Plus setup in the Western Lake Erie Basin (Figure 2). The average size of the subwatersheds in the HUC-12 model was 72 km^2 (range, 25 to 191) while the average watershed size in the

Copyright © 2015 John Wiley & Sons, Ltd.

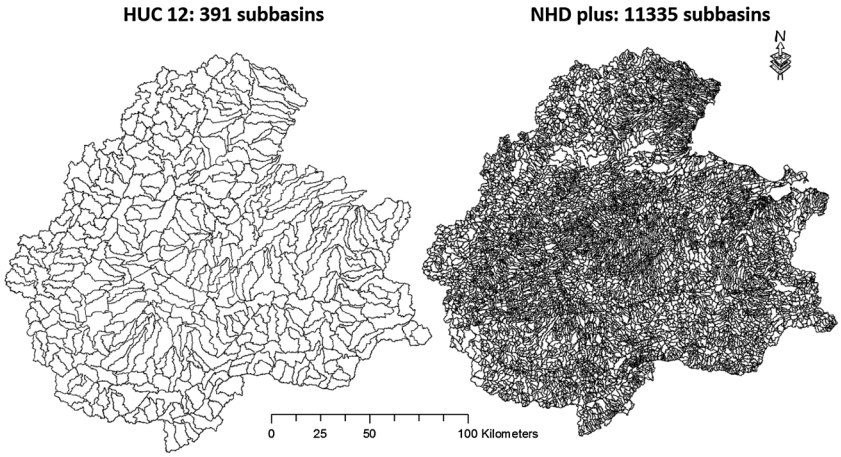


Figure 2. HUC-12 (coarse resolution) and NHDPlus (finer resolution) subwatersheds in WLEB

NHDPlus model was 2.6 km^2 (range, 0.001 to 80). Developing SWAT model at NHDPlus resolution is first of its kind, and no studies were reported in literature that used SWAT at such finer resolution.

Land use was defined using 2010 and 2011 Crop Data Layers (CDLs). The data was processed using techniques recommended by Srinivasan *et al.* (2010) to prepare a single 30-m resolution landuse/landcover layer which includes major crop rotations. Soils were derived from STATSGO (USDA-NRCS, 1995) at a scale of 1: 250 000. All soil properties needed for the SWAT model were extracted from the national STATSGO layer and processed with the ArcSWAT interface.

Land use, soils and slope (derived from DEM) were intersected within each subbasin by ArcSWAT to create unique Hydrologic Response Units (HRUs). Three slope classes (0%–2%, 2%–5% and >5%) were used with landuse, soil and slope (by class) thresholds of 50/50/50 ha. All agricultural crops were exempt from landuse thresholds such that all agricultural crops were included as HRUs. A total of 13 156 and 34 807 HRUs were derived in the WLEB using the HUC-12 and the NHDPlus models.

Both models included daily precipitation and temperature data from 1960 to 2010. This data was derived from the National Oceanic and Atmospheric Administration (NOAA) Cooperative Observer network and Weather-Bureau-Army-Navy stations. Missing data at each station was supplied using an inverse distance weighted interpolation algorithm and observations from the nearest five stations.

Tile drain systems are designed to remove excess field water and lower water tables to reduce crop stresses and allow timely field tillage and planting. However, no clear record of tile locations was available in this basin. It was therefore assumed that tile drains occur in agricultural areas that are located in poorly drained soils and have a slope less than 1%. Poorly drained soils were identified within the basin by processing SSURGO soil using soil data viewer 6.0 program (http://www.nrcs.usda.gov/wps/portal/nrcs/detailfull/soils/home/?cid=nrcs142p2_053620) in ArcGIS.

SWAT management operations (i.e. planting, tillage, harvest and fertilizer application) were assembled from a variety of sources. Operation scheduling was derived from management templates developed by the NRCS for the RUSLE2 model (Foster, 2005). Cropland tillage was derived from Baker (2011). SWAT plant growth-related parameters were developed from local weather statistics. Cropland fertilization was derived from NASS reported county average crop yields. Data was processed and combined into SWAT format management files using software written specifically for this purpose.

Measured streamflow from 12 gauge stations (Figure 3, Table I) was collected from 1 January 1990 to 31 December 2006. The data was used during calibration and validation to facilitate spatial calibration and validation assessments.

METHODS

Intra-watershed processes

The majority of the WLEB is comprised of agricultural land of which more than 85% have tile drainage systems implemented to facilitate artificial drainage and to improve crop yields and field trafficability. Tile drainage is an important and major intra-watershed process in the basin. Representing tile drain in the SWAT model was necessary to accurately capture the spatial hydrological processes and water balance within the watershed. This study evaluated the effects of tile drain to demonstrate the importance of considering intra-watershed processes within the watershed during model development and simulations. The default SWAT model (without

Copyright © 2015 John Wiley & Sons, Ltd.

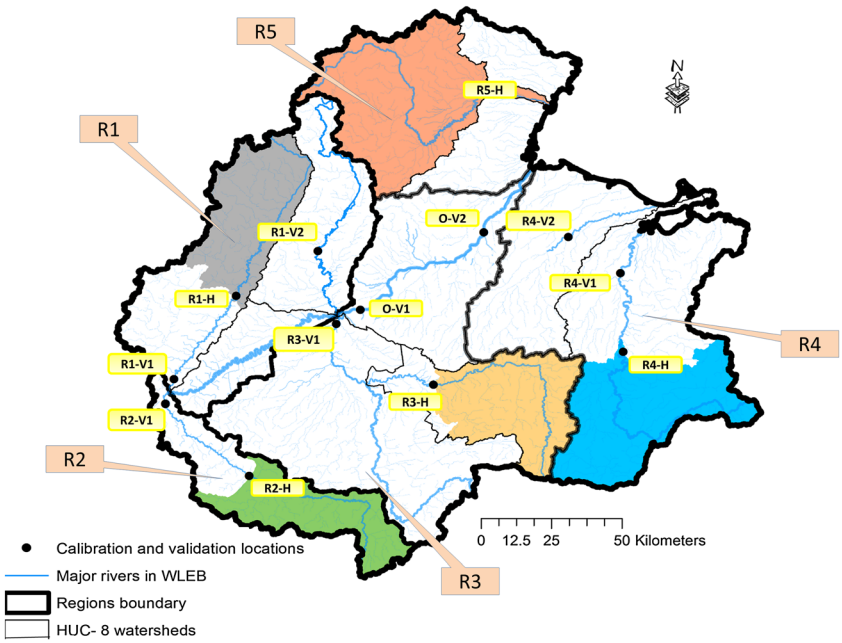


Figure 3. Five regions (separated with thick black border), five head watersheds (represented in shaded colors), and calibration and validation locations in WLEB

calibration) was used to evaluate the effectiveness of tile drains in capturing the overall hydrologic water balance in the watershed. Tile drain parameters including depth to drain (DDRAIN), time to drain (TDRAIN), drain lag time (GDRAIN) and depth to impervious layer (D_IMP) were changed to 1500 (mm), 48 (h), 24 (h) and 1200 (mm), respectively. These tile drain parameters values were based on expert opinions of watershed specialists working in the watershed. The computation of the daily CN value as a function of plant evapotranspiration (ICN = 1) was used because the default soil moisture method (ICN = 0) was predicting too much runoff in shallow soils (Yen *et al.*, 2014c). ICN = 1 (plant-based ET) and ICN COEF = 0.5 along with other tile drain parameters reduced surface runoff and transferred that water as tile flow and thereby simulated tile drains reasonably well in the watershed. The SWAT model was simulated from 1990 to 1999 and a 3-year warm-up period (1987 to 1989) was used prior to the model simulation period as recommended by Daggupati *et al.* (2015). Average annual hydrologic components (surface runoff, ground flow, lateral flow and tile flow) of the water balance along with quantitative statistics and graphical comparisons (discussed in the next section) at R4-H gauge station (Figure 3) were used to assess the performance of simulations with and without tile drains.

Spatial and outlet calibration

The default model after the inclusion of tile drain information (a major intra-watershed process) needed to be calibrated to increase the accuracy of model predictions. A regular calibration procedure where the model is calibrated using observed and simulated data at the outlet may not work well in the WLEB because of the large size of this watershed and potential spatial variability within the basin. In order to capture this spatial variability, spatial calibration is needed. We used a proxy-basin spatial calibration strategy originally proposed by Klemes (1986) and summarized by Daggupati *et al.* (2015). This strategy involves calibration of a model in a gauged watershed and transferring calibrated parameters to nearby or adjacent watersheds within the same eco-region. Further, a spatial validation is performed at various locations to evaluate the performance of model. The logic behind this method is that in a similar eco-region, the climate and watershed conditions vary smoothly over space and the parameters in the region are expected to be similar (Jin *et al.*, 2009).

The WLEB was divided into five different regions (R1, R2, R3, R4 and R5) (Figure 3, Table I). R1, R2 and R3 drain into Maumee River, R4 drains into Sandusky and Cedar-Portage rivers while R5 drains into Raisin and other adjacent tributary rivers. During separation of the regions, landuse, soil, slope and precipitation were used to visualize the spatial variability within the basin. HUC-8 watershed boundaries within the basin were used as guidelines to separate regions. A head watershed (shaded areas in Figure 3) was selected in each region and was calibrated and validated using a temporal split-sample approach in which one period (1990 to 1999) was used for calibration and another period (2000 to 2006) was used for temporal validation. After satisfactory calibration

Copyright © 2015 John Wiley & Sons, Ltd.

Table I. Calibration–validation locations, their station name, stationID, source of data collected at each location and corresponding draining HUC-8 watershed with its name. Locations used for calibration and spatial validation are also shown. In the table, R represents region, H represents head watershed location, V represents validation location and O represents outlets

Calibration /validation locations	Station name	StationID	Source	HUC-8 (name)	Region	Calibration	Validation
R1-H	St. Joseph River near Newville, IN	4178000	United States Geologic Survey	04100003 (St. Joseph)	R1	√	
R1-V1	Tiffin River at Stryker, OH	4185000	United States Geologic Survey	04100006 (Tiffin)	R1		√
R1-V2	St. Joseph River	LEJ100-0003	Indiana Department of Environmental Management	04100003 (St. Joseph)	R1		√
R2-H	St. Marys River	LES040-0007	Indiana Department of Environmental Management	04100004 (St. Marys)	R2	√	
R2-V2	St. Marys River	LES060-0004	Indiana Department of Environmental Management	04100004 (St. Marys)	R2		√
O-V	Maumee River at Waterville, OH	4193500	Heidelberg College River Studies	04100009 (Lower Maumee)	R3		√
R3-H	Blanchard River upstream of Ottawa, OH at County Road 8	500100	Ohio Environmental Protection Agency	04100008 (Blanchard)	R3	√	
R3-V1	Auglaize River upstream of Defiance, OH at Harding Road	500290	Ohio Environmental Protection Agency	04100007 (Auglaize)	R3		√
O-V	Maumee River upstream of Independence Dam	P09W19	Ohio Environmental Protection Agency	04100009 (Lower Maumee)	R3		√
R4-V2	Portage River in Woodville, OH at railroad bridge	4195600	Ohio Environmental Protection Agency	04100010 (Cedar-Portage)	R4		√
R4-H	Sandusky River near Mexico, OH	4197000	Ohio Environmental Protection Agency	04100011 (Sandusky)	R4	√	
R4-V1	Sandusky River near Fremont, OH	4198000	Heidelberg College River Studies	04100011 (Sandusky)	R4		√
R5-H	River Raisin near Monroe, MI	4176500	Heidelberg College River Studies	04100002 (Raisin)	R5	√	

Copyright © 2015 John Wiley & Sons, Ltd.

results in head watersheds in five regions (based on quantitative statistics and graphical comparisons as discussed later), the parameters were transferred to other watersheds in the region. Spatial validation was performed at various locations (Figure 3, Table I) within the transferred region and also at two outlet locations along the Maumee River. During the process of validating the model in regions, the temporal split-sample approach, as discussed previously, was used to complete a comprehensive evaluation and thereby accomplishing spatial calibration.

During the calibration of head watersheds, SWAT was manually calibrated to make sure that the hydrology, overall water balance and general seasonal patterns within the watershed were captured. Additional automated calibration was done using the Sequential Uncertainty Fitting version-2 (SUFI-2) routine in SWAT-CUP program. A monthly time-step was used during calibration. Quantitative statistics and criteria recommended by Moriasi *et al*. (2007) were used to evaluate the simulation performance. The quantitative statistics applied in this study were Nash–Sutcliffe simulation efficiency (NSE) and percentage bias (PBIAS). The model performance for monthly and daily streamflow can be categorized into four classes according to the threshold NSE, and PBIAS values: very good ($0.75 < NSE \leq 1.00$, $PBIAS < \pm 10$); good ($0.65 < NSE \leq 0.75$, $\pm 10 \leq PBIAS < \pm 15$); satisfactory ($0.50 < NSE \leq 0.65$, $\pm 15 \leq PBIAS < \pm 25$); and unsatisfactory ($NSE \leq 0.50$, $PBIAS \geq \pm 25$). Graphical comparisons of time-variable plots of observed and simulated flow provide important insights into model representation of hydrographs, baseflow recession and other pertinent factors often overlooked by quantitative comparisons. In this study, visual comparisons of hydrographs between observed and simulated were evaluated, and the simulation was considered satisfactory only when the shapes (peaks and base flow) of observed and predicted hydrographs were similar.

Outlet calibration strategy was tested by performing an auto calibration using SUFI-2 routine in SWAT-CUP program at a monthly time-step utilizing streamflow at the outlet (O-V2 location, Figure 3). After calibration, the quantitative statistics and graphical comparisons at various locations within the watershed were evaluated. The model performance using spatial and outlet calibration strategies were compared and analysed. In this study, graphical representation of quantitative statistics (only NSE) for spatial and outlet calibration strategies at various locations within WLEB is shown in Figure 4 to have a better view of results spatially and would ensure a more comprehensive evaluation of model performance based on recommendation by Saraswat *et al*. (2015).

Transfer of parameters across scale

Transferring parameters across scales may be desirable as a part of validation option to adapt models to address new issues beyond their original intent. Developing and calibrating a new model to address these new issues may be time consuming. This study investigated the model performance after transferring parameters across spatial and temporal scales to determine if this is a viable option to address novel issues beyond the scope of the original models intent. We transferred parameters from a spatially calibrated and validated HUC-12 model (coarser resolution model) at a monthly time-step to NHDPlus model (finer resolution model) to evaluate the impacts of transferring parameters across spatial scale. Next, in the HUC-12 model, we transferred parameters to daily and yearly time-step to evaluate the temporal scale effects where the model was calibrated on monthly time-step. Also, the temporal scale effects in the NHDPlus model were also evaluated on daily and yearly time-step after the transfer of parameters from a monthly time-step calibrated HUC-12 model. In both the cases, quantitative statistics and graphical criteria (temporal time series plots) at the outlet (O-V2) and one another location (R4-H)

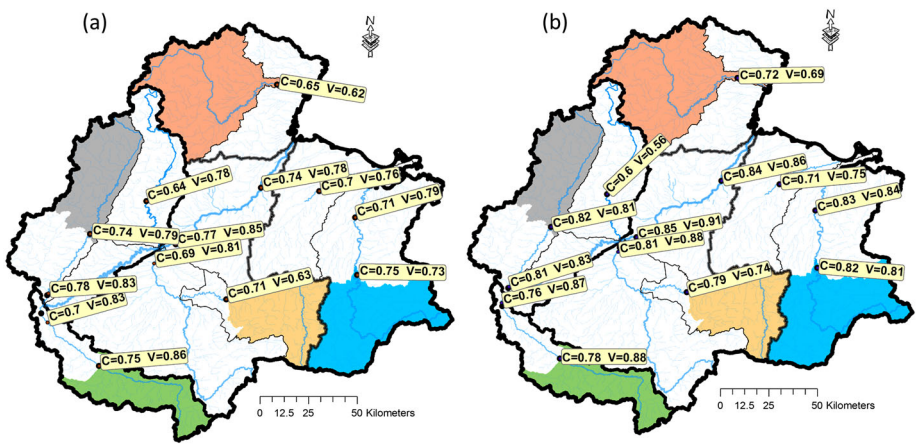


Figure 4. NSE values at various locations in the basin using a) spatial calibration strategy and b) outlet calibration strategy

Copyright © 2015 John Wiley & Sons, Ltd.

Table II. Average annual hydrologic components for S1, S2 and S3 model simulations for the time period of 1990 to 1999 in WLEB

Average annual hydrologic components	S1 (mm)	S2 (mm)	S3 (mm)
% surface runoff (surface runoff/water yield)	71%	11%	28%
% tile flow tile flow/water yield	0%	60%	53%
% ground water (ground water/water yield)	28%	28%	15%
% later flow (lateral/water yield)	1%	1%	3%

Table III. Quantitative statistics for S1, S2 and S3 at R4-H location

Intra-watershed processes	R2	NS	Median simulated	Median observed	PBIAS
Without tile (S1)	0.61	0.60	20.22	14.43	−6.95
After tile (S2)	0.69	0.27	31.71	14.43	−57.29
After calibration/ with tile (S3)	0.83	0.82	13.43	14.21	−4.60

were used to evaluate the impacts of parameter transfer across scale.

RESULTS AND DISCUSSION

Impacts of considering intra-watershed processes

Scenario 1 (S1) is used to denote SWAT model simulation without tile drain and S2 for SWAT model with tile drain. Scenario 3 (S3) which is a SWAT model simulation after calibration is also used for comparison and discussion purposes. However, more discussion on calibration is given later. Average annual hydrologic components of the water balance in WLEB are shown in Table II for S1, S2 and S3 model simulations. Quantitative statistics are presented in Table III for all three scenarios using R4-H gauge location (Sandusky River near Mexico OH) which is in Sandusky watershed (HUC8, 4100011) and is heavily dominated by agricultural land (>83%) and is mostly implemented with tile drains. Graphical comparisons are shown in Figure 5. The surface runoff contribution without tile drainage (S1) was 71% (ratio of surface runoff to total water yield),

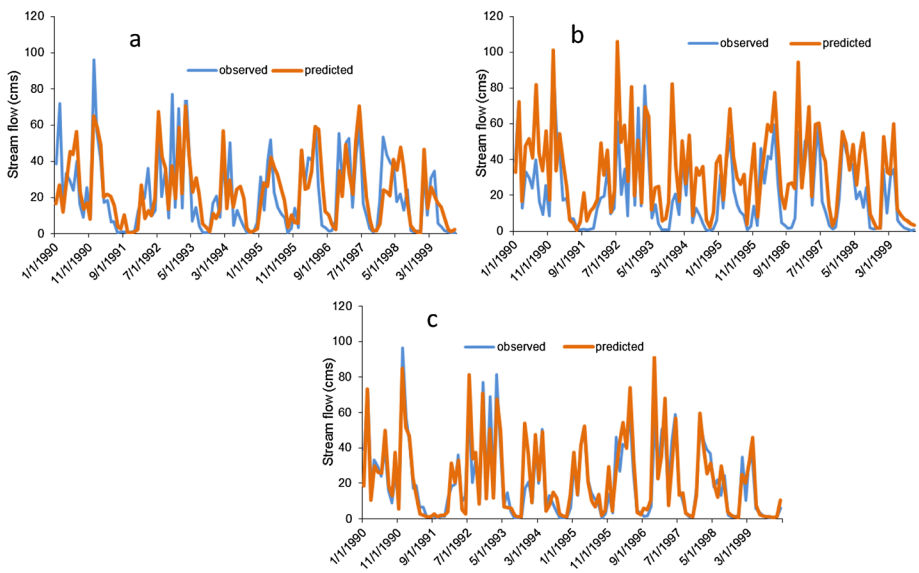


Figure 5. a) Monthly time series comparison in S1, S2 and S3 model simulations for the time period of 1990 to 1999 at R4-H gauge location

Table IV. Quantitative statistics at head watershed locations for Spatial and Outlet calibration strategies

Calibration strategy	Time period	R1-H		R2-H		R3-H		R4-H		R5-H	
		NSE	PBIAS	NSE	PBIAS	NSE	PBIAS	NSE	PBIAS	NSE	PBIAS
Spatial calibration	1990–1999	0.82	11.89	0.78	0.09	0.79	4.81	0.82	−4.72	0.72	0.97
	2000–2006	0.81	5.21	0.88	−3.90	0.74	26.19	0.81	−5.81	0.69	−8.82
Outlet calibration	1990–1999	0.74	−3.98	0.75	4.73	0.71	13.72	0.75	−10.72	0.65	−9.15
	2000–2006	0.79	−14.88	0.86	0.21	0.63	25.94	0.73	−15.24	0.62	−19.41

Copyright © 2015 John Wiley & Sons, Ltd.

while tile flow contribution was 0% (ratio of tile flow to total water yield) (Table II). Monthly quantitative statistics (Table III) show that the model predictions were satisfactory and reliable when tile flow was not included ($R^2 = 0.61$, NSE = 0.60, PBIAS = −6.95). However, having tile flow contribute 0% to total water yield is unrealistic for this watershed. Communication with the experts that work in the watershed suggested that the tile drain contribution is generally around 50% while the surface runoff contribution is around 30%. Thus, despite being unrealistic, the scenario without tile drainage (S1) produced satisfactory performance statistics (Table III). Graphical comparisons between observed and predicted flow data showed that there were differences in timing and magnitude of peak flows and the shape of recession curves (Figure 4a). This was primarily because of soft data such as tile drain information not being included (a major intra-watershed processes in the watershed) during model development. When tile drainage was included without calibration (S2) the surface runoff contribution was 11%, and the tile flow was 60% (Table II). The predictions in S2 were more realistic and close to the opinion of watershed experts. Graphical comparisons of observed and predicted flow showed that the timing and magnitude of peak flows and the shape of recession curves aligned better; however, the predicted flows were higher than observed (Figure 4b). The quantitative statistics were poor ($R^2 = 0.69$, NSE = 0.27, PBIAS = −57.29) despite more realistic contributions of tile flow, mainly because the predicted flows were higher (median observed = 14.43, median simulated = 31. 71) compared to observed flows. This is because of the inclusion of tile drains, which altered the hydrodrological processes and needs calibration to lower predicted flows and align better to improve statistics. When tile drainage was included and the model was calibrated (S3), surface runoff and tile flow contributed 28% and 53%, respectively. These contributions were close to the opinion of watershed experts. Graphical comparison (Figure 4c) shows that the observed and predicted flows align with each other and the quantitative statistics were very good ($R^2 = 0.84$, NSE = 0.82, PBIAS = −4.60). These results showed that using only quantitative performance statistics can be misleading and should not be used alone to make absolute modeling decisions (e.g. Developing Total maximum Daily Loads (TMDLs) or assessing the impacts of best management practices). However, combining quantitative statistics along with graphical comparisons of time series plots and the incorporation of literature or expert knowledge to account for all intra-watershed processes will ensure that hydrological processes and water balance are within reasonable limits and will produce better and more reliable predictions.

Table V. Quantitative statistics at validation locations for Spatial and Outlet calibration strategies

Calibration strategy	Time period	R1-V1		R1-V2		R2-V1		R3-V1		R4-V1		R4-V2		O-V1		O-V2	
		NSE	PBIAS	NSE	PBIAS	NSE	PBIAS	NSE	PBIAS	NSE	PBIAS	NSE	PBIAS	NSE	PBIAS	NSE	PBIAS
Spatial calibration	1990–1999	0.81	17.29	0.60	4.33	0.76	11.96	0.81	8.65	0.83	12.51	0.71	5.34	0.85	9.99	0.84	11.88
	2000–2006	0.83	13.23	0.81	15.05	0.87	7.65	0.88	10.99	0.84	12.68	0.75	17.27	0.91	7.86	0.86	4.42
Outlet calibration	1990–1999	0.78	4.14	0.64	16.46	0.70	15.55	0.69	8.84	0.71	15.26	0.70	9.43	0.77	7.96	0.74	10.27
	2000–2006	0.83	−4.25	0.78	12.10	0.83	10.74	0.81	11.37	0.79	14.98	0.76	10.54	0.85	5.03	0.78	1.88

Copyright © 2015 John Wiley & Sons, Ltd.

Influence of spatial and outlet calibration strategy

Spatial calibration strategy. In the head watershed locations, the NSE values ranged from 0.72 to 0.82 (mean, 0.80) and PBIAS values ranged from −4.72 to 11.89 (absolute [abs] mean 4.49) during the calibration period (Table IV). In the temporal validation period, the NSE and PBIAS ranged from 0.69 to 0.88 (mean, 0.78) and −8.82 to 26.19 (abs.mean, 9.98) (Table IV). In the spatial validation locations and in the calibration period, the NSE ranged from 0.60 to 0.85 (mean, 0.78) and PBIAS ranged from 4.33 to 17.29 (abs.mean, 9.17) (Table V). In the temporal validation period, the NSE ranged from 0.75 to 0.91(mean, 0.84) and PBIAS ranged from 4.42 to 17.27 (abs.mean, 9.99) (Table V). Quantitative statistics showed that spatial calibration strategy performed very good (based on NSE and PBIAS criteria) in the headwater watersheds as well as in the spatial validation locations. The performance was even better in the in the spatial validation locations, especially in the temporal validation period (mean NSE, 0.84 *vs* mean NSE, 0.78).

Outlet calibration strategy. In the headwater calibration locations and in the calibration period, the NSE and PBIAS ranged from 0.65 to 0.75 (mean, 0.73) and −10.72 to 13.72 (abs.mean, 8.46) in the headwater subbasins (Table IV). In the validation period, the NSE and PBIAS ranged from 0.62 to 0.86 (mean, 0.73) and −15.24 to 25.94 (abs.mean, 15.14) (Table IV). In the validation locations and in the calibration period, the NSE and PBIAS ranged from 0.64 to 0.78 (mean, 0.72) and 4.14 to 15.26 (abs.mean, 9.84). In the validation period, the NSE ranged from 0.76 to 0.85 (mean, 0.80) and PBIAS ranged from −4.25 to 14.98 (abs.mean, 7.96) (Table V). The quantitative statistics showed that the outlet calibration strategy rated as good based on NSE criteria and very good based on PBIAS criteria in the various locations of the watershed when it was calibrated at the outlet and verified across various locations in the watershed.

Comparing the two strategies showed that the spatial calibration strategy statistics slightly outperforms the outlet calibration statistics in all locations within the basin (based on NSE and PBIAS statistics). It should also be noted that the statistics at the outlet (O-V2) were better for the spatial calibration during calibration and validation periods mainly because the spatial variations within the basin were more realistically captured. However, the outlet calibration strategy still performed well for this watershed. This could be the result of the basin being very homogeneous with a flat topology and agriculture being the most prominent landuse. In addition, the tile drains which are major intra-watershed processes in the basin are represented in the model. Outlet calibration may perform even less well than the spatial calibration strategy for watersheds that have greater variation in topology and land use.

A time series plot (Figure 6) between observed and predicted flow using outlet calibration strategy and spatial calibration strategy at R4-H location showed that the spatial calibration strategy better represented the peaks and recession of the hydrographs. The outlet calibration strategy was either under or over predicting the peaks and recession. This again showed that using quantitative statistics alone may be misleading and unreliable. The use of graphical comparisons along with quantitative statistics resulted in better evaluation criteria. The results of this study showed that the spatial calibration strategy gave greater confidence in modeling efforts.

Effects of transfering parameters across scale

Spatial scale transfer. The monthly NSE values for NHDPlus model at OV2 and R4-H locations were 0.71 and 0.71, while the PBIAS values were 25.25 and 23.95, respectively (Table VI). Quantitative statistics showed that the NHDPlus model was good in the selected locations based on NSE criteria and satisfactory based on PBIAS criteria. Similar statistics were seen in other locations within the basin. Average annual hydrologic components of the water balance were evaluated in the NHDPlus model by comparing contributions of surface runoff and tile flow. Contribution of surface runoff was 29%, while the contribution of tile flow was 54%, which indicated that the intra-watershed processes (tile drainage)

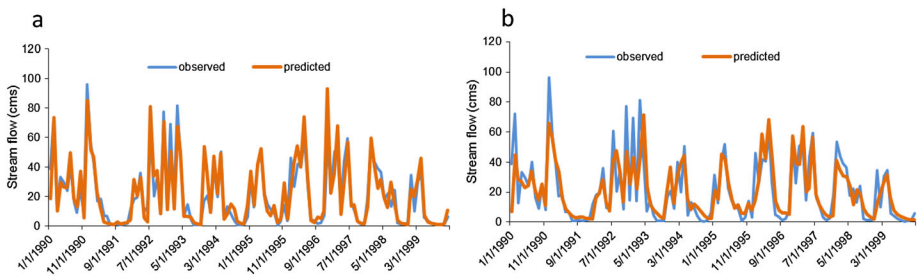


Figure 6. Time series plots at R4-H location using a) spatial calibration strategy and b) outlet calibration strategy

Copyright © 2015 John Wiley & Sons, Ltd.

Table VI. NSE and PBIAS values on daily, monthly and yearly at two locations (O-V2 and R4-H) for NHDPlus and HUC12 models

Spatial scale	Location	Temporal scale	NSE	PBIAS
NHD Plus	O-V2	Daily	−0.02	25.31
		Monthly	0.71	25.25
		Yearly	0.01	25.31
	R4-H	Daily	0.16	23.99
		Monthly	0.71	23.95
		Yearly	0.28	23.99
HUC-12	O-V2	Daily	0.70	10.59
		Monthly	0.82	10.53
		Yearly	0.80	10.60
	R4-H	Daily	0.67	−4.58
		Monthly	0.82	−4.60
		Yearly	0.87	−4.58

were captured reasonably well with spatial transfer of parameters. The time series graphs at both locations showed that there were some differences in timing and magnitude of peak flows (simulated data was under predicting most of the time); however, the NHDPlus model captured the pattern reasonably well (Figure 7). The differences are likely because of the very small (36% smaller) subwatershed size and irregular shape in NHDPlus model compared to the HUC-12 model (Figure 2). In smaller sized and irregular shaped subwatersheds, the length of the reach and time of concentration are small. This would result in faster transport of flow and associated constituents within a day or sometimes within hours after a rainfall event. The resulting daily hydrograph will have higher peaks and a very quick receding curve and thereby under or early prediction at monthly level (Figure 7). In HUC-12 model, the time of concentration and length of reach are longer resulting in more days for the flow to transport and the hydrograph captures peaks and recession reasonably well. Fine-tuning of the general parameters (e.g. time of concentration) may be needed for the NHDPlus model, after spatial transfer of parameters, to accurately capture the timing and magnitude of peak flows and the shape of rising and recession

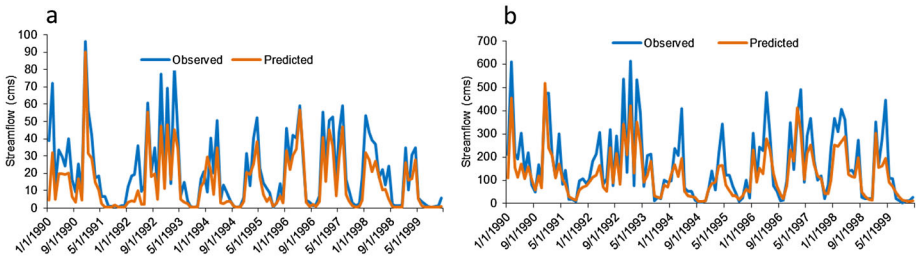


Figure 7. Monthly time series for NHDPlus model at (a) R4-H and (b) OV2 locations

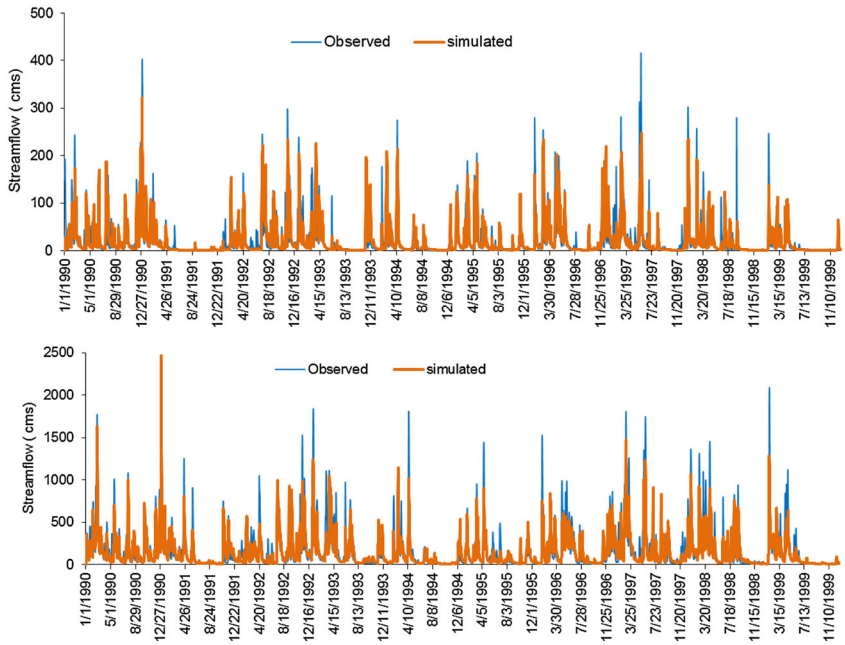


Figure 8. Daily time series plot for HUC-12 model at (a) R4-H (b) OV2 locations

Copyright © 2015 John Wiley & Sons, Ltd.

curves and to improve quantitative statistics. Fine tuning of the parameters and reporting associated results is beyond the scope of this paper and will be presented elsewhere (Yen *et al.* in prep).

Temporal scale transfer. The daily NSE and PBIAS values for the HUC-12 model at O-V2 location are 0.70, 0.80 and 10.59 and 10.60, respectively (Table VI). At the R4-H location, the statistics were 0.67, 0.87 and −4.58 and −4.58, respectively. Quantitative statistics showed that the performance of temporal scale transfer was good to very good based on NSE criteria and very good based on PBIAS criteria on daily and monthly time-step when the original model was calibrated at monthly time-step. The time series plots show that the simulated data was under predicting the peaks flows; however, it captured the timing and the shape of rising and recession curves well (Figure 8). General parameters, such as time of concentration, can be adjusted to better capture the peaks on daily time-step. Overall, the performance of temporal scale transfer was good in the HUC-12 model based on quantitative statistics and temporal time series plots. Quantitative statistics were poor for the NHDPlus model for daily and yearly time-step at both OV2 and R4-H locations (Table VI). The poor quantitative statistics are because of poor performance of the NHDPlus model at the monthly time-step as seen above after the spatial transfer of parameters from the HUC-12 model. Performance on daily and yearly time-steps may be improved after fine tuning the general parameters of the NHDPlus model.

CONCLUSIONS

Coarse resolution (HUC-12) and finer resolution (NHDPlus) models were developed within WLEB to demonstrate the significance of considering intra-watershed processes during model development, compare and contrast two calibration strategies (spatial calibration vs. outlet calibration) and evaluate spatial and temporal transfer of parameters.

We found that including intra-watershed processes (i.e. tile drainage) produced accurate and realistic hydrologic simulations. However, failure to include these processes may still result in a model that performs well according to model performance statistics. Thus, considering only model performance statistics may be misleading. The spatial calibration strategy produced better results in terms of quantitative statistics and graphical comparisons. The outlet calibration strategy also produced decent results in various locations within the watershed. This was likely because the WLEB is a fairly homogenous watershed. However, we believe that the spatial calibration strategy results in greater confidence for modeling efforts that support science-based decisions. Transferring parameters across temporal scales worked well with a similar spatial resolution model; however, additional fine tuning is required when transferring parameters across spatial scales to produce realistic results. This study showed that quantitative statistics should be used in conjunction with graphical comparisons and knowledge of the watershed (e.g. literature sources or expert knowledge) to ensure that hydrological processes and water balance are within reasonable limits and will produce better and more reliable predictions.

ACKNOWLEDGEMENTS

We would like to thank the USDA Natural Resource Conservation Service for funding this work through the Wildlife and Cropland components of the Conservation Effects Assessment Project (CEAP). We particularly want to thank Charlie Rewa for his unwavering support of this project through the Wildlife Component CEAP. This study represents a component of the much broader Western Lakes Erie Basin (WLEB) CEAP effort, and we would like to thank the project team for their contributions: Gust Annis, Kim Hall, Matt Herbert, August Froehlich, Stu Ludsin, Lee Norfleet, Maura O'Brien, Jeff Reuter, Anthony Sasson, Bill Stanley, Mari-Vaughn Johnson, Carrie Vollmer-Sanders and Chris Winslow. Finally, we would like to thank our WLEB CEAP Advisory Panel members that have provided critical input, review and data needed to successfully complete this project: Vicki Anderson, Jeff Ankney, Larry Antosch, Dave Baker, Steve Davis, Jane Frankenberger, Gail Hesse, Kevin King, Mary Knapp, Melinda Koslow, Ken Krieger, Greg LaBarge, Kip Miller, Dale Robertson and Steve Shine.

REFERENCES

Abbaspour KC, Yang J, Maximov I, Siber R, Bogner K, Mieleitner J, Zobrist J, Srinivasan R. 2007. Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT. *Journal of Hydrology* **333**: 413–430. DOI: 10.1016/j.jhydrol.2006.09.014

Arnold JG, Moriasi DN, Gassman PW, Abbaspour KC, White MJ, Srinivasan R, Santhi C, Harmel RD, van Griensven A, Van Liew MW, Kannan N, Jha MK. 2012a. Swat: model use, calibration, and validation. *Transactions of the ASABE* **55**: 1491–1508.

Arnold JG, Kiniry JR, Srinivasan R, Williams JR, Haney EB, Neitsch SL. 2012b. Soil and water assessment tool input/output file documentation: version 2012. Texas Water Resources Institute TR 436. http://swatmodel.tamu.edu/documentation/ (Accessed on May 2014).

Arnold JG, Youssef MA, Yen H, White MJ, Sheshukov AY, Sadeghi AM, Moriasi DN, Steiner JL, Amatya DM, Skaggs RW, Haney EB, Jeong J, Arabi M, Gowda PH. 2015. Hydrological processes and model representation: impact of soft data on calibration. *Transactions of the ASABE* (in Press).

Baker N. 2011. Tillage practices in the conterminous United States, 1989–2004—Datasets Aggregated by Watershed, U.S. Geological Survey Data Series 573.

Benham BL, Baffaut C, Zeckoski RW, Mankin KR, Pachepsky YA, Sadeghi AM, Brannan KM, Soupir ML, Habersack MJ. 2006. Modeling

Copyright © 2015 John Wiley & Sons, Ltd.

bacteria fate and transport in watersheds to support TMDLs. *Transactions of the ASABE* **49**: 987–1002.

Bingner RL, Garbrecht J, Arnold JG, Srinivasan R. 1997. Effect of watershed subdivision on simulation runoff and fine sediment yield. *Transactions of the American Society of Agricultural Engineers* **40**: 1329–1335.

Cao W, Bowden WB, Davie T, Fenemor A. 2006. Multi-variable and multi-site calibration and validation of SWAT in a large mountainous catchment with high spatial variability. *Hydrological Processes* **20**: 1057–1073. DOI: 10.1002/hyp.5933

Chaubey I, Chiang L, Gitau MW, Mohamed S. 2010. Effectiveness of best management practices in improving water quality in a pasture-dominated watershed. *Journal of Soil and Water Conservation* **65**: 424–437. DOI: 10.2489/jswc.65.6.424

Cho J, Bosch D, Vellidis G, Lowrance R, Strickland T. 2013. Multi-site evaluation of hydrology component of SWAT in the coastal plain of southwest Georgia. *Hydrological Processes* **27**: 1691–1700.

Daggupati P, Douglas-Mankin K, Sheshukov A, Barnes P, Devlin D. 2011. Field-level targeting using SWAT: mapping output from HRUs to fields and assessing limitations of GIS input data. *Transactions of the ASABE* **54**: 501–514.

Daggupati P, Pai N, Ale S, Douglas-Mankin KR, Zeckoski R, Jeong J, Parajuli PB, Saraswat D, Youssef MA. 2015. Recommended calibration and validation strategy for hydrologic and water quality models. *Transactions of the ASABE*, (in press).

Douglas-Mankin K, Daggupati P, Sheshukov A, Barnes P. 2013. Paying for sediment: field-scale conservation practice targeting, funding, and assessment using the Soil and Water Assessment Tool. *Journal of Soil and Water Conservation* **68**: 41–51.

Douglas-Mankin K, Srinivasan R, Arnold JG. 2010. Soil and Water Assessment Tool (SWAT) model: current developments and applications. *Transactions of the ASABE* **53**: 1423–1431.

Eckhardt K, Fohrer N, Frede H. 2005. Automatic model calibration. *Hydrological Processes* **19**: 651–658.

Foster GR. 2005. *Revised Universal Soil Loss Equation, Version 2 (RUSLE2)* Science Documentation (In Draft). USDA-ARS: Washington, DC.

Gassman PW, Reyes MR, Green CH, Arnold JG. 2007. The soil and water assessment tool: historical development, applications, and future research directions.

He Y, Bárdossy A, Zehe E. 2011. A review of regionalisation for continuous streamflow simulation. *Hydrology and Earth System Sciences* **15**: 3539–3553.

Jha MK, Gassman PW. 2014. Changes in hydrology and streamflow as predicted by a modelling experiment forced with climate models. *Hydrological Processes* **28**: 2772–2781.

Jin X, Xu C, Zhang Q, Chen YD. 2009. Regionalization study of a conceptual hydrological model in Dongjiang basin, south China. *Quaternary International* **208**: 129–137. DOI: 10.1016/j.quaint.2008.08.006

Kaatz MR. 1955. The Black Swamp, a study in historical geography. *Annals of the Association of American Geographers* **45**: 1–35.

Kane DK, Conroy JD, Richards RP, Baker DB, Culver DB. 2014. Re-eutrophication of Lake Erie, correlations between tributary nutrient loads and phytoplankton biomass. *Journal of Great Lakes Research* **40**: 496–501.

Kannan N, Santhi C, Di Luzio M, Potter S, Arnold JG. 2005. Measuring environmental benefits of conservation practices: the conservation effects assessment project (CEAP)-A model calibration approach at the national level. 2005 ASAE Annual International Meeting.

Karr JR, Toth LA, Dudley DR. 1985. Fish communities of midwestern rivers, a history of degradation. *BioScience* **35**: 90–95.

Klemes V. 1986. Operational testing of hydrological simulation models. *Hydrological Sciences Journal* **31**: 13–24.

Knisel WG, Douglas-Mankin KR. 2012. Creams/gleams: model use, calibration, and validation. *Transactions of the ASABE* **55**: 1291–1302.

Kumar R, Samaniego L, Attinger S. 2013a. Implications of distributed hydrologic model parameterization on water fluxes at multiple scales and locations. *Water Resources Research* **49**. DOI: 10.1029/2012WR012195

Kumar R, Livneh B, Samaniego L. 2013b. Toward computationally efficient large-scale hydrologic predictions with a multiscale regionalization scheme. *Water Resources Research* **49**: 5700–5714. DOI: 10.1002/wrcr.20431

Luo Y, Su B, Yuan J, Li H, Zhang Q. 2011. GIS techniques for watershed delineation of SWAT model in plain polders. *Procedia Environmental Sciences* **10**: 2050–2057. DOI: 10.1016/j.proenv.2011.09.321

Merz R, Blöschl G. 2004. Regionalisation of catchment model parameters. *Journal of Hydrology* **287**: 95–123.

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

Neitsch SL, Arnold JG, Kiniry JR, Williams JR. 2011. Soil and water assessment tool theoretical documentation: version 2009. USDA–ARS, Grassland, Soil and Water Research Laboratory, Temple, TX; and Blackland Research and Extension Center, Texas AgriLife Research, Temple, TX. Texas Water Resources Institute Technical Rep. 406, Texas A&M University System, College Station, TX. http://swatmodel.tamu.edu/documentation/ (Accessed on may 2014).

Parajuli PB, Nelson NO, Frees LD, Mankin KR. 2009. Comparison of AnnAGNPS and SWAT model simulation results in USDA-CEAP agricultural watersheds in south-central Kansas. *Hydrological Processes* **23**: 748–763.

Piniewski M, Okruszko T. 2011. Multi-site calibration and validation of the hydrological component of SWAT in a large lowland catchment. In *Modelling of Hydrological Processes in the Narew Catchment*, Świątek D, Okruszko T (eds). Springer-Verlag: Berlin Heidelberg; 15–41. DOI: 10.1007/978-3-642-19059-9_2

Qi C, Grunwald S. 2005. GIS-based hydrologic modeling in the Sandusky watershed using SWAT. *Transactions of the American Society of Agricultural Engineers* **48**: 169–180.

Santhi C, Arnold JG, Williams JR, Hauck LM, Dugas WA. 2001. Application of a watershed model to evaluate management effects on point and nonpoint source pollution. *Transactions of the American Society of Agricultural Engineers* **44**: 1559–1570.

Santhi C, Kannan N, Arnold JG, Di Luzio M. 2008. Spatial calibration and temporal validation of flow for regional scale hydrologic modeling. *Journal of the American Water Resources Association* **44**: 829–846.

Saraswat D, Frankenberger JR, Pai N, Ale S, Daggupati P, Douglas-Mankin KR, Youssef MA. 2015. Hydrologic and water quality models, documentation and reporting procedures for calibration, validation, and use. *Transactions of the ASABE*, (in Press).

Sears PB. 1941. Postglacial vegetation in the Erie–Ohio Area. *The Ohio Journal of Science* **41**: 225–234.

Seibert J, McDonnell J. 2002. On the dialog between experimentalist and modeler in catchment hydrology: use of soft data for multicriteria model calibration. *Water Resources Research* **38**: 23-1–23-14. DOI: 10.1029/2001WR000978

Seo M, Yen H, Kim MK, Jeong J. 2014. Transferability of SWAT Models between SWAT2009 and SWAT2012. *Journal of Environmental Quality* **43**(3): 869–880.

Sheshukov AY, Daggupati P, Douglas-Mankin K, Lee MM. 2011. High spatial resolution soil data for watershed modeling: 1. Development of a SSURGO-ArcSWAT utility. *Journal of Natural and Environmental Sciences* **2**: 15–24.

Srinivasan R, Zhang X, Arnold J. 2010. SWAT ungauged, hydrological budget and crop yield predictions in the upper Mississippi River basin. *Transactions of the ASABE* **53**: 1533–1546.

Stone MC, Hotchkiss RH, Hubbard CM, Fontaine TA, Mearns LO, Arnold JG. 2001. Impacts of climate change on Missouri River Basin water yield. *Journal of the American Water Resources Association* **37**: 1119–1129.

Trautman MB. 1939. The effects of man-made modifications on the fish fauna in Lost and Gordon Creeks, Ohio, between 1887–1938. *Ohio Journal of Science* **39**: 275–288.

Trautman MB, Gartman DK. 1974. Re-evaluation of the effects of man-made modifications on Gordon Creek between 1887 and 1973 and especially as regards its fish fauna. *The Ohio Journal of Science* **74**: 162–173.

Troy TJ, Wood EF, Sheffield J. 2008. An efficient calibration method for continental-scale land surface modeling. *Water Resources Research* **44**. DOI: 10.1029/2007WR006513

Tuppad P, Santhi C, Srinivasan R. 2010. Assessing BMP effectiveness: multiprocedure analysis of observed water quality data. *Environmental Monitoring and Assessment* **170**: 315–329. DOI: 10.1007/s10661-009-1235-8

Copyright © 2015 John Wiley & Sons, Ltd.

Van Liew MW, Garbrecht J. 2003. Hydrologic simulation of the Little Washita River experimental watershed using SWAT. *Journal of the American Water Resources Association* **39**: 413–426.

Vandewiele GL, Elias A. 1995. Monthly water balance of ungauged catchments obtained by geographical regionalization. *Journal of Hydrology* **170**: 277–291.

Veith T, Wolfe M, Heatwole C. 2003. Optimization procedure for cost effective bmp placement at a watershed scale. *Journal of the American Water Resources Association* **39**: 1331–1343. DOI: 10.1111/j.1752-1688.2003.tb04421.x

Wang S, Zhang Z, Sun G, Strauss P, Guo J, Tang Y, Yao A. 2012. Multi-site calibration, validation, and sensitivity analysis of the MIKE SHE Model for a large watershed in northern China. *Hydrology and Earth System Sciences* **16**: 4621–4632. DOI: 10.5194/hess-16-4621-2012

White KL, Chaubey I. 2005. Sensitivity analysis, calibration, and validations for a multisite and multivariable SWAT model. *Journal of the American Water Resources Association* **41**: 1077–1089. DOI: 10.1111/j.1752-1688.2005.tb03786.x

Xu C. 1999. Estimation of parameters of a conceptual water balance model for ungauged catchments. *Water Resources Management* **13**: 353–368.

Yen H, Bailey RT, Arabi M, Ahmadi M, White MJ, Arnold JG. 2014a. The role of interior watershed processes in improving parameter estimation and performance of watershed models. *Journal of Environmental Quality* **43**: 1601. DOI: 10.2134/jeq2013.03.0110

Yen H, Wang X, Fontane DG, Harmel RD, Arabi M. 2014b. A framework for propagation of uncertainty contributed by parameterization, input data, model structure, and calibration/validation data in watershed modeling. *Environmental Modelling and Software* **54**: 211–221. DOI: 10.1016/j.envsoft.2014.01.004

Yen H, White MJ, Arabi M, Arnold JG. 2014c. Evaluation of alternative surface runoff accounting procedures using the SWAT model. SWAT Special Issue. *International Journal of Agricultural and Biological Engineering* **8**.

Zhang X, Srinivasan R, Bosch D. 2009. Calibration and uncertainty analysis of the SWAT model using Genetic Algorithms and Bayesian Model Averaging. *Journal of Hydrology* **374**: 307–317. DOI: 10.1016/j.jhydrol.2009.06.023

Copyright © 2015 John Wiley & Sons, Ltd.

APPENDIX B: YEN ET AL. 2016. SCIENCE OF THE TOTAL ENVIRONMENT. HTTP://DX.DOI.ORG/10.1016/J.SCITOTENV.2016.06.202.

ARTICLE IN PRESS
STOTEN-20332; No of Pages 17

Science of the Total Environment xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Science of the Total Environment

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

Western Lake Erie Basin: Soft-data-constrained, NHDPlus resolution watershed modeling and exploration of applicable conservation scenarios

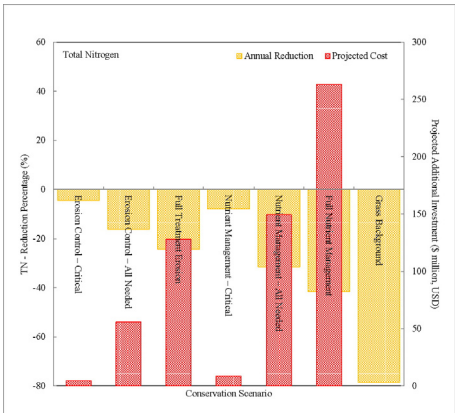
Haw Yen [a,*], Michael J. White [b], Jeffrey G. Arnold [b], S. Conor Keitzer [c], Mari-Vaughn V. Johnson [d], Jay D. Atwood [d], Prasad Daggupati [e], Matthew E. Herbert [f], Scott P. Sowa [f], Stuart A. Ludsin [c], Dale M. Robertson [g], Raghavan Srinivasan [e], Charles A. Rewa [h]

[a] Blackland Research and Extension Center, Texas A&M Agrilife Research, 720 East Blackland Road, Temple, TX 76502, USA
[b] Grassland, Soil & Water Research Laboratory, USDA-ARS, 808 East Blackland Road, Temple, TX 76502, USA
[c] Aquatic Ecology Laboratory, Department of Evolution, Ecology, and Organismal Biology, The Ohio State University, 1314 Kinnear Rd., Columbus, OH 43212, USA
[d] Grassland, Soil & Water Research Laboratory, Resources Assessment Division, USDA-NRCS, 808 East Blackland Road, Temple, TX 76502, USA
[e] Spatial Science Laboratory, Texas A&M University, Suite 221, 1500 Research Plaza, College Station, TX 77843, USA
[f] The Nature Conservancy, Michigan Field Office, 101 East Grand River Ave, Lansing, MI 48906, USA
[g] Wisconsin Water Science Center, U.S. Geological Survey, 8505 Research Way, Middleton, WI 53562, USA
[h] Resources Assessment Division, USDA-NRCS, George Washington Carver Center, Beltsville, MD, 20705, USA

HIGHLIGHTS

- NHDPlus data were conducted for flow and water quality in Western Lake Erie Basin.
- Practicable conservation scenarios were implemented to NHDPlus watershed project.
- Projected cost among different conservation scenarios was compared and investigated.
- Model responses by spring/summer seasons were identified by stream order.
- Improved biological conditions were studied by investment of conservation practices.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 11 May 2016
Received in revised form 23 June 2016
Accepted 25 June 2016
Available online xxxx

Editor: D. Barcelo

ABSTRACT

Complex watershed simulation models are powerful tools that can help scientists and policy-makers address challenging topics, such as land use management and water security. In the Western Lake Erie Basin (WLEB), complex hydrological models have been applied at various scales to help describe relationships between land use and water, nutrient, and sediment dynamics. This manuscript evaluated the capacity of the current Soil and Water Assessment Tool (SWAT) to predict hydrological and water quality processes within WLEB at the finest resolution watershed boundary unit (NHDPlus) along with the current conditions and conservation scenarios. The process based SWAT model was capable of the fine-scale computation and complex routing used in this

* Corresponding author.
E-mail addresses: haw.yen@gmail.com, hyen@brc.tamus.edu (H. Yen).

http://dx.doi.org/10.1016/j.scitotenv.2016.06.202
0048-9697/© 2016 Elsevier B.V. All rights reserved.

Please cite this article as: Yen, H., et al., Western Lake Erie Basin: Soft-data-constrained, NHDPlus resolution watershed modeling and exploration of applicable conservation s..., Sci Total Environ (2016), http://dx.doi.org/10.1016/j.scitotenv.2016.06.202

Keywords:
Lake Erie
NHDPlus model
Soft data
Conservation practice
Swat-SAS
Model calibration

project, as indicated by measured data at five gaging stations. The level of detail required for fine-scale spatial simulation made the use of both hard and soft data necessary in model calibration, alongside other model adaptations. Limitations to the model's predictive capacity were due to a paucity of data in the region at the NHDPlus scale rather than due to SWAT functionality. Results of treatment scenarios demonstrate variable effects of structural practices and nutrient management on sediment and nutrient loss dynamics. Targeting treatment to acres with critical outstanding conservation needs provides the largest return on investment in terms of nutrient loss reduction per dollar spent, relative to treating acres with lower inherent nutrient loss vulnerabilities. Importantly, this research raises considerations about use of models to guide land management decisions at very fine spatial scales. Decision makers using these results should be aware of data limitations that hinder fine-scale model interpretation.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

For the past two decades, complex watershed simulation models such as the Soil and Water Assessment Tool (SWAT; Arnold et al., 2012), Agricultural Policy/Environmental eXtender (APEX, Williams et al., 2012), and Hydrological Simulation Program-Fortran (HSPF, Bicknell et al., 1997) have been implemented for water resources assessment. These and similar advanced technologies enable modelers to investigate challenges associated with water supply, water quality, pollutant control, and other ecosystem services (Chou and Wu, 2014; Jha and Gassman, 2014; Gebremariam et al., 2014). Model outputs are used to inform land management decision-making and policy development at local to national scales (White et al., 2015, in press; Johnson et al., 2015).

As modeling technologies have been further refined to simulate more complex systems in finer spatial and temporal resolution, parameter inputs required to calibrate the models have also increased in complexity (Haan et al., 1995; Ajami et al., 2007). Model simulations of watershed responses to various agricultural management scenarios in terms of hydrological, sediment, and nutrient processes require a large number of physically/empirically based functions, but monitoring efforts in agricultural systems that could provide those inputs are limited. To address the challenges of providing sufficient input data and calibrating complex models (e.g., SWAT) at a fine spatial scale of resolution, efficient optimization techniques and uncertainty analysis need to be applied to ensure that statistically acceptable results are generated affordably and with limited computational time (Duan et al., 1992; Yen et al., 2014a). Because of the lack of monitored data to calibrate the model, assumptions are often necessary during fine-scale model development. Caveats associated with these assumptions should be given due consideration when interpreting modeled results.

Although models have traditionally been calibrated with hard data, it is increasingly necessary to use both hard and soft data to achieve model calibration. Non-temporal model outputs and the associated predictive uncertainty that can be stated as aggregated indices of watershed characteristics are considered soft data (Seibert and McDonnell, 2002). Soft data are typically associated with intra-watershed processes, which have been demonstrated to impact model predictions during model calibration (Yen et al., 2014b, 2014c). It is possible to generate model runs that perform well in terms of hard data outputs (model performance can be tentatively categorized by statistical thresholds; Moriasi et al., 2007), but simultaneously produce unreasonable soft data outputs (e.g. a simulated annual denitrification rate twice the rate of measured data in the field; Yen et al., 2014b, 2014c). Soft and hard data should both be carefully considered to ensure that the calibrated model is appropriate for application in further analysis (Efstratiadis and Koutsoyiannis, 2010).

Trade-offs in the relationship between model resolution and model performance (accuracy) are of keen interest in the field of water resources analyses, because these models increasingly serve as land management and policy development decision support tools (Chaubey et al., 2005). This is particularly true in the Western Lake Erie Basin (WLEB), where nutrient management strategies and water quality impacts have been a focus of modeling for decades (Di Toro et al., 1973). Model calibration strategies have necessarily kept pace with model development, enabling enhanced spatial and temporal resolution, so that model predictions are more relevant to field-scale and local policy level decision-making processes (Cotter et al., 2003). However, efforts related to widespread and long-term monitoring of water quality in WLEB have not kept pace with model development. Robertson and Saad (2011) noted that the reduction in tributary monitoring efforts in this region makes evaluation of current nutrient and sediment loadings difficult. Current monitoring data in the Lake Erie drainage basin is inadequate to measure impacts of agricultural management on water quality (Betanzo et al., 2015) and thus insufficient to support the scales at which models simulating these impacts are currently being applied. This is one reason that the use of soft data is necessary for model calibration and one of the major challenges to model validation.

The SWAT model continues to provide informative analyses at increasingly more refined spatial scales. For example, the National Cropland reports published by the Conservation Effects Assessment Project (CEAP) evaluate the impacts of agricultural management and conservation practice adoption with the APEX (simulation under field scale level) and SWAT (combining outputs from APEX and routing among subwatersheds) models, based on data valid at the 8-digit Hydrologic Unit Code (HUC-8) scale of resolution for cropland in the contiguous USA (USDA-NRCS, 2011; Johnson et al., 2015). The CEAP-Cropland reports inform conservation policy decision making, including the Farm Bill's conservation spending budget. The USEPA (U.S. Environmental Protection Agency) supported Hydrologic and Water Quality System (HAWQS) is a user-friendly online system that applies SWAT to perform scenario analyses in multiple resolution delineation formats, including the HUC-8, -10, and -12 digit scales (USEPA, 2015; Yen et al., 2016a).

The NHDPlus dataset (National Hydrography Dataset) is the finest resolution watershed boundary unit dataset currently available for large-scale modeling applications; the average size of subwatersheds is 2.6 km^2. There is only one large-scale watershed modeling implementation of NHDPlus dataset reported to date. Daggupati et al. (2015) applied the SWAT model to compare accuracy of simulation of hydrologic processes between HUC-12 and NHDPlus scales in the Western Lake Erie Basin (WLEB). Building upon the work performed by Daggupati et al. (2015), this study uses both hard and soft data (annual denitrification rate and nitrate (NO_3) loads contributed from tile drainage system) to develop SWAT models to describe sediment and nutrient (total phosphorus and total nitrogen) dynamics at the NHDPlus scale in WLEB.

In 2002, the Farm Bill passed by the United States Congress included an 80% increase in allocation of federal funds for agricultural conservation relative to the funding amount in the previous Farm Bill (Johnson et al., 2015). CEAP was initiated by the USDA to conduct credible scientific evaluation of the benefits derived from the increased use of federal financial resources for agricultural conservation (Mausbach and Dedrick, 2004). A major goal of CEAP is to quantify the impact of agricultural conservation practices on water quality (e.g. sediment, nitrogen,

phosphorus, and pesticide residues). The focused and scientifically credible outcomes of these analyses can be used to help inform future conservation practice and program investments, and may serve to guide conservation research inquiry and investments (White et al., 2014).

The impetus behind the development of this high resolution modeling project was a need expressed by the Conservation Effects Assessment Project (CEAP) - Wildlife Component for a product capable of accurate predictions of nutrient and sediment dynamics in the many very small streams systems across WLEB in order to conduct aquatic stream health analyses. SWAT model outputs developed from simulating the watershed at this scale could be used as data inputs in biological models, such as those used to assess stream health and inform stream conservation decisions (Keitzer et al., in press). A principal caveat associated with this work is that while the model may be capable of producing appropriately scaled and reasonable-seeming results, insufficient quantitative data exist against which to compare the various model outputs at this scale in order to determine whether the simulated spatial specificity is reflective of reality across the entire simulated WLEB. More monitoring is necessary in order to provide sufficient data for complete model validation.

This implementation of the SWAT model in a sophisticated watershed analysis of WLEB at the finest available, NHDPlus spatial resolution enhances both model performance and the predictive capacity associated with cross-field investigations, enabling better modeling results in the future. In considering the increasing interests in water quality issues in WLEB, this study reports on a unique, novel, and timely application of the SWAT model, in which NHDPlus resolution data are used to: (1) explore soft/hard data applications on NHDPlus scale model development and subsequent predictions of streamflow, and sediment and nutrient dynamics associated with agriculture in the WLEB; (2) demonstrate the utility of modeling at this scale to describe the possible costs and benefits of various conservation practice adoption scenarios in the WLEB. The most important reason that the NHDPlus stream network was adopted in this CEAP-Wildlife study is because the fish community research associated with potential/current contaminants requires detailed information such as simulated flow, sediment, and nutrients at the stream reach scale. In addition, the approach conducted in this study that pairs conservation scenarios with projected additional costs is particularly rare in literature, since that requires mutual work among modelers (engineers), biologists, and economists. Spatially explicit simulations like these provide important information to consider when managing agricultural lands - knowledge gaps could be exposed that focused research efforts could help resolve.

2. Materials and methods

2.1. The SWAT model

The SWAT model was used to simulate streamflow and water quality in WLEB (Fig. 1). SWAT is a process/empirical-based, quasi-distributed, continuous time-step model developed by the United States Department of Agriculture – Agricultural Research Service (USDA-ARS) to perform large-scale watershed simulations based on specified soil, landuse, weather, and topographic data (Arnold et al., 1993, 2012). Potential impacts of current and changing land management and conservation practice adoption on ecological goals and concerns can be evaluated by simulating scenarios and analyzing predicted impacts on hydrologic, sediment, and nutrients processes (Daggupati et al., 2011; Keitzer et al., 2016; Scavia et al., 2016). SWAT is one of the most broadly applied large-scale watershed simulation models in the field of water resource planning and management (Gassman et al., 2007) and is among the toolsets used to inform conservation practice policy within the USA (Johnson et al., 2015; Mausbach and Dedrick, 2004).

The publicly available version of SWAT is currently parameterized at three spatial scales: (1) basin level; (2) subbasin level; and (3) HRU (Hydrologic Response Unit) level. HRUs are a unique combination of slope, landuse, and soils; there can be many HRUs per subbasin whereas parameters for a HRU can be very specialized by user discretion. The HRU level parameters are specifically assigned to all HRU files so that the parameter values may be altered by the modeler for selected HRUs. The subbasin level parameters are assigned uniformly within a given subbasin, and the basin level parameters are uniformly applied to the entire basin defined by the SWAT project. It was stated that SPCON (linear parameter to adjust sediment load in channel sediment routing (Neitsch et al., 2011)) is one of the most sensitive parameters involved in sediment and sediment-associated phosphorus calibration and calculating transportable maximum sediment load (Chu et al., 2004; White and Chaubey, 2005). However, SPCON is a parameter specified at the basin level, so it is possible for SWAT simulations to over- or under-estimate the basin sediment and sediment-associated phosphorus loads, especially because most watersheds do not have homogeneously distributed topography (Neitsch et al., 2011). To improve model predictions of sediment and sediment-associated phosphorus dynamics in this study, a revised version of SWAT, named SWAT-SAS (Subregional Adjustment of Sediment) was developed, which allows SWAT users to specify SPCON values for each subbasin within the basin of interest. These subbasin specific adjustments in the *.rte files enhance the spatial accuracy of sediment and sediment-associated phosphorus predictions. More details of SWAT-SAS are described in Appendix A.

2.2. Soft-data-constrained calibration

Large-scale watershed modeling efforts with sophisticated streamflow, sediment, and nutrient dynamic simulation processes necessarily require the incorporation of a large number of model parameters, some of which are difficult to populate with currently available datasets. Various parameter estimation (or, calibration) techniques have been proposed to solve challenging high-dimensional watershed calibration problems (Yen et al., 2014a). The parameter estimation process (e.g., automatic or manual calibration) is designed to minimize the error term between the observed and simulated data in a given time series (e.g., daily streamflow, monthly sediment load). The temporal data used to calculate error statistics during calibration process can be categorized as hard data (Seibert and McDonnell, 2002). Modelers can use manually defined statistical guidelines, such as the General Performance Ratings (Table 2), to evaluate the performance of error statistics (Moriasi et al., 2007). However, it was stated in literature that watershed simulation outputs may be inconsistent with real hydrological mechanisms and/or intra-watershed processes (Yen et al., 2014b).

In addition to hard data in the form of temporal series, soft data are defined as non-temporal measures of a watershed in reflecting intra-watershed processes such as denitrification, average annual sediment loading, or ratio of nitrate attributed from subsurface versus surface flow (Yen et al., 2016b). Developing calibration routines without due consideration of soft data may produce excellent model outputs in terms of error statistics, but the simulations may violate actual watershed behavior. For example, in a case study at the Eagle Creek Watershed, Indiana, USA, an auto-calibrated SWAT model without consideration of the region's prevalent tile drainage system might provide accurate stream flow predictions, but would incorrectly attribute the majority of flow to surface runoff losses (Yen et al., 2014b). The presence of the tile drainage system requires the modeler to apply additional constraints to the model simulation rather than relying on the auto-calibration system to correctly designate flow to the appropriate loss pathways (Yen et al., 2014b).

In much of the Midwest region (MWR) of the USA, tile flow pathways contribute to significant subsurface flow pathways; if model results are used to inform conservation practice adoption decisions or

ARTICLE IN PRESS

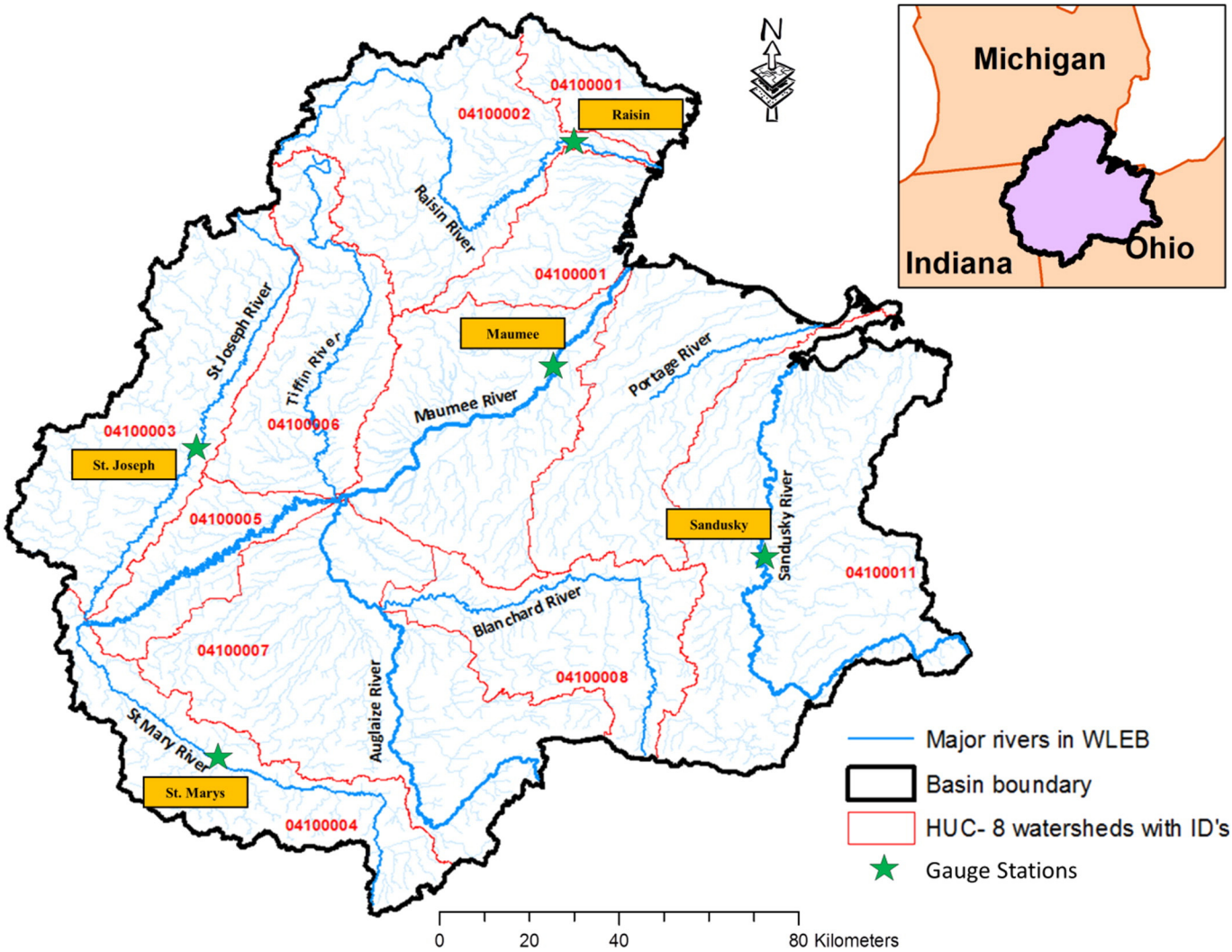


Fig. 1. Location of the Western Lake Erie Basin and the representative gauging stations.

conservation-related policy, overlooking the significance of this loss pathway could lead to inappropriate land management decisions. In addition, the associated predictive uncertainty may also be affected considerably (Yen et al., 2014c). It was indicated that the use of soft data may have direct impact on both parameter and predictive uncertainty. Some relevant parameters may not explore the full spectrum of given ranges because of the incorporation of additional soft-data-constraints (however, exploration of uncertainty of hard/soft data is not the primary goal in this study so it will not be fully investigated). In this study, soft-data-developed constraints were added to the denitrification rate and amount of nitrate (NO_3) contributed from tile flow to help govern the simulated intra-watershed processes in the region. The denitrification rate is constrained to be <50 kg/ha (David et al., 2009) and the ratio of NO_3 contributed by tile flow is constrained to be no less than two-thirds of total NO_3 losses (Schilling, 2002). For auto-calibration routines in practice, projected (or proposed) candidate parameter sets are generated by the current best solution evaluated by statistical performance. By the incorporate soft-data-constrained calibration, simulated results which violate the two constraints are automatically rejected and will not be considered as the candidate parameter set. In this case, the finalized calibration results will follow the constrained watershed behavior automatically. The approach of soft-data-constrained calibration guarantees that quality of derived model outputs to be representative to the real world in terms of better reflecting hydrological and water quality processes. On the other hand, modelers and engineers can take advantage of this approach to avoid the risk of generating good modeling results for the wrong reasons.

2.3. Study area description

The WLEB is located in the Midwestern United States of Ohio (76%), Indiana (17%) and Michigan (7%), has a drainage area of 23,817 km^2, and is bounded by the Raisin River in the north and the Sandusky River in the south (Fig. 1). Elevation changes in the WLEB are fairly mild (elevation varies from 246 m to 387 m above sea level), with an average slope of around 2%. The mean winter temperature is −5 °C and temperatures can rise to 29 °C in summer (available from the National Oceanic and Atmospheric Administration (NOAA) Cooperative Observer network and Weather- Bureau-Army-Navy stations). Average annual precipitation ranges between 838 and 940 mm. The major land use in WLEB is cultivated cropland (70%), dominated by corn-soybean rotations; forest (12%) and urban (12%) lands comprise the other dominant land uses in the region. A significant portion of the agricultural lands in WLEB, like most agricultural lands across the MWR, are installed with well-organized tile drainage systems, which enable agricultural activities (Kaatz, 1955) in a region previously dominated by a wetland (4000 km^2) called the Great Black Swamp (Mitsch and Gosselink, 2007). Although agricultural tile drainage is used in other regions of the globe, such as Canada and Europe, the extensiveness and level of sophistication of the MWR's large-scale agricultural tile drainage system is

Please cite this article as: Yen, H., et al., Western Lake Erie Basin: Soft-data-constrained, NHDPlus resolution watershed modeling and exploration of applicable conservation s..., Sci Total Environ (2016), http://dx.doi.org/10.1016/j.scitotenv.2016.06.202

unique. Model simulations of the impacts of various land management scenarios on conservation concerns in this unconventional watershed must consider both natural and anthropogenic influences on the mechanisms of surface and subsurface flow (groundwater, tile drainage flow) in order to assess the actual watershed behavior.

2.4. Model setup/incorporation using NHDPlus data

Two SWAT models with different spatial resolutions were developed for the WLEB region: (1) a HUC-12 scale model; and (2) an NHDPlus scale model, at the 30-meter Digital Elevation Model (DEM) for topographical features (Fig. 2). The HUC-12 model was built on the Watershed Boundary Dataset (WBD) (1:24,000 scale-level) with 391 subbasins, which is available online (http://datagateway.nrcs.usda.gov). The sizes of subbasins in the HUC-12 model ranged from 25 to 191 km^2 (average: 72 km^2). The WBD dataset is conjointly synchronized by the United States Geological Survey (USGS), United States Department of Agriculture – Natural Resources Conservation Service (USDA-NRCS), and the United States Environmental Protection Agency (USEPA). The NHDPlus dataset is coordinated by USGS and USEPA; associated data (at 1:100,000 scale-level) is available online (http://www.horizon-systems.com/NHDPlus/index.php). The NHDPlus model includes 11,128 subbasins, which range in size from 0.001 to 80 km^2 (average: 2.6 km^2). Daily streamflow calibrations in the NHDPlus version were developed by Daggupati et al. (2015).

SWAT models were set up using the Geographic Information System (GIS) interface of ArcSWAT2012 (rev593) (supported by ArcGIS10.1). Since it is time-consuming to execute a SWAT project in NHDPlus resolution (e.g. 10 h to complete a single 13-year SWAT run using an Intel ® Core™ i5-2500 k CPU @ 3.30 GHz, 64-bit operating system, Microsoft Windows 7 Professional), a HUC-12 resolution SWAT model of WLEB was built and calibrated for streamflow (each iteration consumes <2 h for the same running period in NHDPlus). The calibrated streamflow parameters were then transferred from the HUC-12 designed model to the NHDPlus project. Details can also be found in Daggupati et al. (2015).

Soils data were acquired from the publicly available USDA-NRCS State Soil Geographic database for the Conterminous United States (STATSGO at 1:250,000 scale-level) (USDA-NRCS, 1995). The simulated land uses were derived from the 30-meter resolution Crop Data Layers (CDLs), which include major rotations of local crops (Srinivasan et al., 2010). The slope information is incorporated with the soil and landuse data by HRU, where three slope categories (0–2%, 2–5%, and >5%) were defined, with the threshold area of 50 ha for each. In addition, row-crop agriculture was represented in HRUs without implementing pre-defined thresholds. The total number of HRUs for the HUC-12 and the NHDPlus models was 13,156 and 34,807, respectively.

Management practice data for agricultural activities were obtained from multiple sources. Cropland tillage practices were available from USGS survey data (Baker, 2011). Operation schedules designed for the RUSLE2.0 erosion model were acquired from USDA-ARS (Foster, 2005). Fertilization rates were obtained from National Agricultural Statistics Service (NASS) reports (USDA-NASS, 2014) of average crop yield at the county level. The assembled data were organized in the format of SWAT management files by a software package in VB.net (code available upon request). The management operations for agricultural activities (e.g., harvest/kill operations) were defined by date, and heat units were applied to those without data. In SWAT operation, historical precipitation was used for the years being simulated (e.g., from 1990 to 1999). On the other hand, it is very difficult to obtain the corresponding management details in each and every year in practice. Therefore, the projected operation schedule was implemented to reflect the management operation during simulation. Since spatial data on tile drainage systems are not publicly available, it is not feasible to allocate tiles specifically to currently drained field sites. Therefore, to accommodate for the impacts of tile drainage, the existence of tile drainage was assumed on all agricultural soils with <1% slope and soils categorized as poorly drained per the Soil Survey Geographic database (http://www.nrcs.usda.gov/wps/portal/nrcs/detailfull/soils/home). The HUC-12 and the NHDPlus SWAT models use the same temperature and precipitation data (from 1960 to 2010) inputs, available from the Weather Bureau Army Navy (WBAN) and Cooperative Observer Network Stations collected by National Oceanic and Atmosphere Administration (NOAA).

Available measured streamflow, sediment, and water quality record data were used to conduct model calibration (1990–1999) and validation (2000–2006) (Table 1). More details on streamflow calibration in the NHDPlus scale model can be found in Daggupati et al. (2015). Daily loads for total nitrogen, total phosphorus, and total suspended solids data used in calibration and validation of nutrient and sediment processes were estimated from measured USGS values with the Fluxmaster load estimation program (Schwarz et al., 2006; Robertson

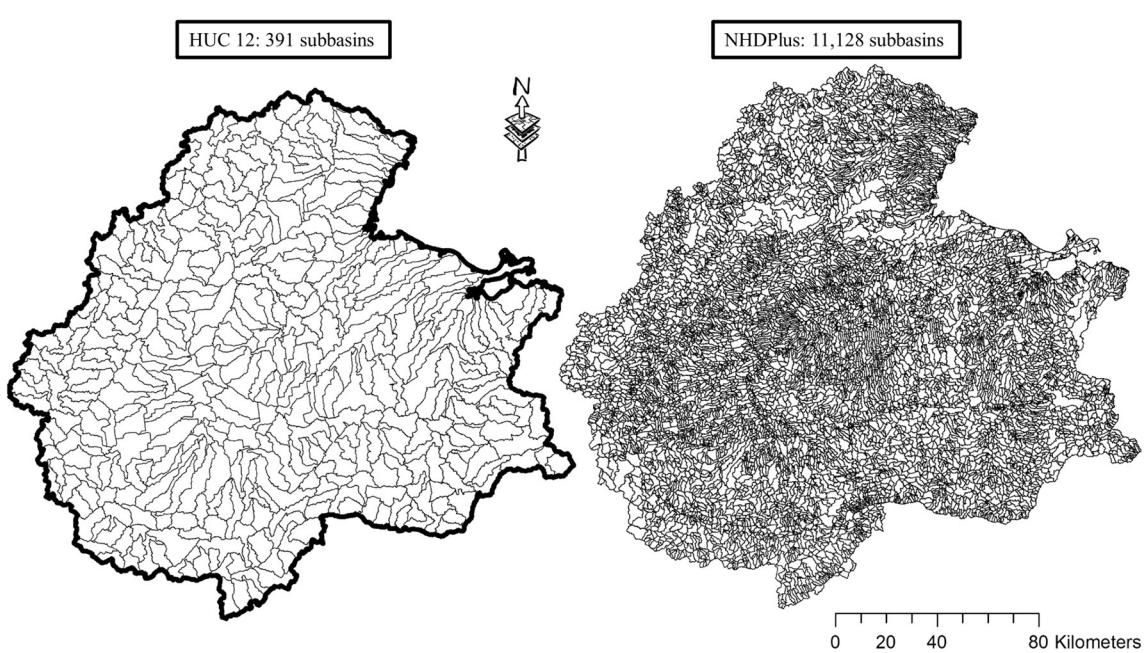


Fig. 2. The HUC-12 and the NHDPlus models in WLEB (Daggupati et al., 2015).

Please cite this article as: Yen, H., et al., Western Lake Erie Basin: Soft-data-constrained, NHDPlus resolution watershed modeling and exploration of applicable conservation s..., Sci Total Environ (2016), http://dx.doi.org/10.1016/j.scitotenv.2016.06.202

ARTICLE IN PRESS

Table 1
General information of calibration/validation data sources.

Station ID	HUC-8 number	Drainage area (km^2)	Location	Data source (streamflow)	Data source (nutrients)
4176500	04100002 (Raisin)	2680 (11.3%)	River Raisin near Monroe, MI	Heidelberg College River Studies	Heidelberg University & USGS
4178000	04100003 (St. Joseph)	1541 (6.5%)	St. Joseph River near Newville, IN	United States Geologic Survey	USGS
LES040-0007	04100004 (St. Marys)	1311 (5.5%)	St. Marys River	Indiana Department of Environmental Management	Indiana Department of Environmental Management
4193500	04100009 (Maumee)	16,270 (68.3%)	Maumee River at Waterville, OH	Heidelberg College River Studies	Heidelberg University, USGS & OH Environmental Protection Agency
4198000	04100011 (Sandusky)	2015 (8.5%)	Sandusky River near Fremont, OH	Heidelberg College River Studies	Heidelberg University, USGS & OH Environmental Protection Agency

The selected representative gauge stations are used in calibration/validation processes where details of the selection procedure can be found in Daggupati et al. (2015).

and Saad, 2011) (Table 1). The LOAD ESTimator (LOADEST) program (Runkel et al., 2004) was used to estimate daily loads for soluble phosphorus and NO_3 based on measurements of grab samples taken at the gauging stations (Table 1).

2.5. Model calibration/validation

SWAT has been applied in WLEB at coarser scales than the analyses presented here, including 350 subbasins and 4416 HRUs in total; (Bosch et al., 2011, 2013) and 4-digit HUCs (USDA-NRCS, 2011). In this study SWAT simulated 391 subbasins and 13,156 HRUs at HUC-12 scale; 11,128 subbasins and 34,807 HRUs at NHDPlus scale.

The potential impacts of nitrogen, phosphorus, and sediment loadings on aquatic species can be evaluated through standardized indices such as IBI (Index of Biotic Integrity) and dissolved oxygen in a given water body. Excess nitrogen and phosphorous concentrations in freshwater can pose a serious threat to freshwater biodiversity and many of the valuable ecosystem services provided by freshwater ecosystems (Carpenter et al., 1998; Smith et al., 1999). For instance, nutrient loading from WLEB watersheds has contributed to eutrophication of Lake Erie and threatens commercial and recreational fisheries, tourism, and the supply of safe drinking water (Michalak et al., 2013; Kane et al., 2014). In the United States, nutrient enrichment is a major contributor to degraded stream conditions (USEPA, 2006). Water quality (drinkability, fishability, swimability, etc.) issues in WLEB continue to concern municipalities and the general public. In 2014, potentially 500,000 water users faced a drinking water crisis in and around Toledo, Ohio, USA related to an algal bloom, a symptom of eutrophication (Dungjen and Patch, 2014). This study provides scientific reference to nitrogen, phosphorus, and sediment dynamics, which play a role in the development of harmful algal blooms (HABs) and toxicity of the blooms. Although the drivers of HABs and their toxicity are poorly understood, they are clearly an emerging threat to freshwater systems (Brooks et al., 2016.). This report and the SWAT model development associated with it may help inform land management and policy decisions to help reduce the likelihood and/or frequency of recurrences of similar emergency situations.

As mentioned earlier, SWAT streamflow calibration and validation in WLEB was conducted in previous work (Daggupati et al., 2015). Weather data used for the calibration period for the nutrients and sediment dynamics was from 1990 to 1999, with a three-year additional simulation warm-up period (1987 to 1989); the validation was conducted on 2000 to 2006 data. Since the NHDPlus project was computationally demanding, the sediment and nutrients calibration of WLEB was conducted manually by expert judgement.

To evaluate the model performance under the developed calibrations, two quantitative error statistics were used: the Nash-Sutcliffe Efficiency Coefficient (NSE) and the Percent Bias (PBIAS). The NSE (Nash and Sutcliffe, 1970) has been implemented in a wide variety of watershed modeling topics (Servat and Dezetter, 1991; ASCE, 1993) and has proven effective at predicting reasonable temporal outputs, especially on seasonal peaks (Santhi et al., 2001). The PBIAS statistical measure, on the other hand, performs better at capturing trends in consistent watershed responses, such as average flow rate, which is important in maintaining stabilized fish populations (USDA-NRCS, 2014). The potential magnitudes of NSE range from $-\infty$ to 1. Perfect matches between model predictions and observation data are indicated when NSE equals one. On the other hand, the best PBIAS value is 0%; underestimation is apparent when PBIAS >0%, and overestimation is signified by PBIAS <0%. General performance ratings of NSE and PBIAS values follow the convention developed by Moriasi et al. (2007) (Table 2).

2.6. Conservation scenarios

The CEAP-Croplands report on the Great Lakes region reports on conservation practices and agricultural practices in use from 2003 to 06 (USDA-NRCS, 2011). The report classifies all cropland acres into "needs classes" based on inherent vulnerabilities to leaching and or runoff and levels of conservation treatment. In addition to analyses of "current conditions" simulations and their impacts on conservation concerns, the CEAP-Croplands National reports contain analyses on modeled hypothetical conservation scenarios that estimate potential impacts of various conservation strategies associated with prioritizing treatment of acres according to "needs classes" (USDA-NRCS, 2011). Simulation models were used to estimate potential benefits and costs associated with applying prescribed conservation practices to any or all of the classes of acres. Per the CEAP-Croplands classification conventions, "Critical Needs" acres have a high need for additional conservation treatment; these acres include the most vulnerable of under-treated acres, those with the fewest conservation practices in place, where the highest losses of sediment and/or nutrients can be expected if no further conservation actions are taken. "Moderate Needs" acres have an intermediate need for additional conservation treatment; these acres are under-treated, but have lower inherent levels of vulnerability to losses or have some effective conservation practices in place. "Low Needs" acres have low inherent vulnerabilities to losses, or are adequately treated to address those vulnerabilities; these acres may still suffer annual nutrient and sediment losses that could be lessened with additional treatment. Further, current treatment levels on Low Needs acres must be maintained in order to ensure low nutrient and sediment losses in the future.

One of the objectives is to demonstrate the potential outcomes of watershed modeling at NHDPlus resolution by incorporating possible costs via different scenarios of conservation practices. Therefore, settings of parameters in each scenario are the same in order to conduct less biased comparisons. The major differences in each scenario are the level of nutrient reduction by alternated conservation practices applied (e.g., erosion control, nutrient management) and resultant impacts on losses of nutrients and sediment from farm fields. The scenarios were selected based on scenarios that CEAP-Croplands had applied in previous reports (USDA-NRCS, 2014, 2015, 2016). Using these scenarios broadens the potential implications of the conducted analyses, as it allows for future comparison across regions that were

Please cite this article as: Yen, H., et al., Western Lake Erie Basin: Soft-data-constrained, NHDPlus resolution watershed modeling and exploration of applicable conservation s..., Sci Total Environ (2016), http://dx.doi.org/10.1016/j.scitotenv.2016.06.202

Table 2
General performance ratings.
Moriasi et al. (2007).

Categories of general performance rating	NSE	PBIAS (%)	
		Streamflow	Sediment & nutrients
Very good	0.75 < NSE ≤ 1.00	PBIAS < ±10	PBIAS < ±25
Good	0.65 < NSE ≤ 0.75	±10 ≤ PBIAS < ±15	±25 ≤ PBIAS < ±40
Satisfactory	0.50 < NSE ≤ 0.65	±15 ≤ PBIAS < ±25	±40 ≤ PBIAS < ±70
Unsatisfactory	NSE ≤ 0.50	PBIAS ≥ ±25	PBIAS ≥ ±70

NSE: Nash-Sutcliffe efficiency coefficient.
PBIAS: Percent bias.

similarly simulated in CEAP-Croplands reports. In this study, eight simulations were performed with the SWAT model developed for these analyses. The calibrated SWAT model represents the Baseline condition (BL) and serves as the basis from which the other scenarios were developed. The conditions observed in the BL scenario are the result of extensive conservation practices already in use in the region. Every cultivated cropland acre in the WLEB region was treated with an average of 1.8 conservation practices in 2003–06 and 2.4 conservation practices in 2012. Conservation practices in place in WLEB in 2003–06 represented a $208 million annual investment ($43.39 per acre), whereas conservation practices in place in 2012 represented a $277 million annual investment ($56.98 per acre; USDA-NRCS, 2016). The data on practices and management used to develop the BL were primarily from the time period between 2003–06 and 2012.

Conservation treatment types were classified into two groups: erosion control practices alone or erosion control practices in conjunction with nutrient management practices. Simulation of six treatment scenarios explores the potential impacts of treating acres in particular needs classes with one or both treatment practices at the watershed scale. Acres in each of the three needs classes defined in the CEAP-Croplands report on the Great Lakes (NRCS-2011) were treated with either structural practices or both structural and nutrient management practices (Table 3).

An additional Grass Background (GBG) scenario was simulated to provide context for the predicted impacts of "current" agricultural management. The GBG scenario provides an estimate of sediment and nutrient losses in the absence of agriculture (replaced by grass cover). The eight simulated scenarios were analyzed to provide estimations of annual financial costs and conservation impacts of various watershed-scale conservation strategies (Table 3 and Fig. 4).

The primary conservation practice cost data source was the 2010 official state USDA-NRCS Payment Schedule database, augmented with cropland rental rates, commodity prices, and fertilizer prices from other published sources (USDA-NRCS, 2012). The NRCS National Conservation Plans (NCP) database was used to estimate the average number of units of practice per protected acre by state and practice. Rates from the NRCS Technical Service Provider database were used as a proxy for Technical Assistance (TA) costs. Since a farm field may have a variety of practices applied, each with a different useful life span, a special Equivalent Net Annual Value formula was used to amortize the cost of each practice, sum the costs across practices for each sample point, and calculate the annualized cost per acre of the treatment alternative (Boardman et al., 2001). The costs used here include the full cost of planning, installation, maintenance, and forgone income on land converted to conservation cover, regardless of whether some or all of the cost would be partially reimbursed to the farmer or incurred by a non-farm entity such as a federal agency. State average cropland rental rates were used as the cost of land converted from active crop production to a conserving use such as filter strips or buffers (USDA-NASS, 2011). All conservation practice costs were converted to units of practice per protected acre, using data from the NCP database. For example, the NCP data may show that on average 1.6 acres of buffer strip are used per 40 acres of cropland field, for a ratio of 0.04 acres of buffer practice per acre of protected cropland. If the buffer strip annualized cost per acre of buffer was $100, then for this example, the modeled cost would be $4 per protected acre ($100 ∗ 0.04).

Table 3
Definitions of CEAP scenarios and the corresponding cost.
USDA-NRCS (2015, 2016).

Conservation practices	Content	Projected additional cost $ (USD)	Treatment needs level		
			High	Moderate	Low
Baseline	BL	162,500,000†	–	–	–
Erosion control - critical	ECC	4,504,349	EC	–	–
Erosion control - all needed	ECA	55,782,326	EC	EC	–
Full treatment erosion control	FT	128,262,006	EC	EC	EC
Nutrient management - critical	NMC	8,394,761	EC + NM	–	–
Nutrient management - all needed	NMA	149,253,423	EC + NM	EC + NM	–
Full nutrient management	NM	263,393,414	EC + NM	EC + NM	EC + NM
Grass background	GBG	–	–	–	–

BL: The default scenario that incorporates current watershed information and the associated management practices in the field. The baseline scenario will be calibrated before conducting the following assessment in varying scenarios.
ECC: The ECC scenario includes region in the watershed with high treatment need for the target of erosion control.
ECA: The ECA scenario includes region in the watershed with high and moderate treatment need for the target of erosion control.
FT: The FT scenario includes region in the watershed with high, moderate, and low treatment need for the target of erosion control.
NMC: The NMC scenario includes region in the watershed with high treatment need for the target of erosion control and nutrient management.
NMA: The NMA scenario includes region in the watershed with high and moderate treatment need for the target of erosion control and nutrient management.
NM: The NM scenario includes region in the watershed with high, moderate, and low treatment need for the target of erosion control and nutrient management.
GBG: The GBG scenario represents the assumption that all available croplands are forced to apply grassland practice with no exception. In this case, the expected cost could be the highest of all scenarios.
† Per-acre cost data for individual conservation practices were available for the 2010 crop year. The following four sources were used to compile baseline conservation cost estimates: 1. The NRCS payment schedule, 2. The NRCS conservation plans database, 3. The Technical Service Provider rate database, and 4. Average cropland rental rates.

Please cite this article as: Yen, H., et al., Western Lake Erie Basin: Soft-data-constrained, NHDPlus resolution watershed modeling and exploration of applicable conservation s..., Sci Total Environ (2016), http://dx.doi.org/10.1016/j.scitotenv.2016.06.202

2.7. Potential biological applications

The potential for conservation practices to improve stream biological conditions was explored by quantifying the amount of streams potentially degraded by excess nutrients in each conservation scenario. Excess nutrients are a pervasive threat to stream biodiversity (Richter et al., 1997; Dudgeon et al., 2006; Vörösmarty et al., 2010). For example, recent surveys suggest that excess nutrients are potentially degrading stream biological conditions in >40% of the rivers and streams in the United States (United States EPA, 2016). For the purposes of this study, a stream segment was classified as degraded if average annual total nitrogen (TN) or total phosphorus (TP) concentrations were above established nutrient criteria shown to increase primary production excessively and alter stream algae, aquatic insect, and fish communities (Evans-White et al., 2014). A stream segment's average concentration was calculated as the mean of median annual concentrations from the 21-year simulation for each conservation scenario. Outcomes for each conservation scenario were compared to the baseline scenario to assess the benefits provided by that scenario.

A number of nutrient criteria have been developed to assess stream health and while there is some agreement in the general magnitude of threshold values, the actual thresholds can vary by >6-fold (Evans-White et al., 2014). Nutrient criteria are particularly influenced by the statistical method (e.g., percentile approaches, 2DKS, regression tree), location (e.g., ecoregion), and biological response (e.g., algae, aquatic insects, or fish) used in their development (Evans-White et al., 2014). To account for this variability explicitly, 43 different nutrient criteria from a recent review of stream nutrient criteria in the United States (Table 4 in Evans-White et al., 2014) were used to assess potential stream degradation. These nutrient criteria included thresholds for both TN and TP based on empirical observations and were developed using a variety of statistical methods and biological responses (Evans-White et al., 2014).

A linear mixed model was used to assess the return on investment of different conservation practice types (erosion control only vs. erosion control and nutrient management). This model included the different nutrient criteria as random intercepts, and the additional cost of a conservation scenario, practice types included in a conservation scenario, and their interaction as fixed effects. The model used was:

$$y_i = \alpha_{j(i)} + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \epsilon_i$$

$$\alpha_j \sim N(\mu_\alpha, \sigma_\alpha^2)$$

where in the first level of the model y_i is the percent of the watershed improved in conservation scenario i using nutrient criteria j as a function of practice types included X_1, the amount of additional investment X_2, and their interaction X_3. In the second level of the model, nutrient criterion α_j was normally distributed with a mean of μ_α and residual variance of $\sigma^2{}_\alpha$ across all nutrient criteria. The effects of cost, practice types included, and their interactions were considered significant if the bootstrapped 95% confidence intervals of the estimates did not include zero (replicates = 10,000). All models were fit using the lme4 package (Bates et al., 2014) in the R statistical environment (R Core Team, 2008).

3. Results and discussion

3.1. Results of calibration/validation

As it was stated previously that watershed simulation of the NHDPlus project is an extremely time-consuming task. Therefore, a HUC-12 model was initiated and calibrated for streamflow in advance of the development of the NHDPlus scale model (Daggupati et al., 2015). One of the most important conclusions was made by Daggupati et al. (2015) was that some subbasins are similar to other adjacent ones in terms of geographical differences and also modeling behavior. Therefore, a total of five major subregions were categorized to conduct the following computationally expensive NHDPlus project. Five representative gauge stations (Daggupati et al., 2015) selected for watershed calibration and validation are shown in Fig. 1 and Table 4. According to the General Performance Ratings (Table 2), the daily streamflow calibration and validation can be considered between "Satisfactory" and "Good" (calibration: NSE: 0.54–0.87; PBIAS (%): −11.76–22.70; validation: NSE: 0.43–0.88; PBIAS (%): −26.07–25.25). In general, higher statistical standards are expected for streamflow calibration since flow is the fundamental medium of hydrologic, sediment, and nutrient processes in watersheds. However, streamflow calibration could be compensated for by sediment and nutrient mechanisms, especially in the NHDPlus resolution. It is extremely difficult to achieve "Very Good" simulation of streamflow processes while concurrently calibrating for other output variables to be within appropriate ranges in terms of statistical performance.

Statistical metrics for monthly sediment calibration and validation ranged from "Good" to "Very Good" (calibration: PBIAS (%): −10.43–18.63); validation: PBIAS (%): −35.01–35.46) It is important to have sediment processes well-calibrated because of the interaction between sediment and total phosphorus dynamics, due to sediment-associated phosphorus. Results for monthly total phosphorus calibration and validation were "Very Good" (calibration: PBIAS (%): −12.78–8.42; validation: PBIAS (%): 0.61 to −22.74). Model predictions for monthly total nitrogen ranged from "Satisfactory" (validation at St. Joseph) to "Very Good" (all other stations).

Model performance varied between individual gauging stations. The model provided better performance for the Maumee River gauge ("Very Good" results in all output variables) compared to the other stations (temporal results of streamflow, sediment, and nutrients are shown in Fig. 3). Among the five selected gauge stations, the Maumee River Station is the most representative outlet, since it receives 68.3% of the whole drainage area. In addition to hard data calibration, soft data used in these analyses, can have substantial impact on model predictions (Seibert and McDonnell, 2002; Yen et al., 2014b). The biological analyses in this study required simulation results which mimic realistic watershed responses. To simulate actual watershed behavior correctly, denitrification rate (DENI) and NO_3 contributed from tile flow (SSQ_Ratio) were constrained within certain ranges (DENI ≤ 50 kg/ha; SSQ_Ratio ≥ 0.6). The calibrated SWAT model had a DENI of 23.31 kg/ha and SSQ_Ratio of 0.72 (72% of NO_3 attributed to subsurface flow pathways) and both soft data outputs were reasonable (in quantity and in ratio) according to the available literature (Schilling, 2002; David et al., 2009).

As mentioned previously, the NHDPlus SWAT simulation of WLEB is exceptionally computationally expensive. It is not feasible to conduct thousands of model simulations during parameter estimation. Instead, expert judgement was used, especially for sediment and nutrient calibration. Therefore, necessary compensation (trade-off) has been made during the calibration process.

3.2. Applications of conservation scenarios

The conservation strategies simulated here explore potential conservation gains and associated financial costs that may be expected from the application of various suites of practices to various acres (Fig. 4, Table 3). The application of structural practices that provide erosion control without complementary nutrient management practices (ECC, ECA, and FT; Table 3) is much less expensive than is inclusion of structural practices and nutrient management. Application of structural practices on Critical Needs, All Needed, or all cropland acres, (ECC, ECA, and FT scenarios) could be implemented at 54, 37, and 49% of the respective costs required to supplement structural practices with nutrient management practices on these same acres (NMC, NMA, and NM, respectively).

ARTICLE IN PRESS

Table 4
Statistics of the calibration (1990–1999) and validation (2000–2006) periods.

Station	Streamflow		Sediment	TP	TN
	NSE	PBIAS (%)	PBIAS (%)	PBIAS (%)	PBIAS (%)
Raisin	0.70[a]/0.43[b]	−11.76/−26.07	16.71/35.46	−3.55/−22.74	14.66/3.59
St. Joseph	0.73/0.74	22.70/18.66	−10.43/−20.3	5.33/4.95	−25.39/−68.71
St. Marys	0.54/0.43	17.94/25.25	17.99/19.57	6.52/9.42	8.23/21.00
Maumee	0.87/0.88	18.03/13.56	10.07/−10.59	8.42/3.42	16.87/13.27
Sandusky	0.82/0.75	18.67/7.00	18.63/−35.01	−12.78/0.61	−15.26/−12.37

[a] Quantitative measures on the left-hand side represent the statistics for the calibration period.
[b] Quantitative measures on the right-hand side represent the statistics for the validation period.

The NM scenario suggests adoption of erosion control and nutrient management on all cropland acres could reduce annual loads of sediment being lost from croplands by up to 45%, total nitrogen by up to 41%, nitrate nitrogen by up to 32%, total phosphorus by up to 54%, and dissolved phosphorus by up to 28% (Fig. 4), but adoption of NM does not alleviate all of the sediment, nitrogen, or phosphorus conservation concerns in all of the streams. Even simulated elimination of agriculture from the landscape and simulated conversion all agricultural lands to vegetated easements (GBG scenario) did not reduce any nutrient or sediment losses by >80%, suggesting that other land uses currently contribute to sediment and nutrient loads (Fig. 4). Restoring stream health across the region will likely require efforts from all stakeholders and not just the agricultural community.

Maintaining current conservation and applying nutrient and structural erosion treatment to all agricultural acres in WLEB would cost around half a billion dollars every year. These funds are not currently available; therefore, the idea of prioritizing conservation spending by targeting Critical Needs acres for treatment before all other acres has been posited. However, an important and interesting finding of this work is that treating acres classified as Low Needs provides substantial

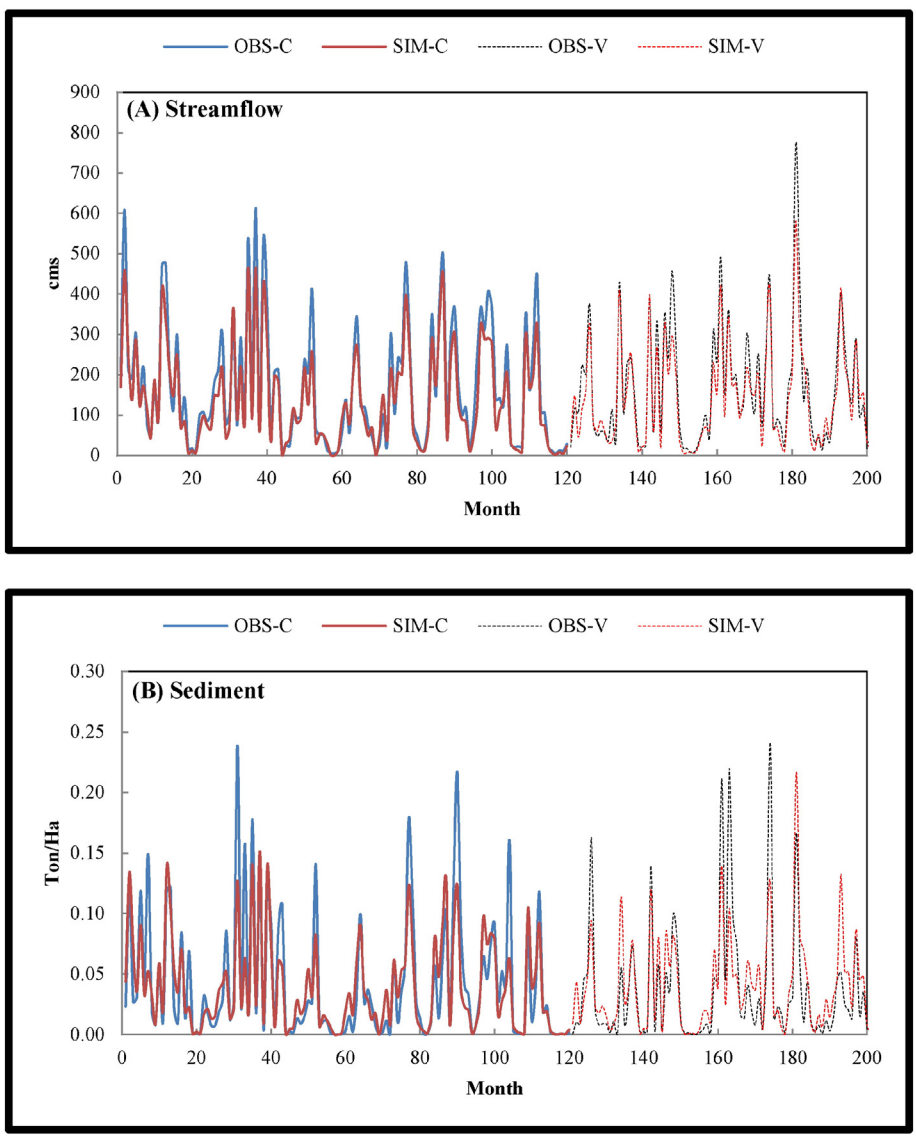


Fig. 3. Calibration and validation results of temporal processes of Maumee Station: (A) streamflow; (B) sediment; (C) total phosphorus; and (D) total nitrogen (OBS: observation data; SIM: simulation output; -C: calibration period; -V: validation period).

Please cite this article as: Yen, H., et al., Western Lake Erie Basin: Soft-data-constrained, NHDPlus resolution watershed modeling and exploration of applicable conservation s..., Sci Total Environ (2016), http://dx.doi.org/10.1016/j.scitotenv.2016.06.202

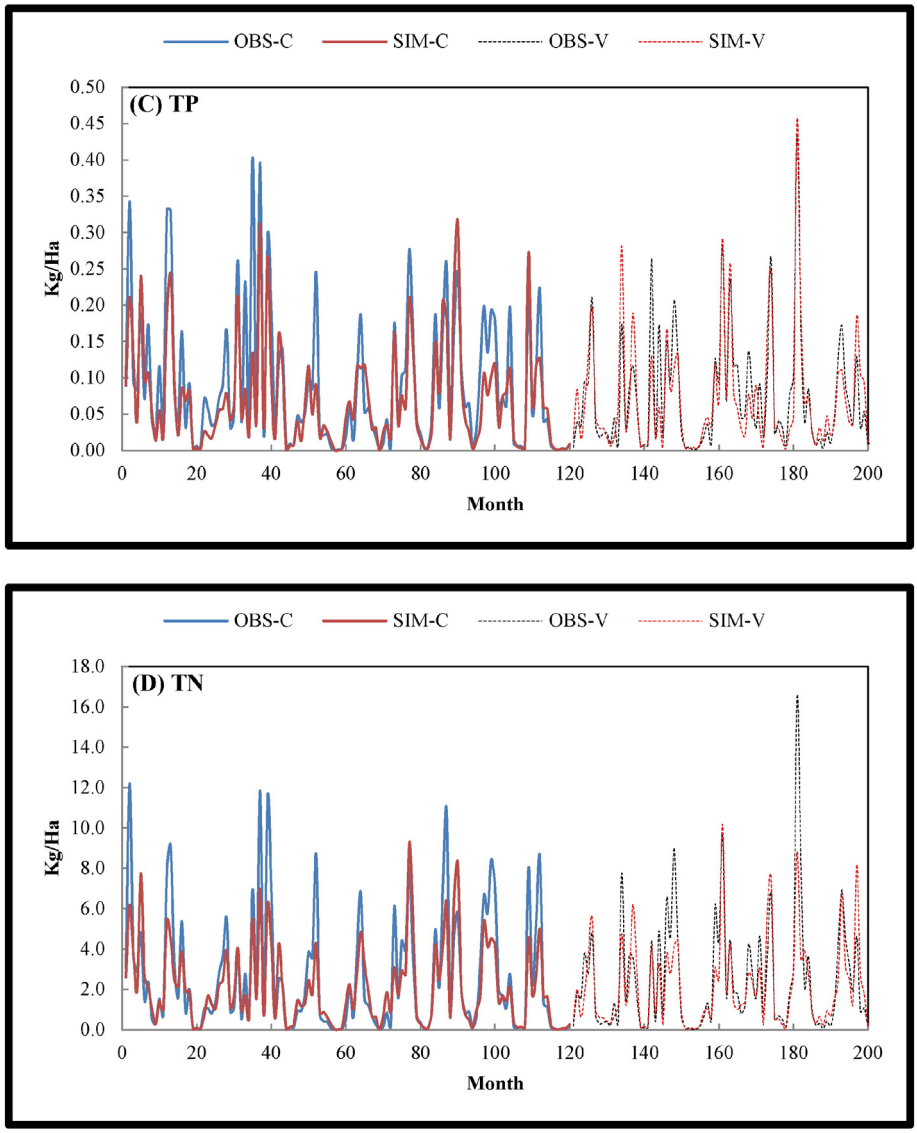


Fig. 3 (*continued*).

benefits in almost all cases, relative to treating only Critical Needs or Critical and Moderate Needs acres. These simulations suggest that an effective conservation strategy for cropland acres in WLEB should consider implementation of improved conservation practices on *all* cropland acres in the region, including those that have a low inherent vulnerability to losses and those that are already well treated.

It should be noted that the conservation scenarios explored here were developed to reduce edge-of-field losses, not to meet the needs of the streams or the ultimate receiving water body, Lake Erie. Further, these scenarios were not developed to represent the most cost effective policies to reduce nutrient and sediment loads to the streams or to Lake Erie. Land managers and communities must identify their conservation goals and develop comprehensive plans to achieve those goals; these plans will likely include both on-field and off-field conservation practices.

As expected, targeting treatment to critical needs acres provides the largest conservation return per dollar investment. Critical Needs acres can be treated with erosion control (ECC) at 4% of the investment required to treat all acres with erosion control (FT) and provides 13% of FT's annual sediment conservation benefits. All Needed acres can be treated with erosion control (ECA) at 43% of the cost of FT and will provide 56% of FT's annual sediment loads reduction benefits. Therefore, although low needs acres have the potential to provide nearly half of all possible sediment reduction benefits, there is a higher cost per unit of sediment reduction benefit on these acres than on Critical Needs and Moderate Needs acres.

Estimations of the annual reductions in sediment loads lost from agricultural acres in WLEB under alternative scenarios are shown in Fig. 4(A). As would be expected, structural erosion control practices have a significant impact on reduction of sediment loads. Increasing treatment to include nutrient management practices drastically increases the cost of treatment, but provides no benefit to sediment loss reductions (Table 3, Fig. 4). For example, all acres can be treated with erosion control practices (FT) at 49% of the cost of treating all acres with erosion control and nutrient management (NM), but both options reduce sediment loads being lost from agricultural acres by 45%, relative to baseline conditions.

Estimations of the annual reductions in total nitrogen loads being lost from agricultural acres in WLEB under alternative scenarios are shown in Fig. 4(B). Inclusion of nutrient management with structural erosion control practices had a much greater impact on total nitrogen

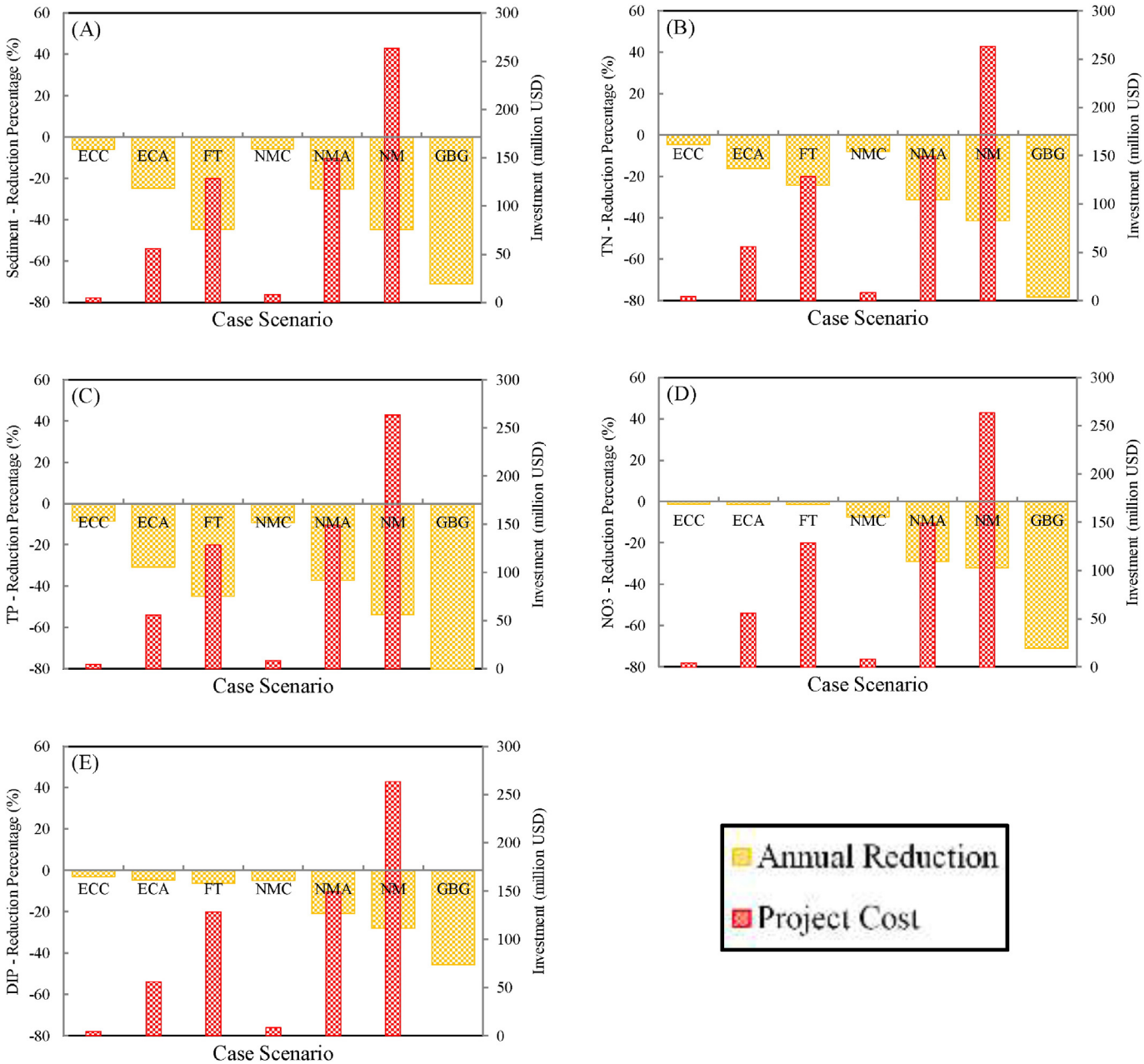


Fig. 4. Demonstration of percentage reduction of targeted output responses with alternative conservation scenarios in terms of projected additional cost ($ million, USD): (A) sediment; (B) total nitrogen; (C) total phosphorus; (D) NO_3; and (E) dissolved phosphorus (ECC: Erosion Control – Critical; ECA: Erosion Control – All Needed; FT: Full Treatment Erosion; NMC: Nutrient Management – Critical; NMA: Nutrient Management – All Needed; NM: Full Nutrient Management; GBG: Grass Background).

loss reduction than did structural practice adoption alone. The simulated benefits of enhanced nutrient management practices on annual nitrate reductions are dramatic, likely due to the fact that erosion control practices only address overland flow losses and nitrate may be lost to subsurface flow pathways (Fig. 4D). The application of structural practices to control erosion (ECC, ECA, and FT) achieves only 56, 51, and 59% of the potential annual total nitrogen load reduction benefits achieved in scenarios that apply structural practices and nutrient management practices (NMC, NMA, and NM) on Critical Needs, All Needed, or all cropland acres, respectively (Fig. 4B). However, ECC, ECA, and FT provide only 15, 5, and 4% of the nitrate load reduction benefits achieved in the NMC, NMA, and NM scenarios, respectively (Fig. 4D). Treating all acres with erosion control practices (FT; $128.3 million, annually) provides over three times the annual benefits in total nitrogen load reduction, but provides only 19% of the nitrate loss benefits relative to gains made with adoption of erosion control and nutrient management on Critical Needs acres only (NMC; $8.4 million, annually). Thus, nutrient management is clearly necessary if nitrate load reductions are the principal conservation goal. Investment decisions should carefully consider the resource concern of interest, as different management strategies are required to meet various conservation goals.

Estimations of the annual reductions in total phosphorus loads being lost from agricultural acres under alternative scenarios are shown in Fig. 4(C). Phosphorus, unlike nitrogen, is commonly bound to sediment. Therefore, erosion control practices tend to have more impact on total phosphorus losses than on total nitrogen losses. However, application of nutrient management in conjunction with structural erosion control practices has a much greater impact on total phosphorus loss reduction than does structural practice adoption alone. This effect is markedly

more dramatic when impacts on annual dissolved phosphorus load reductions are considered, likely due to the fact that erosion control practices only address overland flow losses, whereas dissolved phosphorus may also be lost to subsurface flow pathways (Fig. 4E).

The application of structural practices to control erosion (ECC, ECA, and FT) achieves 93, 83, and 84% of the annual total phosphorus load reduction benefits achieved in scenarios that apply structural practices in conjunction with nutrient management practices (NMC, NMA, and NM), respectively (Fig. 4C). However, ECC, ECA, and FT provide only 62, 23, and 22% of the potential dissolved phosphorus load reduction benefits that could be achieved with inclusion of nutrient management alongside erosion control (in NMC, NMA, and NM, respectively (Fig. 4E)). This suggests that on Critical Needs acres a significant amount of the dissolved phosphorus being lost could be conserved through structural practice adoption, whereas on the moderate and low needs acres nutrient management practices are imperative in order to achieve comparable reductions in dissolved phosphorus losses. Model simulations suggest that treatment of all acres with structural erosion control practices (FT; $128.3 million, annually) provides over 4.8 times the annual benefits

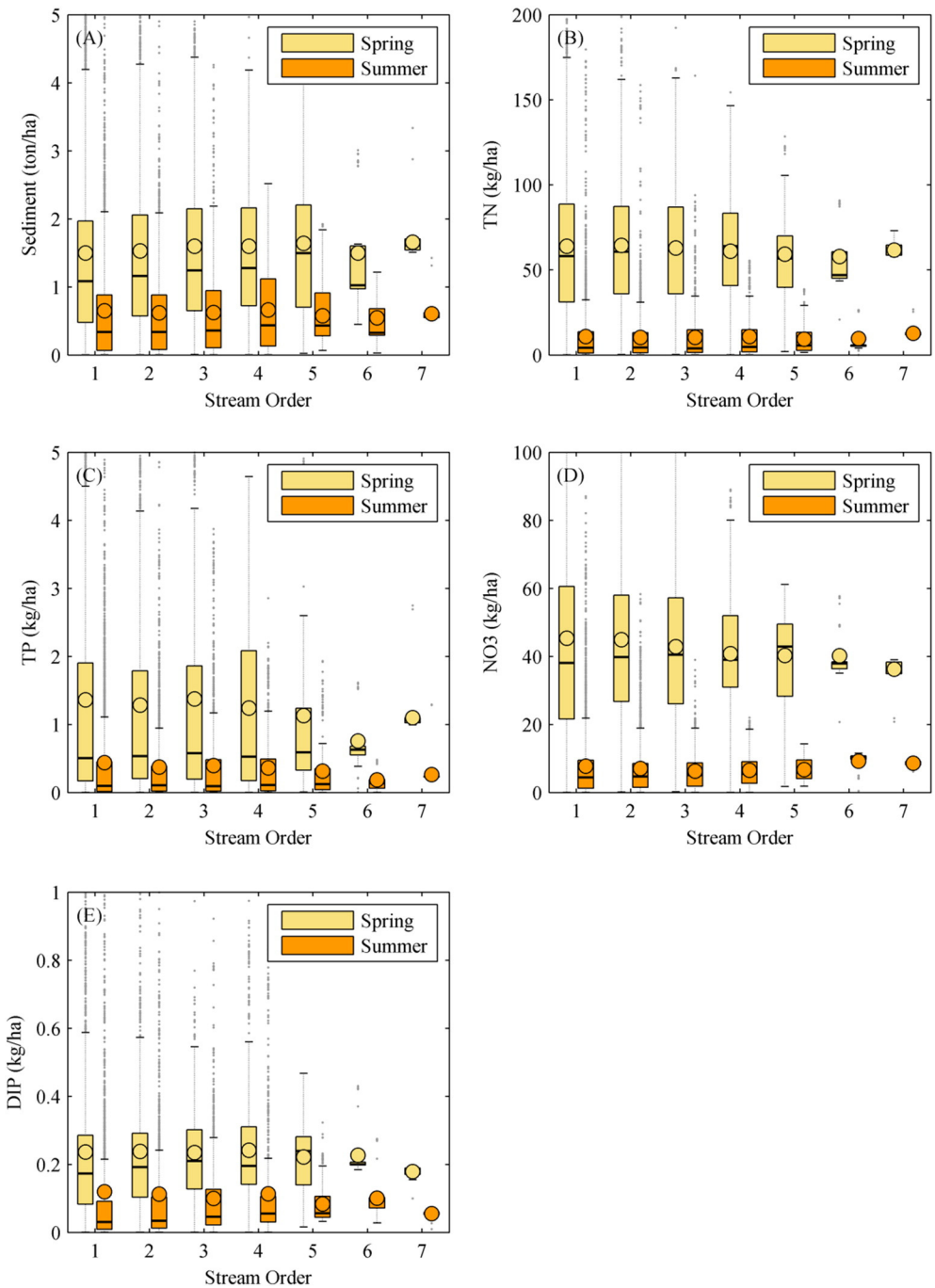


Fig. 5. Model responses for spring and summer seasons by different stream order (baseline scenario): (A) sediment; (B) total nitrogen; (C) total phosphorus; (D) NO_3; and (E) dissolved phosphorus.

Please cite this article as: Yen, H., et al., Western Lake Erie Basin: Soft-data-constrained, NHDPlus resolution watershed modeling and exploration of applicable conservation s..., Sci Total Environ (2016), http://dx.doi.org/10.1016/j.scitotenv.2016.06.202

in total phosphorus load reduction, but provides only 1.3 times the dissolved phosphorus loss benefits relative to gains that could be made with adoption of structural erosion control and nutrient management practices on Critical Needs acres only (NMC; $8.4 million, annually). As with nitrogen management decisions, phosphorus management strategies must take into account the importance of treating for total phosphorus reduction or targeting dissolved phosphorus reduction. The conservation goal informs the appropriate conservation strategy.

Beyond the general comparisons for each output variable, cross comparisons between total nitrogen and total phosphorus are also interesting. WLEB and most of the MWR are heavily tiled, so that most excess precipitation is transported through tile drainage systems. Rerouting water through tiles also reroutes nutrients to subsurface pathways. Tile drains are not only loss pathways for nitrate nitrogen and dissolved phosphorus, but have also been implicated as a loss pathway for sediment and sediment-associated phosphorus losses (Gaynord and Findlay, 1995; Molder et al., 2015). These results suggest that

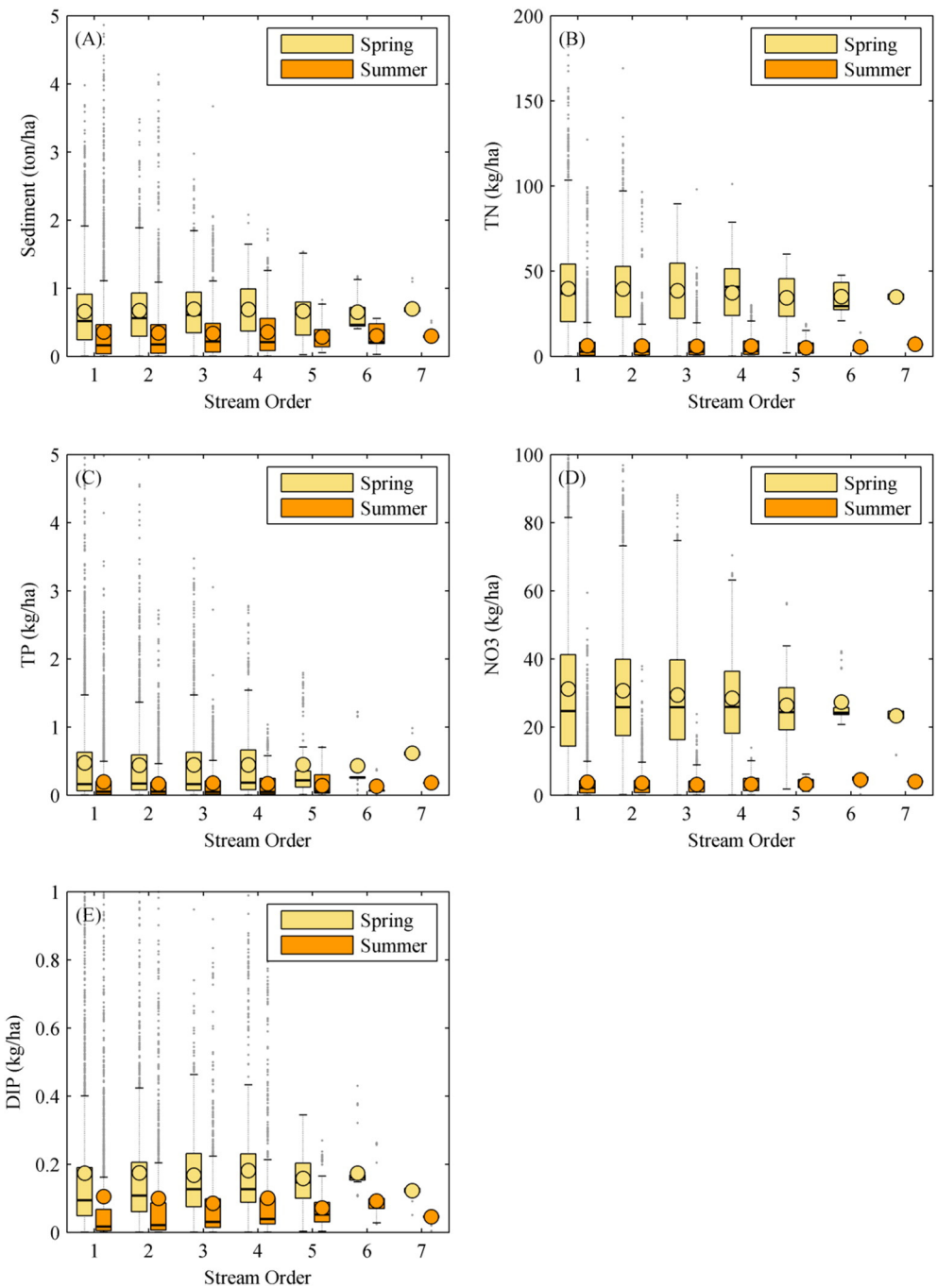


Fig. 6. Model responses for spring and summer seasons by different stream order (NM scenario): (A) sediment; (B) total nitrogen; (C) total phosphorus; (D) NO_3; and (E) dissolved phosphorus.

Please cite this article as: Yen, H., et al., Western Lake Erie Basin: Soft-data-constrained, NHDPlus resolution watershed modeling and exploration of applicable conservation s..., Sci Total Environ (2016), http://dx.doi.org/10.1016/j.scitotenv.2016.06.202

careful and comprehensive conservation planning that provides nutrient loss reductions in all loss pathways is needed to provide desired nitrogen and phosphorus loss reductions in this unique area (Fig. 4B and C).

3.3. Model responses enclosed in stream order levels

To examine additional information that can be obtained from the NHDPlus resolution SWAT model, model responses for spring and summer seasons by different stream order were examined (Figs. 5 and 6). Larger loads are predicted during spring as compared to summer, which is consistent with regional weather. Greater flows are expected in WLEB during spring, when snow melt occurs. Snow melt and spring rains may mobilize nutrients applied during spring planting or in the previous fall.

Modeling at this scale exposed the fact that rivers with lower stream orders have a higher degree of uncertainty. There are multiple reasons for this uncertainty. One reason is that calibration at this scale is difficult due to lack of calibration data; there is uncertainty associated with spatial interpretations when data are lacking. Low stream order streams and rivers are generally smaller in size, so may be less resilient to disturbance, such as flooding events and or anthropogenic disturbances. Therefore, simulated nutrient and sediment quantities in low order streams may vary substantially from reality, which estimates in higher order streams should be better calibrated and have less uncertainty.

Because of lower order stream vulnerabilities to perturbation, conservation practices that benefit lower order streams are likely to be beneficial to stream health in the region. These lower order streams feed into higher order streams, which would then also benefit from the conservation practices that reduced loads to the lower order streams. The biological indicators used in this work suggest that increased focus should be placed on conservation practices that target loss reduction in the spring season instead of the entire year (Keitzer et al., in press).

3.4. Potential reductions in degraded biological conditions

Anthropogenic disturbance, including nutrient enrichment is a serious threat to stream biodiversity (Richter et al., 1997; Dudgeon et al., 2006; Vörösmarty et al., 2010). Nutrient enrichment can cause excessive algal growth, alter algal community structure and food quality, and decrease dissolved oxygen levels, all of which can degrade stream biodiversity (Miltner and Rankin, 1998, Dodds, 2006; Wang et al., 2007; Evans-White et al., 2009; Miltner, 2010, Taylor et al., 2014). Understanding the effectiveness of conservation actions on mitigating nutrient enrichment of streams and rivers at the watershed-scale is therefore a key component of effective stream biodiversity conservation.

Averaged across all nutrient criteria, we estimate that 95% of WLEB streams have some level of nutrient enrichment that could degrade habitat and decrease biodiversity (about 19,256 km of streams). This includes nearly 73% of the streams in the watershed estimated to have annual nutrient levels above TP criteria, 93% above TN criteria, and nearly 68% of stream segments have both TP and TN concentrations above the nutrient criteria used for this study. These estimates suggest that while both are widespread in WLEB streams, streams are more likely to have excess nitrogen concentrations than excess phosphorus concentrations. However, high levels of both nutrients co-occur across a substantial portion of the watershed. Nutrient management strategies likely need to address both nitrogen and phosphorus losses through all potential loss pathways to be effective at providing benefits to stream biodiversity and health.

Our results suggest that the conservation practices simulated here are more effective at reducing stream degradation as the result of excess phosphorus than nitrogen (Fig. 7). We found that across all nutrient criteria and conservation scenarios, the percent of streams with improved water quality associated with scenario adoption ranged from <1% to 34% based on TP criteria and from <1% to 12% for TN criteria. Regardless of whether one or both nutrients are selected as the dominant conservation concern in the region, achieving large reductions will likely require

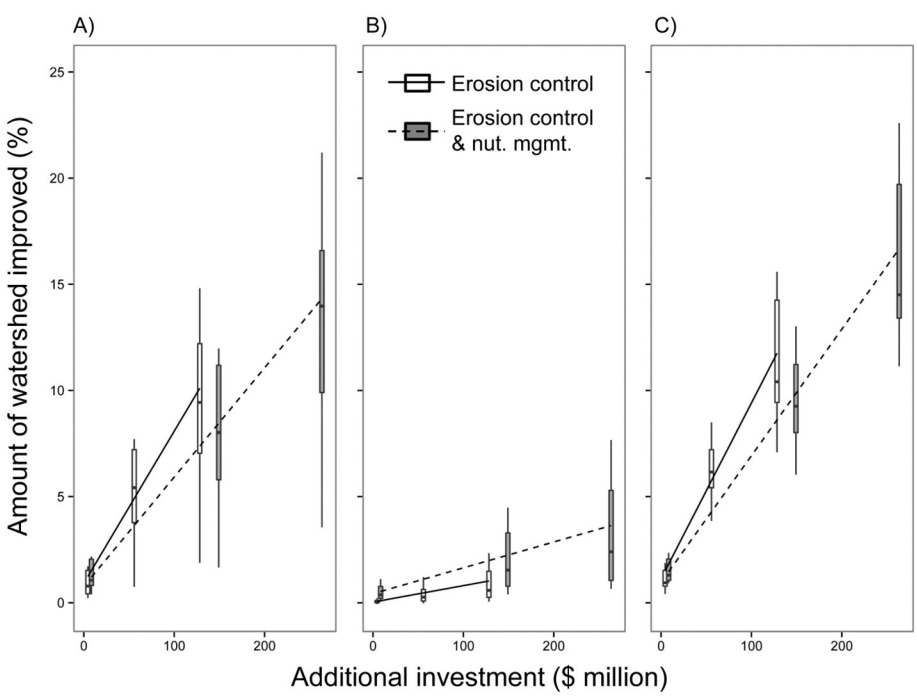


Fig. 7. Amount of the watershed in which stream biological conditions improved as a result of additional investment in conservation practices reducing nutrient enrichment: (A) improved percentage for streams that were above TP criteria; (B) improved percentage for streams that were above TN criteria; and (C) improved percentage for streams that were above both TP and TN. Boxplots represent the interquartile range of improvements across nutrient criteria with whiskers extending to the 10th and 90th percentiles. Lines show the predicted relationship between additional investment for conservation scenarios that included only erosion control practices (solid line) and scenarios that included erosion control and nutrient management practices (dashed line).

Please cite this article as: Yen, H., et al., Western Lake Erie Basin: Soft-data-constrained, NHDPlus resolution watershed modeling and exploration of applicable conservation s..., Sci Total Environ (2016), http://dx.doi.org/10.1016/j.scitotenv.2016.06.202

Table 5
Coefficient estimates (95% confidence interval) from multilevel models describing the effects of additional investment (Cost), including nutrient management (Nutrient management), and their interaction (Cost x nutrient management) on the amount of the watershed that is improved according to nutrient criteria for total phosphorus (TP) and total nitrogen (TN). Estimates were considered significant if bootstrapped (n = 10.000) confidence intervals did not include zero.

Nutrient	Intercept	Cost	Nutrient management	Cost × nutrient management
TP	0.91 (0.27,1.57)	0.072 (0.063, 0.079)	−0.15 (−1.07,0.78)	−0.02 (−0.029, −0.011)
TN	0.029 (−0.27,0.31)	0.008 (0.005, 0.01)	0.40 (0.03,0.80)	0.004 (0.0005, 0.008)
Both TN & NP	1.13 (0.56,1.68)	0.08 (0.075, 0.089)	−0.17 (−0.9376,0.6494)	−0.02 (−0.03, −0.015)

treating a substantial portion of the watershed. For instance, treating only Critical Needs acres with erosion control and nutrient management provides at most a 3% reduction in the percent of streams impacted by phosphorus; treating all agricultural acres in the region with structural controls and nutrient management can provide up to a 34% reduction in streams impacted by phosphorus. Similar results from a variety of watersheds suggest that conservation practices must be widely implemented to achieve the large reductions in nutrient and sediment losses from agricultural lands needed to improve stream biodiversity throughout the watershed (Einheuser et al., 2012; Bosch et al., 2013).

Our results also suggest that there are differences in potential return on investment of depending on the conservation concern being addressed (Table 5, Fig. 7). Erosion control practices alone represent a more cost effective strategy for reducing degradation by TP, while including nutrient management proves to be a more cost effective means for reducing degradation by TN (Table 5, Fig. 7). Because both nitrogen and phosphorus can negatively affect stream biodiversity (Miltner and Rankin, 1998; Wang et al., 2007; Evans-White et al., 2009), it is likely necessary that both must be managed appropriately if stream biodiversity is to be restored across the watershed. This effort will require development of comprehensive conservation plans and investments in both structural erosion control and nutrient management.

This work only considered the benefits of conservation investments on in-field and edge-of-field practices as they relate to nutrient losses. However, conservation practices can also reduce sediment loading to streams, reduce the frequency of extreme flows (e.g., stream drying), and improve in-stream habitat, all of which should benefit stream biodiversity (Wang et al., 2002, 2006). We may therefore have underestimated the potential benefits of in-field and edge-of-field conservation practices, particularly if multiple stressors interact to affect stream biodiversity, which appears to be common (Townsend et al., 2008; Matthaei et al., 2010; Wagenhoff et al., 2011, 2012). In-stream practices were not explored here and may be valuable complements to on-farm conservation practices.

While nutrient enrichment is a pervasive threat to stream biodiversity (Richter et al., 1997; Dudgeon et al., 2006; Vörösmarty et al., 2010; US EPA, 2015, 2016), it should be noted that it is not the only factor contributing to habitat degradation in the region. Among other stressors (e.g., invasive species), changes to in-stream habitat quality (e.g., siltation), flow regimes, and riparian vegetation are also prevalent concerns in the WLEB (Trautman, 1939; Trautman and Gartman, 1974; Karr et al., 1985). Conservation practices simulated here are limited to in-field and edge-of-field practices and their impacts on overland and subsurface water flows and associated nutrients. Other stream restoration techniques, including in-stream practices, dam removal, road improvements, riparian restoration and improved floodplains connectivity, were not considered here and may be essential to successful holistic agroecosystem planning and stream habitat improvement at the watershed scale. Results presented here should be considered with the caveat that the stressors and practices explored are limited in scope; these scenarios may under or overestimate the potential benefits of conservation practices in conjunction with other management across the landscape.

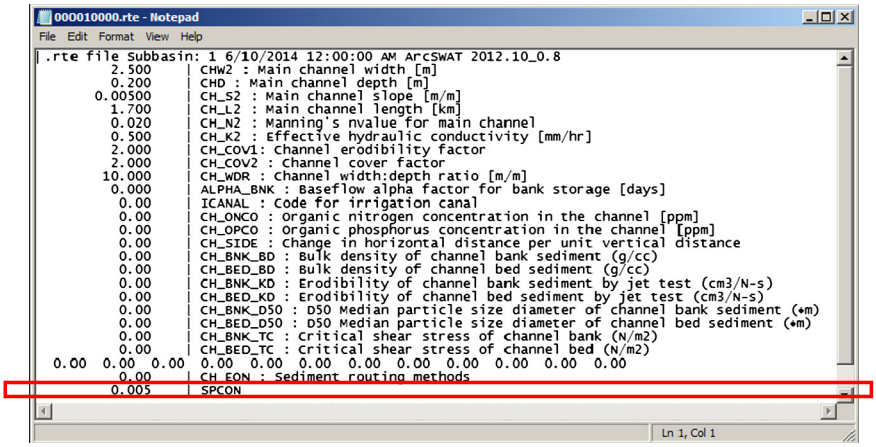
4. Conclusion

In this study, SWAT was applied at the NHDPlus resolution to explore the model's capacity to estimate potential environmental impacts of alternative conservation practices in terms of sediment and nutrient losses reduction with corresponding projected cost. An important finding is that SWAT is able to parse the varied impacts of conservation practices on total nitrogen, total phosphorus, nitrate, and dissolved phosphorus in response to the same conservation practice. However, without adequate monitored data, the fine-scale results cannot be appropriately validated. This work improves SWAT's capacity to serve as a decision support tool to determine conservation strategies associated with various conservation concerns and nutrient loss pathways. Structural erosion practices alone provide a sufficient remedy to sediment loss and could reduce current sediment loads lost from agricultural acres by up to 45%. The model also demonstrated that nutrient loss reduction strategies, especially for nitrate and dissolved phosphorus, benefit from the inclusion of nutrient management plans. In the WLEB, dissolved phosphorus is a critical concern (Daloglu et al., 2012; Baker et al., 2014). Therefore, realistically predicting dynamic responses to management can be valuable to researchers from various disciplines, as well as land use planners.

The work presented here is not just another piece of a calibration/validation project in watershed modeling. Instead, it is the first time in history that the most advanced technology of delineating stream networks at the fine resolution (NHDPlus) was implemented on a large-scale watershed like the WLEB. The goal of this study was to determine if the SWAT model was capable of identifying environmental impacts caused by agricultural activities, including various conservation strategies, at a refined spatial scale by applying the state-of-the-art techniques, including soft-data calibration. The model proved capable to simulating seemingly reasonable statistical results. While this manuscript and the model developed to support it provide scientists, engineers, and stakeholders with increased science-based information to augment current decision support tools for determining appropriate land management and policy development, it should be noted that there is insufficient spatially explicit monitoring data in the region to support appropriate validation of the model's spatially explicit results. It is true that the proposed finest resolution model may produce uncertainty and artificial errors because of the limitation. However, potential errors could cause more negative impacts on model predictions if the associated biological analysis was conducted in larger scale such as HUC-12 (average 72 km^2 in unit size which is >27 times larger than average unit area of NHDPlus). Implementations of finer resolution data are not only to scale down the individual unit of simulation, but it also can provide more information as indications that particular issues needed to improve in the future to enhance the accuracy of the modeling work.

Acknowledgements

This project was funded by grants from the United States Department of Agriculture - Natural Resources Conservation Service (USDA-NRCS) Conservation Effects Assessment Project (CEAP) - Wildlife and Cropland components (Cooperative Agreement: 68-7482-12-504). USDA is an equal opportunity employer and provider!



References

Ajami, N.K., Duan, Q., Sorooshian, S., 2007. An integrated hydrologic Bayesian multimodel combination framework: confronting input, parameter, and model structural uncertainty in hydrologic prediction. Water Resour. Res. 43 (1), 1–19.

Arnold, J.G., Allen, P.M., Bernhardt, G., 1993. A comprehensive surface-groundwater flow model. J. Hydrol. 142, 47–69.

Arnold, J., Moriasi, D., Gassman, P., Abbaspour, K., White, M., Srinivasan, Santhi, C., Harmel, R.D., van Griensven, A., Van Liew, M.W., Kannan, N., Jha, M.K., 2012. SWAT: model use, calibration, and validation. Trans. ASABE 55 (4), 1491–1508.

ASCE, 1993. Criteria for evaluation of watershed models, ASCE task committeeon definition of criteria for evaluation of watershed models of the watershed management, irrigation, and drainage division. J. Irrig. Drain. Eng. 119, 429.

Baker, N.T., 2011. Tillage Practices in the Conterminous United States, 1989–2004—Datasets Aggregated by Watershed: U.S. Geological Survey Data Series 573.

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 Lakes Res. 40, 502–517.

Bates, D., Maechler, M., Bolker, B., Walker, S., 2014. lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7.

Betanzo, E.A., Choquette, A.F., Reckhow, K.H., Hayes, L., Hagen, E.R., Argue, D.M., Cangelosi, A.A., 2015. Water Data to Answer Urgent Water Policy Questions: Monitoring Design, Available Data and Filling Data Gaps for Determining the Effectiveness of Agricultural Management Practices for Reducing Tributary Nutrient Loads to Lake Erie, Northeast-Midwest Institute Report. (169 pp., http://www.nemw.org?).

Bicknell, B.R., Imhoff, J.C., Kittle Jr., J.L., Donigian, A.S., Johanson, R.C., 1997. Hydrological Simulation Program–Fortran: User's Manual for Version 11. U.S. Environmental Protection Agency, National Exposure Research Laboratory, Athens, GA, EPA/600/R-97/080 (755 pp.).

Boardman, A.E., Greenberg, D.H., Vining, A.R., Weimer, D.L., 2001. Cost-Benefit Analysis. Concepts and Practice. second ed. Prentice Hall, Upper Saddle River, NJ.

Bosch, N.S., Allan, J.D., Dolan, D.M., Han, H., Richards, R.P., 2011. Application of the soil and water assessment tool for six watersheds of Lake Erie: model parameterization and calibration. J. Great Lakes Res. 37, 263–271.

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 Lakes Res. 39, 429–436.

Brooks, B., Lazorchak, J.M., Howard, M.D.A., Johnson, M.-V.V., Morton, S.L., Perkins, D.A.K., Reavie, E.D., Scott, G.I., Smith, S.A., Steevens, J.A., 2016. Are harmful algal blooms becoming the greatest inland water quality threat to public health and aquatic systems? Environ. Toxicol. Chem. 35, 6–13.

Carpenter, S.R., Caraco, N.F., Correll, D.L., Howarth, R.W., Sharpley, A.N., Smith, V.H., 1998. Nonpoint pollution of surface waters with phosphorus and nitrogen. Ecol. Appl. 8, 559–568.

Chaubey, I., Cotter, A.S., Costello, T.A., Soerens, T.S., 2005. Effect of DEM data resolution on SWAT output uncertainty. Hydrol. Process. 19, 621–628. http://dx.doi.org/10.1002/hyp.5607.

Chou, F.N.-F., Wu, C.-W., 2014. Determination of cost coefficients of a priority-based water allocation linear programming model – a network flow approach. Hydrol. Earth Syst. Sci. 18, 1857–1872. http://dx.doi.org/10.5194/hess-18-1857-2014.

Chu, T.W., Shirmohammadi, A., Montas, H., Sadeghi, A., 2004. Evaluation of the SWAT Model's sediment and nutrient components in the piedmont physiographic region of Maryland. Trans. ASABE 47 (5), 1523–1538.

Cotter, A.S., Chaubey, I., Costello, T.A., Soerens, T.S., Nelson, M.A., 2003. Water quality model output uncertainty as affected by spatial resolution of input data. J. Am. Water Resour. Assoc. 39, 977–986. http://dx.doi.org/10.1111/j.1752-1688.2003.tb04420.x.

Daggupati, P., Douglas-Mankin, K.R., Sheshukov, A.Y., Barnes, P.L., Devlin, D.L., 2011. Field-level targeting using SWAT: mapping output from HRUs to fields and assessing limitations of GIS input data. Trans. ASABE 54, 501–514.

Daggupati, P., Yen, H., White, M.J., Srinivasan, R., Arnold, J.G., Keitzer, C.S., Sowa, S.P., 2015. Impact of model development, calibration and validation decisions on hydrological simulations in West Lake Erie Basin. Hydrol. Process. http://dx.doi.org/10.1002/hyp.10536.

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

David, M.B., Del Grosso, S.J., Hu, X., Marshall, E.P., McIsaac, G.F., Parton, W.J., Tonitto, C., Youssef, M., 2009. Modeling denitrification in a tile-drained, corn and soybean agroecosystem of Illinois, USA. Biogeochemistry 93, 7–30. http://dx.doi.org/10.1007/s10533-008-9273-9.

Di Toro, D.M., O'Connor, D.J., Mancini, J.L., Thomann, R.V., 1973. A preliminary phytoplankton-zooplankton-nutrient model of western Lake Erie. Systems Analysis & Simulation in Ecology Vol. 3. Academic Press.

Dodds, W.K., 2006. Eutrophication and trophic state in rivers and streams. Limnol. Oceanogr. 51, 671–680.

Duan, Q., Sorooshian, S., Gupta, V., 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. Water Resour. Res. 28 (4), 1015–1031.

Dudgeon, D., Arthington, A.H., Gessner, M.O., Kawabata, Z.I., Knowler, D.J., Leveque, C., Naiman, R.J., Prieur-Richard, A.H., Soto, D., Stiassny, M.L.J., Sullivan, C.A., 2006. Freshwater biodiversity: importance, threats and conservation challenges. Biol. Rev. 81, 163–182.

Dungjen, T., Patch, D., 2014. Toledo-area water advisory expected to continue through Sunday as leaders await tests. The Blade. Published on Aug. 2, 2014. (Available online): http://www.toledoblade.com/local/2014/08/02/City-of-Toledo-issues-do-no-drink-water-advisery.html.

Efstratiadis, A., Koutsoyiannis, D., 2010. One decade of multi-objective calibration approaches in hydrological modeling: a review. Hydrol. Sci. J. 55 (1), 58–78.

Einheuser, M.D., Nejadhashemi, A.P., Sowa, S.P., Wang, L., Hamaamin, Y.A., Woznicki, S.A., 2012. Modeling the effects of conservation practices on stream health. Sci. Total Environ. 435-436, 380–391.

Evans-White, M.A., Dodds, W.K., Huggins, D.G., Baker, D.S., 2009. Thresholds in macroinvertebrate biodiversity and stoichiometry across water-quality gradients in Central Plains (USA) streams. J. N. Am. Benthol. Soc. 28, 855–868.

Evans-White, M.A., Haggard, B.E., Scott, J.T., 2014. A review of stream nutrient criteria development in the United States. J. Environ. Qual. 42, 1002–1014.

Foster, G.R., 2005. Revised Universal Soil Loss Equation, Version 2.0 (RUSLE2 2.0): Science Documentation. USDA-Agricultural Research Service, Washington, DC, p. 286.

Gassman, P.W., Reyes, M.R., Green, C.H., Arnold, J.G., 2007. The soil and water assessment tool: historical development, applications and future research directions. Trans. Am. Soc. Agric. Biol. Eng. 50 (4), 1211–1250.

Gaynord, J.D., Findlay, W.I., 1995. Soil and phosphorus loss from conservation and conventional tillage in corn production. J. Environ. Qual. 24, 734–741.

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. Softw. 71, 121–134.

Haan, C.T., Allred, B., Storm, D.E., Sabbagh, G.J., Prabhu, S., 1995. Statistical procedure for evaluating hydrologic/water quality models. Trans. ASAE 38 (3), 725–733.

Jha, M.K., Gassman, P.W., 2014. Changes in hydrology and streamflow as predicted by a modelling experiment forced with climate models. Hydrol. Process. 28, 2772–2781.

Johnson, M.V.V., Norfleet, M.L., Atwood, J.D., Behrman, K.D., Kiniry, J.R., Arnold, J.G., White, M.J., Williams, J., 2015. The Conservation Effects Assessment Project (CEAP): A National Scale Natural Resources and Conservation Needs Assessment and Decision Support Tool. IOP Conference Series: Earth and Environmental Science Vol. 25. http://dx.doi.org/10.1088/1755-1315/25/1/012012.

Kaatz, M.R., 1955. The black swamp: a study in historical geography. Ann. Assoc. Am. Geogr. 45 (1), 1–35. http://dx.doi.org/10.1111/j.1467-8306.1955.tb01481.x.

Kane, D.D., Conroy, J.D., Richards, P.R., Baker, D.B., Culver, D.A., 2014. Re-eutrophication of Lake Erie: correlations between tributary nutrient loads and phytoplankton biomass. J. Great Lakes Res. 40, 496–501.

Karr, J.R., Toth, L.A., Dudley, D.R., 1985. Fish communities of Midwestern rivers: a history of degradation. BioScience 35, 90–95.

Keitzer, S., Ludsin, S.A., Sowa, S., Annis, G., Daggupati, P., Froelich, A.M., Herbert, M.E., Johnson, M.V., Yen, H., White, M.J., Arnold, J.G., Sasson, A.M., Rewa, C.A., 2016.

Thinking outside of the lake: can controls on nutrient inputs into Lake Erie benefit stream conservation in its watershed? J. Great Lakes Res. http://dx.doi.org/10.1016/j.jglr.2016.05.012 (Available online).
Matthaei, C.D., Piggott, J.J., Townsend, C.R., 2010. Multiple stressors in agricultural streams: interactions among sediment addition, nutrient enrichment and water abstraction. J. Appl. Ecol. 47, 639–649.
Mausbach, M.J., Dedrick, A.R., 2004. The length we go: measuring environmental benefits of conservation practices. J. Soil Water Conserv. 59 (5), 96–103.
Michalak, A.M., Anderson, E.J., Beletsky, D., Boland, S., Bosch, N.S., Bridgemen, 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, P.R., Scavia, D., Steiner, A.L., Verhamme, E., Wrigtht, 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.
Miltner, R.J., 2010. A method and rationale for deriving nutrient criteria for small rivers and streams in Ohio. Environ. Manag. 45, 842–855.
Miltner, R.J., Rankin, E.T., 1998. Primary nutrients and the biotic integrity of rivers and streams. Freshw. Biol. 40, 145–158.
Mitsch, W.J., Gosselink, J.G., 2007. Wetlands. John Wiley and Sons, pp. 65–66 (ISBN 978-0-471-69967-5).
Molder, B., Cockburn, J., Berg, A., Lindsay, J., Woodrow, K., 2015. Sediment-assisted nutrient transfer from a small, no-till tile drained watershed in Southwestern Ontario, Canada. Agric. Water Manag. 152, 31–40.
Moriasi, D.N., Arnold, J.G., Liew, M.W.V., Bingner, R.L., Harmel, R.D., Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. ASABE 50 (3), 885–900.
Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models: part I. A discussion of principles. J. Hydrol. 10 (3), 282–290.
Neitsch, S.L., Arnold, J.G., Kiniry, J.R., R, W. J., 2011. Soil & Water Assessment Tool Theoretical Documentation Version 2009. Texas Water Resource Institute Technical Report No. 406. Texas A & M University System.
R Development Core Team, 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria http://www.R-project.org.
Richter, B.D., Braun, D.P., Mendelson, M.A., Master, L.L., 1997. Threats to imperiled freshwater fuana. Conserv. Biol. 11, 1081–1093.
Robertson, D.M., Saad, D.A., 2011. Nutrient Inputs to the Laurentian Great Lakes by Source and Watershed Estimated Using SPARROW Watershed Models. Journal of the American Water Resources Association 47 (5), 1011–1033.
Runkel, R., Crawford, C., Cohn, T., 2004. Load Estimator (LOADEST): A Fortran Program for Estimating Constituent Loads in Streams and Rivers 2004. US Geological Survey Techniques and Methods Book Vol. 4.
Santhi, C., Arnold, J.G., Williams, J.R., Dugas, W.A., Srinivasan, R., Hauck, L.M., 2001. Validation of the SWAT model on a large River Basin with point and nonpoint sources. J. Am. Water Resour. Assoc. 37, 1169–1188.
Scavia, D., Kalcic, M., Muenich, R.L., Aloysius, N., Boles, C., Confessor, R., DePinto, J., Gildow, M., Martin, J., Read, J., Redder, T., Sowa, S., Wang, Y., Yen, H., 2016. Shaping Lake Erie Agriculture Nutrient Management Via Scenario Development. University of Michigan, Ann Arborhttp://tinyurl.com/pp4umuz.
Schilling, K.E., 2002. Chemical transport from paired agricultural and restored prairie watersheds. J. Environ. Qual. 31, 1184–1193. http://dx.doi.org/10.2134/jeq2002.1184.
Schwarz, G.E., Hoos, A.B., Alexander, R.B., Smith, R.A., 2006. The SPARROW Surface Water-Quality Model—Theory, Applications and User Documentation. U.S. Geological Survey, Techniques and Methods 6-B3 http://pubs.usgs.gov/tm/2006/tm6b3, accessed October 15, 2012.
Seibert, J., McDonnell, J.J., 2002. On the dialog between experimentalist and modeler in catchment hydrology: use of soft data for multicriteria model calibration. Water Resour. Res. 38 (11), 1241. http://dx.doi.org/10.1029/2001WR000978.
Servat, E., Dezetter, A., 1991. Selection of calibration objective functions in the context of rainfall-runoff modeling in a Sudanese savannah area. Hydrol. Sci. J. 36, 307–330.
Smith, V.H., Tilman, G.D., Nekola, J.C., 1999. Eutrophication: impacts of excess nutrient inputs on freshwater, marine, and terrestrial ecosystems. Environ. Pollut. 100, 179–196.
Srinivasan, R., Zhang, X., Arnold, J., 2010. SWAT ungauged: hydrological budget and crop yield predictions in the upper Mississippi River basin. Trans. ASABE 53, 1533–1546.
SWAT Literature Database Website, d. https://www.card.iastate.edu/swat_articles/ (Last accessed 4/2/2014, 4:58 pm).
Taylor, J.M., King, R.S., Pease, A.A., Winemiller, K.O., 2014. Nonlinear response of stream ecosystem structure to low-level phosphorus enrichment. Freshw. Biol. 59, 969–984.
Townsend, C.R., Uhlmann, S.S., Matthaei, C.D., 2008. Individual and combined responses of stream ecosystems to multiple stressors. J. Appl. Ecol. 45, 1810–1819.
Trautman, M.B., 1939. The effects of man-made modifications on the fish fauna in Lost and Gordon Creeks, Ohio, between 1887-1938. The Ohio Journal of Science 39, 275–288.
Trautman, M.B., Gartman, D.K., 1974. Re-evaluation of the effects of man-made modifications on Gordon Creek between 1887 and 1973 and especially as regards its fish fauna. The Ohio Journal of Science 74, 162–173.
USDA - Natural Resources Conservation Service (NRCS), 1995. Soil Survey Geographic (SSURGO) Database - Data Use Information. Misc. Publication No. 1527. National Soil Survey Center, Lincoln, NE.
USDA - Natural Resources Conservation Service (NRCS), 2011. Assessment of the Effects of Conservation Practices on Cultivated Cropland in the Great Lakes Region, Conservation Effects Assessment Project (CEAP). http://www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/stelprdb1045480.pdf.
USDA - Natural Resources Conservation Service (NRCS), 2012. FY2013 Practice Payment Schedule for EQIP/WHIP, Guidance Document. http://www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/nrcs141p2_035967.pdf.
USDA - Natural Resources Conservation Service (NRCS), 2014. Western Lake Erie Basin (WLEB) CEAP Project, Semiannual Progress Report.
USDA - Natural Resources Conservation Service (NRCS), 2015. Western Lake Erie Basin (WLEB) CEAP project, Semiannual Progress Report.
USDA - Natural Resources Conservation Services (NRCS), 2016. Effects of conservation practice adoption on cultivated cropland acres in Western Lake Erie Basin, 2003-06 and 2012. (120 pp.) http://www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/nrcseprd889806.pdf.
USDA-National Agricultural Statistics Service (NASS), 2011. Annual report of cash rents per acre. http://www.nass.usda.gov/Statistics_by_State/Ohio/Publications/Reports_by_Title/cshrnt11.pdf (last available Jan 11, 2016).
USDA-National Agricultural Statistics Service (NASS), 2014. http://www.nass.usda.gov.
USEPA, 2006. Wadeable Streams Assessment. USEPA-841-B-06-002. USEPA, Washington, D.C.
USEPA, 2015. https://epahawqs.tamu.edu/projects.
USEPA, 2016. National Rivers and Stream Assessment 2008–2009: A Collaborative Survey. USEPA-841-R-I6-007. USEPA, Washington, D.C.
Vörösmarty, C.J., McIntyre, P.B., Gessner, M.O., Dudgeon, D., Prusevich, A., Green, P., Glidden, S., Bunn, S.E., Sullivan, C.A., Liermann, C.R., Davies, P.M., 2010. Global threats to human water security and river biodiversity. Nature 467, 555–561.
Wagenhoff, A., Townsend, C.R., Phillips, N., Matthaei, C.D., 2011. Subsidy-stress and multiple-stressor effects along gradients of deposited fine sediment and dissolved nutrients in a regional set of streams and rivers. Freshw. Biol. 56, 1916–1936.
Wagenhoff, A., Townsend, C.R., Matthaei, C.D., 2012. Macroinvertebrate responses along broad stressor gradients of deposited fine sediment and dissolved nutrients: a stream mesocosm experiment. J. Appl. Ecol. 49, 892–902.
Wang, L., Lyons, J., Kanehl, P., 2002. Effects of watershed best management practices on habitat and fish in Wisconsin streams. J. Am. Water Resour. Assoc. 38, 663–680.
Wang, L., Lyons, J., Kanehl, P., 2006. Habitat and fish responses to multiple agricultural best management practices in a warm water stream. J. Am. Water Resour. Assoc. 42, 1047–1062.
Wang, L., Robertson, D.M., Garrison, P.J., 2007. Linkages between nutrients and assemblages of macroinvertebrates and fish in wadeable streams: implication to nutrient criteria development. Environ. Manag. 39, 194–212.
White, K.L., Chaubey, I., 2005. Sensitivity analysis, calibration, and validations for a multisite and multivariable SWAT model. J. Am. Water Resour. Assoc. 41 (5), 1077–1089.
White, M.J., Santhi, C., Kannan, N., Arnold, J.G., Harmel, R.D., Norfleet, L., Allen, P., DiLuzio, M., Wang, X., Atwood, J.D., Haney, E., Johnson, M., 2014. Nutrient delivery from the Mississippi River to the Gulf of Mexico and effects of cropland conservation. J. Soil Water Conserv. 69 (1), 26–40.
White, M., Harmel, R.D., Yen, H., Arnold, J.G., Gambone, M., Haney, R., 2015. Development of sediment and nutrient export coefficients for U.S. ecoregions. J. Am. Water Resour. Assoc. 51 (3), 758–775. http://dx.doi.org/10.1111/jawr.12270.
White, M.J., Gambone, M., Yen, H., Arnold, J.G., Harmel, R.D., Santhi, C., Haney, R.L., 2015. "Regional Blue and Green Water Balances and Use by Selected Crops in the U.S.". Journal of the American Water Resources Association (JAWRA) 51 (6), 1626–1642. http://dx.doi.org/10.1111/1752-1688.12344.
Williams, J.W., Izaurralde, R.C., Steglich, E.M., 2012. Agricultural Policy/Environmental EXtender Model Theoretical Documentation Version 0806 (131 pp.).
Yen, H., Jeong, J., Tseng, W., Kim, M., Records, R., Arabi, M., 2014a. Computational procedure for evaluating sampling techniques on watershed model calibration. J. Hydrol. Eng. 20 (7), 04014080. http://dx.doi.org/10.1061/(ASCE)HE.1943-5584.0001095.
Yen, H., Bailey, R.T., Arabi, M., Ahmadi, M., White, M.J., Arnold, J.G., 2014b. The role of interior watershed processes in improving parameter estimation and performance of watershed models. Journal of Environmental Quality 43 (5), 1601–1613. http://dx.doi.org/10.2134/jeq2013.03.0110.
Yen, H., Wang, X., Fontane, D.G., Harmel, R.D., Arabi, M., 2014c. A framework for propagation of uncertainty contributed by parameterization, input data, model structure, and calibration/validation data in watershed modeling. Environ. Model. Softw. 54, 211–221. http://dx.doi.org/10.1016/j.envsoft.2014.01.004.
Yen, H., Daggupati, P., White, M.J., Srinivasan, R., Gossel, A., Wells, D., Arnold, J.G., 2016a. Application of Large-Scale, Multi-resolution Watershed Modeling Framework Using the Hydrologic and Water Quality System (HAWQS). Water 8 (4), 164 0.3390/w8040164.
Yen, H., White, M.J., Ascough III, J.C., Smith, D.R., Arnold, J.G., 2016b. Augmenting watershed model calibration with incorporation of ancillary data sources and qualitative soft data sources. J. Am. Water Resour. Assoc. 52 (3), 788–798. http://dx.doi.org/10.1111/1752-1688.12428.

Important Corrections of STOTEN20332
Ms. Ref. No.: STOTEN-D-16-02377R2

1. Abstract: Please replace "(SWAT)" to "(SWAT2012)"
2. Abstract: Limitations to the model's predictive capacity were due to a paucity of data at the NHDPlus scale (1-2 km2) rather than due to SWAT functionality.
3. Introduction: Replace "since that requires mutual work…" to "because that requires mutual work…".
4. Sec. 2.2: Replace "accurate stream flow predictions" to "accurate streamflow predictions".
5. Sec. 2.4: Replace "Since spatial data" to "Because spatial data".
6. Sec. 2.5: Replace "phosphorous" to "phosphorus".
7. Sec. 2.5: Replace "This report and the SWAT…" to "This study and the SWAT…".
8. Sec. 2.6: Replace "(NRCS-2011)" to "(USDA-NRCS, 2011)".
9. Sec. 3.1: Replace "As it was stated previously that" to "As it was stated previously, that".
10. Sec. 3.1: Replace "made by Daggupati et al. (2015) was" to "made by Daggupati et al. (2015)".
11. Sec. 3.1: Replace "streamflow calibration since" to "streamflow calibration because".
12. Sec. 3.1: Replace ", since it receives" to ", because it receives".
13. Conclusion: In addition, without monitored data at the NHDPlus scale, the fine-scale results cannot be appropriately (thoroughly) validated.
14. Acknowledgments: Right before "USDA is an equal opportunity employer and provider", please also add "Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government".
15. Table 1: Change "Heidelberg College" to "Heidelberg University".
16. Table 1: Change "United States Geologic Survey" to "U.S. Geological Survey".
17. Table 1: The original sources of the streamflow data used in the Heidelberg University River Studies were from USGS.
18. Figure 5 & 6: Boxplots represent the interquartile range with whiskers extending to the 10th and 90th percentiles, with the center of the box representing mean values.
19. Figure 7: Change "Boxplots represent..." to "Boxplots (lines in the center are the mean values) represent…"
20. Reference: Scavia, D., M. Kalcic, R. L. Muenich, N. Aloysius, J. Arnold, C. Boles, R. Confessor, J. DePinto, M. Gildow, J. Martin, J. Read, T. Redder, D. Robertson, S. Sowa, Y. Wang, M. White, H. Yen, 2016. Informing Lake Erie agriculture nutrient management via scenario evaluation. University of Michigan, Ann Arbor. http://graham.umich.edu/water/project/erie-western-basin
21. Acknowledgements: This project was funded by grants from the United States Department of Agriculture - Natural Resources Conservation Service (USDA-NRCS) Conservation Effects Assessment Project (CEAP) - Wildlife and

Cropland components. Any use of trade, firm, or product names for descriptive purposes only and does not imply endorsement by the U.S. Government. USDA is an equal opportunity employer and provider!

APPENDIX C: KEITZER ET AL. 2016. JOURNAL OF GREAT LAKES RESEARCH. HTTP://DX.DOI.ORG/10.1016/J.JGLR.2016.05.012.

ARTICLE IN PRESS

JGLR-01087; No. of pages: 10; 4C:

Journal of Great Lakes Research xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Journal of Great Lakes Research

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

Thinking outside of the lake: Can controls on nutrient inputs into Lake Erie benefit stream conservation in its watershed?

S. Conor Keitzer [a,*], Stuart A. Ludsin [a], Scott P. Sowa [b], Gust Annis [b], Jeff G. Arnold [c], Prasad Daggupati [d], August M. Froehlich [e], Matt E. Herbert [b], Mari-Vaughn V. Johnson [f], Anthony M. Sasson [e], Haw Yen [g], Mike J. White [c], Charles A. Rewa [h]

[a] Aquatic Ecology Laboratory, Department of Evolution, Ecology, and Organismal Biology, The Ohio State University, 1314 Kinnear Rd., Columbus, OH 43212, USA
[b] The Nature Conservancy Michigan Field Office, 101 East Grand River Ave., Lansing, MI 48906, USA
[c] Grassland Soils and Water Research Laboratory, USDA-ARS, 808 East Blackland Rd., Temple, TX 76502, USA
[d] Texas A&M AgriLife Research, Texas A&M University, 221B Spatial Science Laboratory, 1500 Research Plaza, College Station, TX 77943, USA
[e] The Nature Conservancy Ohio Field Office, 6375 Riverside Drive, Suite 100, Dublin, OH 43017, USA
[f] Soil Science and Resource Assessment Division, USDA-NRCS, 720 East Blackland Rd., Temple, TX 76502, USA
[g] Blackland Research and Extension Center, Texas A&M University, 720 East Blackland Rd., Temple, TX 76502, USA
[h] Resource Assessment Division, USDA-NRCS, 6501 Sunnyside Ave., Beltsville, MD 20705, USA

ARTICLE INFO

Article history:
Received 14 December 2015
Accepted 3 May 2016
Available online xxxx

Communicated by Joseph DePinto

Index words:
Best management practices
SWAT
Non-point source pollution
Great Lakes
Ecosystem-based management
Index of Biotic Integrity

ABSTRACT

Investment in agricultural conservation practices (CPs) to address Lake Erie's re-eutrophication may offer benefits that extend beyond the lake, such as improved habitat conditions for fish communities throughout the watershed. If such conditions are not explicitly considered in Lake Erie nutrient management strategies, however, this opportunity might be missed. Herein, we quantify the potential for common CPs that will be used to meet nutrient management goals for Lake Erie to simultaneously improve stream biological conditions throughout the western Lake Erie basin (WLEB) watershed. To do so, we linked a high-resolution watershed-hydrology model to predictive biological models in a conservation scenario framework. Our modeling simulations showed that the implementation of CPs on farm acres in critical and moderate need of treatment, representing nearly half of the watershed, would be needed to reduce spring/early summer total phosphorus loads from the WLEB watershed to acceptable levels. This widespread CP implementation also would improve potential stream biological conditions in >11,000 km of streams and reduce the percentage of streams where water quality is limiting biological conditions, from 31% to 20%. Despite these improvements, we found that even with additional treatment of acres in low need of CPs, degraded water quality conditions would limit biological conditions in >3200 stream km. Thus, while we expect CPs to play an important role in mitigating eutrophication problems in the Lake Erie ecosystem, additional strategies and emerging technologies appear necessary to fully reduce water quality limitation throughout the watershed.

© 2016 International Association for Great Lakes Research. Published by Elsevier B.V. All rights reserved.

Introduction

Reducing nutrient inputs from the western Lake Erie basin (WLEB) watershed is integral to reversing Lake Erie's recent re-eutrophication (Ohio EPA, 2010, 2013; Scavia et al., 2014; Annex 4, 2015). This large watershed (~26,000 km^2) drains a landscape that is >70% agricultural and contains nearly 2 million ha of farmland that is mostly in corn and soybean crop rotations (USDA NRCS, 2011). Multiple changes in local agricultural practices have occurred during the past 30 years, including the type of fertilizer used, the timing of fertilizer application, tillage practices, and increased artificial drainage, the combination of which has increased the potential for nutrient runoff into Lake Erie from the WLEB watershed (Richards et al., 2002; Daloğlu et al., 2012; Smith et al., 2015). When combined with an increasing frequency of single and multi-day severe storms during the winter and spring (Hayhoe et al., 2010) and the widespread nature of legacy loads (Sharpley et al., 2013; Powers et al., 2016), these changes in agricultural practices have contributed to increased loading of highly bioavailable dissolved reactive phosphorus into Lake Erie (Richards et al., 2010; Daloğlu et al., 2012; Scavia et al., 2014). This excess phosphorus loading, in turn, has helped fuel Lake Erie's re-eutrophication (Stumpf et al., 2012; Michalak et al., 2013; Kane et al., 2014; Scavia et al., 2014). Because eutrophication poses a threat to important ecosystem services

* Corresponding author. Tel.: +1 614 688 2299.
E-mail address: keitzer.2@osu.edu (S.C. Keitzer).

http://dx.doi.org/10.1016/j.jglr.2016.05.012
0380-1330/© 2016 International Association for Great Lakes Research. Published by Elsevier B.V. All rights reserved.

Please cite this article as: Keitzer, S.C., et al., Thinking outside of the lake: Can controls on nutrient inputs into Lake Erie benefit stream conservation in its watershed? J. Great Lakes Res. (2016), http://dx.doi.org/10.1016/j.jglr.2016.05.012

provided by Lake Erie (Ludsin et al., 2001; Hobbs et al., 2002), reducing phosphorus loading from WLEB tributaries is a management priority (Ohio EPA, 2010, 2013; Scavia et al., 2014; Annex 4, 2015).

While efforts to reduce phosphorus loading will benefit Lake Erie, the extent to which these efforts will help improve water quality and biological conditions in the ecologically, culturally, and economically important stream network of the WLEB watershed remains uncertain. This network contains >20,000 km of streams and rivers that historically supported a rich diversity of invertebrates and fish (Trautman, 1981; Krebs et al., 2010). The WLEB watershed, much like Lake Erie proper, provides valuable ecosystem services (e.g., drinking water; recreational opportunities such as fishing and canoeing) to residents in Indiana, Michigan, and Ohio. Unfortunately, stream water quality in the watershed also has become degraded during the past century, owing in large part to the same agricultural sediment and nutrient non-point source (NPS) runoff that has degraded Lake Erie (Karr et al., 1985; Ohio EPA, 2014). Thus, reducing agricultural NPS runoff to help clean up Lake Erie may offer an opportunity to improve water quality and biological conditions throughout the WLEB stream network. Because farmers in the area feel a strong sense of responsibility to protect water quality in their local watersheds (Burnett et al., 2015), they might be more willing to adopt voluntary and potentially costly agricultural conservation practices (referred to as CPs hereafter), if they knew that such practices would benefit their local watershed in addition to benefiting downstream Lake Erie. Such adoption, in turn, could lead to a potential "win-win" for user groups of both Lake Erie and its watershed.

At present, however, perceived benefits of targeted phosphorus load reductions for Lake Erie have not included consideration of the possible benefits to the large stream network contained within the WLEB watershed. Thus, the extent to which targeted load reductions to Lake Erie also might improve water quality, biological conditions, and ecosystem services throughout WLEB tributaries remains an important information gap. A better understanding of where and by how much water quality and biological conditions would change throughout the WLEB watershed because of targeted load reductions to Lake Erie also could help prioritize nutrient management strategies.

Because agriculture is the dominant form of land use in the WLEB watershed, one approach to reducing nutrient loading from this watershed is to increase implementation of CPs. These CPs could include erosion control practices such as filter strips and cover crops, as well as nutrient management, which includes altering the rate, timing, amount, and method of fertilizer application. Since the mid-1970s, CPs, in particular erosion control practices such as conservation tillage, have been widely adopted in the WLEB watershed (Richards et al., 2002). These practices appear to have reduced nutrient and suspended sediment concentrations in some Lake Erie tributaries (Richards and Baker, 2002; Richards et al., 2009), and are correlated with recent improvements in stream biological conditions (Miltner, 2015). How effective additional investment in these and other widely adopted CPs would be for meeting Lake Erie nutrient reduction goals remains unknown. Even more uncertain is how additional conservation treatment of cropland would affect stream conditions and the resident aquatic biota within Lake Erie's watershed.

Herein, we provide findings from a coupled physical–biological modeling study that sought to quantify the potential benefits of increasing investment in CPs to stream biological conditions within Lake Erie's watershed. More specifically, we linked an existing high-resolution watershed-hydrology model for the WLEB watershed (Daggupati et al., 2015a) to a predictive statistical model of an Index of Biotic Integrity (IBI) developed from several long-term state-agency datasets to forecast potential benefits of additional investment in CPs. While our simulations were not designed to provide the most cost-effective solutions nor model stream impacts of reducing phosphorus loads to the levels recommended for Lake Erie, several of them more than satisfactorily met the targeted reductions in phosphorus loading to the lake. Ultimately, we discuss the potential of CPs to simultaneously meet water quality goals in Lake Erie and benefit stream biological conditions within the WLEB watershed.

Methods

Study area and species

We focused on the WLEB watershed because it is integral to effective Lake Erie nutrient management (Ohio EPA, 2010, 2013; Scavia et al., 2014; Annex 4, 2015). This relatively flat watershed (average slope is <2%) drains an ~26,000 km^2 area in portions of Ohio, Indiana, and Michigan (Fig. 1). Most of the watershed falls within the Eastern Corn Belt Plains or the Huron/Erie Lake Plains Ecoregions, although a small portion (<2%) is in the Southern Michigan/Northern Indiana Drift Plains. Historically, this watershed was comprised of a mixture of hardwood forests, wetlands, and prairie, which eventually succumbed to rapid and widespread land clearing, wetland draining, and stream channelization that began during the mid-1800s (Trautman, 1981). Today, >70% of the watershed is in row-crop agriculture, with patchily distributed urban and forested lands each making up ~12% of the remaining area. Because of this topography and land-use history, most streams in the WLEB watershed are low gradient and slow flowing, carrying heavy nutrient and sediment loads that have negatively impacted native stream biodiversity and Lake Erie (Trautman, 1939; Trautman and Gartman, 1974; Karr et al., 1985; Scavia et al., 2014).

The stream network of the WLEB watershed historically supported a diverse fish fauna (Trautman, 1981). At least 98 native fish species that span a wide range of reproductive (e.g., nest builders, crevice spawners, broadcast spawners), feeding (e.g., detritivores, herbivores, invertivores, piscivores), and habitat (e.g., benthic, pelagic, littoral) guilds have been observed in the watershed. These species have different sensitivities to nutrient and sediment pollution (Trautman, 1981; Ohio EPA, 1987). In turn, different fish communities occur throughout the WLEB watershed, with their composition likely determined to some degree by the magnitude and intensity of agricultural runoff impacts on water quality. Degraded water quality in this watershed has indeed negatively affected piscivores, herbivores, and insectivores in particular, leading to fish communities dominated by omnivorous species (Karr et al., 1985).

Modeling stream water quality

We simulated sediment and nutrient processes and stream hydrology using the Soil and Water Assessment Tool (SWAT; Arnold et al., 1998). SWAT is a semi-distributed, continuous-time model developed by the United States Department of Agriculture — Agricultural Research Service for large-scale watershed simulation. SWAT is a robust and flexible approach for simulating agricultural effects on hydrologic processes that performs well relative to other watershed models in the WLEB watershed (Gebremariam et al., 2014).

We used SWAT to develop a watershed model (Daggupati et al., 2015a) at the 1:100,000 resolution using the National Hydrography Database Plus Version 2 (NHDPlusV2) dataset (http://www.horizon-systems.com/NHDPlus/NHDPlusV2_home.php). However, because conducting simulations at this resolution was too computationally expensive, we initially calibrated model parameters at a broader watershed resolution (12-digit hydrologic unit code, HUC-12). Afterwards, we transferred those parameters to the NHDPlusV2 model to provide reasonable starting points for parameter values for this finer-resolution model. We further calibrated monthly stream flow, suspended sediment, total phosphorus (TP), and total nitrogen (TN) for the NHDPlusV2 model using five river gauges that had historical data with these attributes: 1) the Raisin River near Monroe, MI; 2) St. Joseph River near Newville, IN; 3) St. Marys River at Wilshire, OH; 4) Maumee River at Waterville, OH; and 5) Sandusky River near Fremont, OH. Detailed descriptions of the calibration and validation

ARTICLE IN PRESS

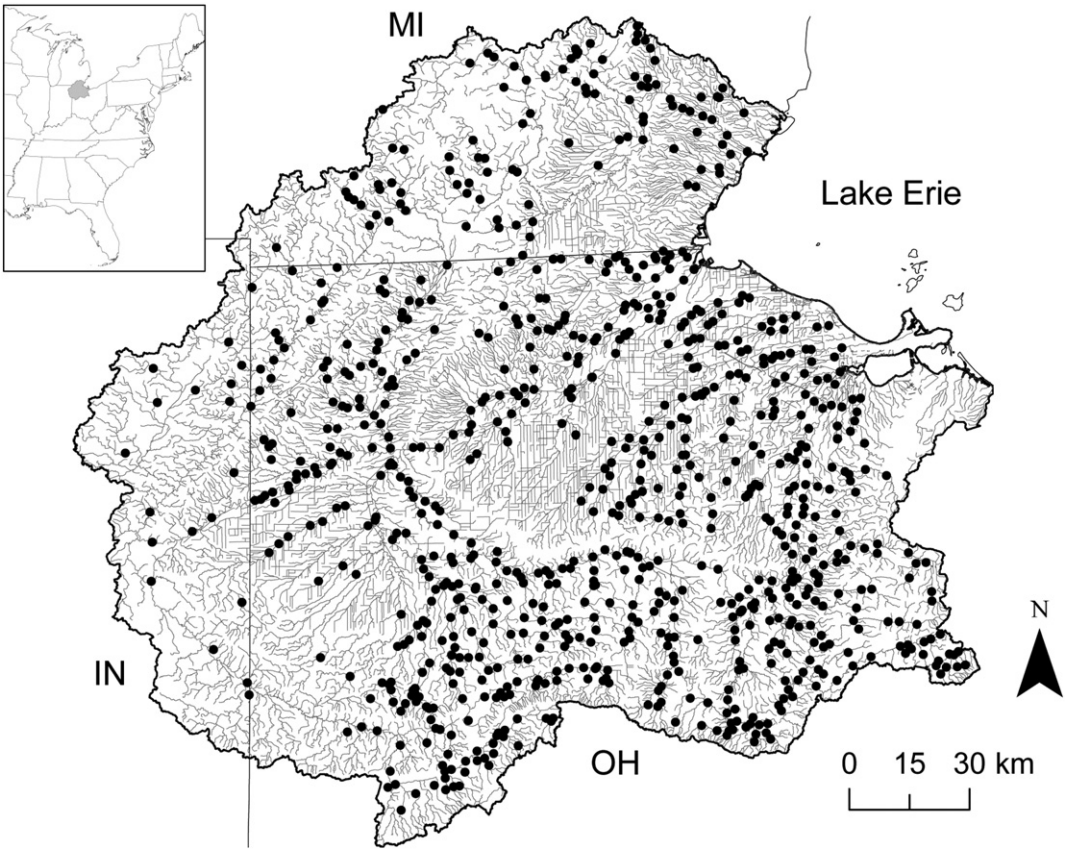


Fig. 1. The western Lake Erie basin watershed (IN-MI-OH, USA) contains >200,000 km of streams that drain a largely agricultural area. Points indicate the approximate locations of fish sample collections during 1990 to 2012 (n = 841).

procedure can be found in Daggupati et al. (2015a), but followed general guidelines for SWAT model development (Daggupati et al., 2015b). We chose to focus on percent bias (PBIAS) over other performance measures, such as Nash–Sutcliffe efficient (NSE), because PBIAS measures how well the model captures long-term average conditions, whereas measures such as NSE better assess the ability to capture extreme events (e.g., storm runoff). Because we expected long-term average conditions would have a larger effect on fish communities than extremes, our effort focused on reducing the model's PBIAS.

Predictive models of fish community health

We used a multi-metric fish community index, the IBI, to assess stream biological condition (Karr, 1981; Karr et al., 1986). Indices of biotic integrity, which compare biotic communities in a given stream reach to those expected in reference conditions (i.e., streams less disturbed by human activity), are commonly used by state and federal management agencies, academic researchers, and conservation organizations as a measure of stream biological condition. To calculate IBI values, we used fish community data collected at numerous stream sites throughout the WLEB (Fig. 1) by the Indiana Department of Environmental Management (n = 18 segments), Michigan Department of Environmental Quality and Michigan Department of Natural Resources (n = 101 segments), and the Ohio Environmental Protection Agency (n = 722 stream segments) during 1990 to 2012 (n = 841 unique stream segments). Fish sampling methods were similar among states, with electro-shocking procedures that varied more within each agency (based on the size of the stream sampled) than among agencies (Ohio EPA, 1987; Michigan DEQ, 1997; Indiana DEM, 2007). Because sampling protocols were designed to comprehensively evaluate fish communities based on species relative abundances, we assumed that they generated accurate samples of the fish community across the basin (Esselman et al., 2013). We rigorously evaluated each dataset to ensure that all collections were indeed community samples, species occurrences accurately reflected their range, and site locations were geographically accurate, similar to the approach used in Esselman et al. (2013). From these fish community data, we calculated the IBI according to Michigan's Procedure 51 (P51; Michigan DEQ, 1997). We rescaled P51 scores to have a minimum of 0 and a maximum of 1, with a higher score indicating a better biological condition.

Fish community samples were spatially linked to an NHDplusV2 stream segment based on latitude and longitude, as well as written descriptions of their location. We assumed a sample was representative of the entire stream segment. In cases where multiple sites were sampled within the same stream segment in the same year, we used the average IBI of those samples to represent the stream segment. For stream segments sampled during multiple years, we used the most recent fish community sample to reflect conditions during our watershed simulation period as closely as possible. This procedure resulted in a varying number of yearly samples through time, although some years had particularly large sample sizes (Fig. A1).

We were particularly interested in understanding how nutrients and suspended sediments affected stream biological conditions. We chose these stressors because they can affect stream biotic communities in multiple ways (e.g., Waters, 1995; Evans-White et al., 2009; Miltner, 2010; Taylor et al., 2014), and a primary approach to mitigate eutrophication in Lake Erie is implementation of CPs to reduce sediment and nutrient runoff from farm fields in the watershed. Because we were uncertain as to whether annual or sub-annual (seasonal) nutrient and sediment concentrations would better predict IBI scores, we included average annual, springtime, and summertime concentrations as predictors in our models. We focused on spring and summer because we expected agricultural pollution to have its largest impacts during these two eco-hydrological periods. These two periods also reflect

Please cite this article as: Keitzer, S.C., et al., Thinking outside of the lake: Can controls on nutrient inputs into Lake Erie benefit stream conservation in its watershed? J. Great Lakes Res. (2016), http://dx.doi.org/10.1016/j.jglr.2016.05.012

different aspects of the seasonal hydrograph and are important to the life-history of most fishes in the WLEB watershed. We defined the spring period as 1 March through 30 June, which encompasses the falling limb in the typical WLEB hydrograph (SCK, unpublished data) and the period when most fishes spawn (Auer, 1982, Ohio Division of Natural Resources, website contains fish reproductive behavior, accessed on 20 Aug. 2014: http://wildlife.ohiodnr.gov/species-and-habitats/species-guide-index/fish; M. Kibbey, The Ohio State University Museum of Biological Diversity, 2015, personal communication). We defined the summer period as 1 July through 30 September, which covers base-flow conditions and encompasses the growing season of most fishes.

To assess whether stream water quality was limiting stream-fish community health, we used multiple quantile regression to model the relationship between observed IBI scores and water quality attributes generated by our watershed model (i.e., stream discharge and TP, TN, and suspended sediment concentrations). Water quality attributes for each stream segment were derived by determining the annual median concentration (mg/L) for each time period and then calculating the average of these values from the 21-year SWAT simulation (1990–2010; Table A1). Thus, our measures of water quality reflect long-term average conditions within a stream segment. We used quantile regression because it can reveal reliable relationships despite not all influential (i.e., limiting) factors being measured, which often is the case with ecological data (Cade and Noon, 2003). We modeled the upper 97th percentile of IBI scores (hereafter referred to as the maximum IBI) in relation to water quality attributes because modeling upper percentiles offers a means to identify limiting relationships (Cade et al., 1999).

We used Akaike's Information Criterion adjusted for small sample size (AIC_c) to identify the best supported models that could reliably predict a maximum IBI score. Because multiple water quality stressors often co-occur and interact to affect stream biota (Townsend et al., 2008; Matthaei et al., 2010; Piggott et al., 2015), we included different additive and interactive effects of water quality stressors in our initial candidate set of models (Table A2). We did not include water quality attributes from multiple time periods (i.e., annual, spring, or summer) in the same model to avoid fitting overly complex models. We also included average stream discharge (L/s) as an additive effect in all models to account for the influence that stream size has on fish communities (Sheldon, 1968; Horwitz, 1978; Angermeier and Schlosser, 1989).

We did not include covariates that were highly correlated ($r \geq |0.7|$; Table A3) in the same model to avoid potential prediction errors associated with multicollinearity (Dormann et al., 2013). As a result, we could not include TN or TP in the same models. Because both TN and TP can degrade stream fish biotic integrity (Miltner and Rankin, 1998; Wang et al., 2007), we felt it was important to evaluate the potential limiting effect of TN and TP. We therefore developed a separate set of candidate models containing TN or TP (Table A2), and selected the best supported model from each candidate set as the model with the lowest AIC_c score (Burnham and Anderson, 1998). We then evaluated the accuracy of the best TN and TP model using k-fold cross-validation ($k = 10$). Because predictions of the maximum IBI from our quantile regression models were expected to be higher than observed values, we adopted assessment statistics from Vaz et al. (2008). For each k-fold training and test dataset, we calculated the Spearman's rank correlation coefficient (r_s) between observed and predicted maximum IBI and the percentage of observed scores that fell below our predictions. We considered the models acceptable, if 1) we found a statistically significant and positive correlation for each k-fold and 2) at least 97% of the observed values, on average, fell below our predicted value.

We used the best supported and validated TN model and TP model to make separate forecasts of the maximum IBI with a stream segment based on water quality conditions generated from our watershed simulation. We then selected the lowest score from these forecasts to use as the maximum IBI score within a stream segment. This approach assumes that either TN or TP, along with potential interactions with stream discharge and suspended sediment concentration, is limiting stream fish communities in a stream segment. Thus, whichever model predicts the lowest score represents the maximum IBI given the simulated water quality conditions.

We standardized all water quality variables to have a mean of zero and standard deviation of one to place water quality attributes on the same scale. While discharge data were $\log_{10}$-transformed to achieve linearity, transformations of other water quality variables were not necessary. We logit-transformed observed IBI scores because values were bounded between zero and one. All biological models were developed and analyzed with the quantreg package (Koenker, 2015) in the R statistical environment (R Core Team, 2015). We used the MuMIn package (Barton, 2015) to calculate model selection statistics.

Conservation scenarios

We used conservation scenarios developed as part of the USDA's Conservation Effects Assessment Project (CEAP) Cropland National Assessment to examine the potential benefits of increasing investment in CPs for improving stream biological conditions and Lake Erie nutrient management. These scenarios consisted of applying erosion control and nutrient management CPs at three treatment levels in the WLEB. The erosion control practices included were filter strips, field borders, surface roughening, herbaceous and forest riparian buffers, wind erosion control, cover crops, and residue tillage management. Nutrient management consisted of altering the timing, amount, rate, and/or form of nutrient application.

The CPs were applied to 1) only farm acres in critical need of treatment, 2) farm acres in critical and moderate need of treatment, and 3) all acres. An acre's need for CPs was determined by its level of inherent vulnerability for nutrient and sediment loss and its level of treatment based on farmer surveys from a subset of the National Resources Inventory (USDA NRCS, 2011). According to these farmer surveys, about 8% (384,160 acres) of the watershed is in critical need of treatment and 40% (1,920,800 acres) is in moderate need of treatment for at least one resource concern (i.e., sediment, phosphorus, or nitrogen loss through surface or subsurface flows; USDA NRCS, 2011). In most cases, several CPs were applied to the same field to address multiple resource concerns (USDA NRCS, 2011). We assumed a CP adoption rate of 80% on the acres in need of treatment and that the "best" option would be chosen on 75% of the treated acres to implicitly account for limitations in CP adoption associated with individual farmer behavior. In addition to these three conservation scenarios, we included a "grassland" scenario in which we replaced all agricultural land with native grasses. This grassland scenario served as a theoretical upper benchmark to assess impacts of agricultural runoff and the effectiveness of CPs.

Assessing the benefits of conservation scenarios

We simulated edge-of-field reductions in agricultural runoff from CP implementation using the Agricultural Policy/Environmental eXtender (APEX) model (Gassman et al., 2009). APEX is a flexible model capable of simulating a broad array of agricultural management options, including nutrient management, tillage operations, cropping systems, and CPs (Wang et al., 2011). A major advantage of APEX is that spatially explicit relationships in field units allow for realistic and physically-based simulations of CP reductions in surface and subsurface runoff (Wang et al., 2011). Reductions in agricultural runoff were simulated in APEX through alteration of a number of model parameters (e.g., curve number) according to a mixture of empirical and theoretical data (Wang et al., 2011). Reductions at the field-scale were aggregated for an NHDplusV2 catchment and coupled to our watershed model to simulate changes in water quality and phosphorus loads during 1990 to 2010 in each conservation scenario.

ARTICLE IN PRESS

We calculated the combined spring/early summer TP loads (1 Feb. to 31 July) from four major WLEB tributaries (River Raisin, Maumee River, Portage River, and Sandusky River) to examine the effectiveness of conservation scenarios for meeting Lake Erie nutrient management goals. These nutrient management goals recommend a 40% reduction in TP loads relative to 2008 loads to reduce the likelihood of harmful algal bloom formation in Lake Erie (Ohio EPA, 2013; Annex 4, 2015). However, observed TP loading data were not available for spring/early summer in 2008 for the River Raisin or Portage River (tributary loading data from National Center for Water Quality Research, accessed on 3 March 2016: www.heidelberg.edu/academiclife/distinctive/ncwqr/data/data). We therefore used a target TP load of 974.39 MT, which represents a 40% reduction from our simulated load for major WLEB tributaries in 2008.

To assess the potential benefits of CP implementation for improving stream biotic integrity, we first forecasted the potential maximum IBI in each stream segment for the three conservation scenarios. We used one-sided paired *t*-tests to determine if the maximum IBI score was greater in conservation scenarios than in the baseline scenario. We then calculated the amount of stream kilometers in which the maximum IBI improved as our first stream assessment measure. For our second stream assessment measure, we examined the average change in maximum IBI scores that occurred in each scenario to assess the magnitude of improvements in stream fish communities that occurred with additional CP investment. We used one-sided paired *t*-tests to determine if the change in maximum IBI scores continued to increase with increasing CP implementation. To better understand what changes in water quality were responsible for changes in the maximum IBI, we examined correlations between changes in the maximum IBI and changes in water quality attributes across all scenarios.

For our third stream assessment measure, we used different percentiles of the observed fish community data to assess where water quality was limiting fish community conditions. We considered streams where the potential maximum IBI score was below the 70th percentile of the observed data (IBI < 0.65) to be "limited" by water quality, those between the 70th and 90th percentiles (0.65 ≤ IBI < 0.80) to be "moderately limited" by water quality, and those greater than the 90th percentile (IBI ≥ 0.80) of the observed data to be "not limited" by water quality. We assessed a conservation scenario's effectiveness by examining how the percentage of the total stream length in each of these categories changed from baseline conditions.

Results

Watershed model performance

We considered watershed model performance acceptable if the PBIAS during model validation was ±25% for stream discharge, ±55% for suspended sediment loads, and ±70% for TN and TP loads (Moriasi et al., 2007). Based on these performance ratings, our watershed model performed well for all water quality attributes but stream discharge in the River Raisin (Table 1). Although we did not specifically calibrate for NSE or the coefficient of determination (R^2), our watershed model performed reasonably well for the Maumee River and Sandusky River, with NSE and $R^2 > 0.5$ for most parameters (Table 1). Performance was less accurate at other gauges based on these measures for stream discharge, and was poor in many cases for other parameters. Thus, our model accurately captured long-term average conditions across the watershed, but performed less well at capturing extreme events. In terms of simulating annual spring/early summer TP loads from WLEB tributaries, the watershed model performed well ($R^2 = 0.70$) and tended to underestimate TP (PBIAS = −15.5) loads from WLEB tributaries (Fig. 2a).

Biological model performance

The best-supported TN model for predicting maximum IBI scores in the WLEB watershed included spring discharge, spring suspended sediment concentration, and an interaction between spring TN concentration and spring suspended sediment concentration as predictors (Table 2). The best-supported TP model also included an interaction with suspended sediment concentration, but in this case, annual water quality conditions were selected. Validation statistics suggest that both the TN model and TP model performed well (Table 2). The average r_s for the TN model was 0.52 (range = 0.43–0.63) and 0.51 (range = 0.40–0.58) for the TP model, with all relationships statistically significant (*p*-values < 0.001). The average percentage of the observed IBI scores that fell below the predicted values was 97% for the TN model (range = 94–100%) and TP model (range = 93–99%). Overall, these statistics indicated that both models correctly identified the upper limits in stream biotic integrity (maximum IBI) based on stream discharge and the water quality attributes that we modeled.

Baseline conditions

Fish community health

According to our classifications of water quality limitation, we found that 31% (6323 km) of streams were limited, 43% (8781 km) were moderately limited, and 25% (5154 km) were not limited by water quality in the baseline scenario (Fig. 2a; Table 3). Thus, if we assume that removing water quality as a limiting factor for all streams represents an ideal goal for stream conservation, water quality appears to be limiting stream fish communities to some degree in ~75% (15,104 km) of WLEB streams.

Lake Erie loads

Inputs of TP into Lake Erie from major WLEB tributaries varied considerably during our 21-year simulation (Fig. 2a). The average spring/early summer TP load from major WLEB tributaries into Lake Erie was estimated as 1186 metric tons (MT). The Maumee River contributed nearly 74% of this load (873 MT). The next highest contributor was the Sandusky River (195 MT), followed by the River Raisin (66 MT) and the Portage River (51 MT). The combined TP loads from major WLEB tributaries exceeded management targets in 15 of the 21 years.

Table 1
Model validation statistics for the watershed model at the monthly time-step for stream discharge and suspended sediment, total phosphorus (TP), and total nitrogen (TN) loads. We report the coefficient of determination (R^2), Nash–Sutcliffe efficiency (NSE), and percent bias (PBIAS), but focused on minimizing PBIAS, which reflects long-term average conditions, during model development. Bold values indicate acceptable model performance based on the PBIAS (Moriasi et al., 2007).

	Stream discharge			Suspended sediment			TP			TN		
Station	R^2	NSE	PBIAS	R^2	NSE	PBIAS	R^2	NSE	PBIAS	R^2	NSE	PBIAS
River Raisin	0.72	0.43	26%	0.36	0.27	**−35%**	0.44	−0.96	**23%**	0.5	0.35	**−4%**
St. Joseph River	0.78	0.74	**−19%**	0.26	−0.01	**20%**	0.26	−0.56	**−5%**	0.69	−0.64	**69%**
St. Marys River	0.49	0.43	**−25%**	0.13	0.13	**−19%**	0.24	0.07	**−9%**	0.31	0.27	**−21%**
Maumee River	0.92	0.88	**−14%**	0.66	0.66	**10%**	0.79	0.78	**−3%**	0.71	0.69	**−13%**
Sandusky River	0.76	0.75	**−7%**	0.35	−0.97	**35%**	0.54	0.38	**−1%**	0.64	0.57	**12%**

Please cite this article as: Keitzer, S.C., et al., Thinking outside of the lake: Can controls on nutrient inputs into Lake Erie benefit stream conservation in its watershed? J. Great Lakes Res. (2016), http://dx.doi.org/10.1016/j.jglr.2016.05.012

ARTICLE IN PRESS

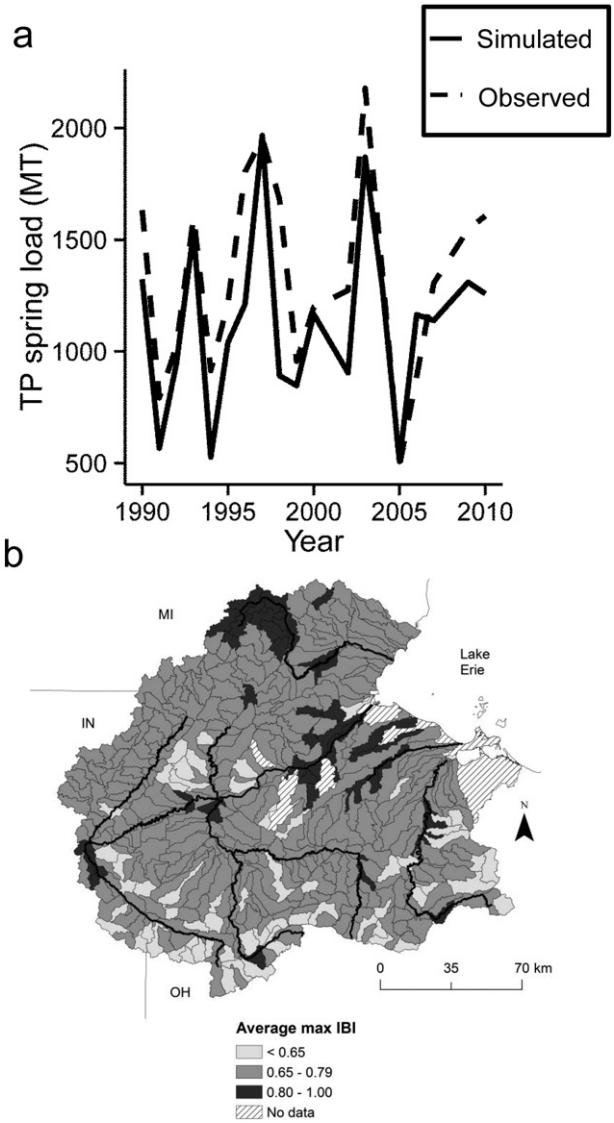


Fig. 2. Modeled baseline conditions (1990–2010) of (a) and (b) distribution of stream biological conditions, as measured by an Index of Biotic Integrity for fish (max IBI). Combined total phosphorus (TP) during spring/early summer from major tributaries in the western Lake Erie basin (WLEB). Observed nutrient loading data are from the National Center for Water Quality Research at Heidelberg University (accessed on 10 Nov 2015: www.heidelberg.edu/academiclife/distinctive/ncwqr). Totals do not include the contribution of the Portage River because no observed loading data were available from the 1990s.

Potential benefits of conservation scenarios

Stream water quality

The implementation of CPs improved stream water quality throughout the watershed (Table 4). Predicted nutrient and sediment concentrations varied across scenarios, with average in-stream suspended sediment and nutrient concentrations declining as the number of acres treated increased (Table 4). These improvements in water quality were small when only critical acres were treated with CPs (<7% on average). However, across spring and summer seasons, average reductions of 38% for TN, 45% for TP, and 51% for suspended sediments occurred when all farm acres were treated (Table 4). These reductions were not due to CP effects on river flows, as average stream discharge either increased slightly (<3%) with CP implementation (Critical & Moderate and All scenarios) or declined negligibly (0.02%) relative to baseline conditions (Critical scenario; Table 4).

We also found that water quality stressors may increase in some stream segments with CP implementation. These increases occurred in 18% of the watershed on average when only critical farm acres were treated, 5% when critical and moderate acres were treated, and 2% when all farm acre types were treated. The largest increases occurred when only critical farm acres were treated and related to increases during the spring. Changes were generally small, with mean changes of 3, 0.004, 0.09 mg/L for suspended sediments, TP, and TN, respectively. However, large changes did occur in some stream segments; the largest increase in suspended sediment concentration was 79 mg/L, 0.05 mg/L for TP, and 1.04 mg/L for TN.

Fish community health

One-sided paired *t*-tests showed that improvements in the average maximum IBI score were statistically significant with CP implementation compared to the baseline scenario in our simulations (*p*-values of all comparisons <0.001). The length of streams with a predicted increase in their maximum IBI ranged from 5436 km (27% of the watershed; critical acres treated) to 13,297 km (66% of the watershed; all acre types treated) of the watershed (Fig. 3a). While the average change in the maximum IBI across the watershed was generally small (range = 0.005 to 0.04; Fig. 3b), comparisons of conservation scenarios indicate that increasing CP implementation continued to increase the change in maximum IBI (critical acres vs. critical and moderate acres: $t_{9158} = 46.74$, *p*-value < 0.001; critical and moderate acres vs. all acres: $t_{9158} = 43.87$, *p*-value < 0.001). Despite relatively small changes

Table 3
Percentage of the total western Lake Erie Basin stream length that fell within different stream biological categories for the Index of Biotic Integrity (IBI) under baseline conditions, the three conservation scenarios (Critical, Critical & Moderate, All), and the grassland scenario. The mean (s.d.) maximum IBI that occurred based on simulated water quality conditions is also shown. "Acres treated" refers to the type of farm acres treated in each conservation scenario based on their need for conservation treatment. The baseline and grassland conditions are shown for reference, respectively representing average water quality conditions (1990–2010) and the theoretical upper benchmark.

Acres treated	Average maximum IBI	Limited	Moderately limited	Not limited
Baseline	0.71 (0.13)	31	43	25
Critical	0.72 (0.13)	29	45	26
Critical & Moderate	0.74 (0.12)	20	50	30
All	0.75 (0.12)	16	53	31
Grassland	0.76 (0.11)	11	54	35

Table 2
The best supported predictive total nitrogen (TN) and total phosphorus (TP) models of Index of Biotic Integrity scores for fish in the western Lake Erie basin watershed during 1990–2012, based on Akaike's Information Criterion adjusted for small sample size (Table A2). Parameter estimates on the logit-scale are shown for the model intercept, stream discharge (Q), TN, TP, suspended sediments (SS), and interactions between TN and SS (TN × SS) and TP and SS (TP × SS). Model validation statistics shown are the mean (standard deviation) of the Spearman's rank coefficient (r_s) and the percentage of observed values that were below predictions (% below) from the *k*-fold cross-validation. Note that the TP model is for average summer conditions and the TN model is for average spring conditions.

Model	Intercept	Q	TN	TP	SS	TN × SS	TP × SS	r_s	% below
TN	1.63	0.69	−0.41	–	0.16	0.12	–	0.52 (0.06)	97 (2)
TP	1.60	0.66	–	0.08	−0.21	–	0.25	0.51 (0.07)	97 (2)

Please cite this article as: Keitzer, S.C., et al., Thinking outside of the lake: Can controls on nutrient inputs into Lake Erie benefit stream conservation in its watershed? J. Great Lakes Res. (2016), http://dx.doi.org/10.1016/j.jglr.2016.05.012

ARTICLE IN PRESS

Table 4
Average water quality parameters for spring and summer discharges, total nitrogen (TN), total phosphorus (TP), and suspended sediment. "Acres treated" refers to the type of farm acres treated in each conservation scenario based on their need for conservation treatment. The baseline and grassland conditions are shown for reference, respectively representing average water quality conditions (1990–2010) and the theoretical upper benchmark.

Acres treated	Discharge (L/s)		TN (mg/L)		TP (mg/L)		Suspended sediment (mg/L)	
	Spring	Summer	Spring	Summer	Spring	Summer	Spring	Summer
Baseline	3727	1243	8.75	2.29	0.25	0.16	171.49	99.06
Critical	3726	1248	8.17	2.15	0.23	0.15	157.89	94.01
Critical & Moderate	3770	1256	6.18	1.68	0.16	0.12	113.96	76.21
All	3809	1262	5.4	1.4	0.12	0.1	70.48	55.65
Grassland	3797	1261	1.94	0.68	0.05	0.07	29.24	33.6

in the maximum IBI overall, substantial improvements were possible in some stream segments; the largest change in the maximum IBI ranged from 0.21 to 0.36 across the three CP scenarios, with changes in maximum IBI scores positively related to the number of acres treated. Across all scenarios, the change in the maximum IBI was strongly correlated with reductions in TN concentrations ($r_s = -0.60$, $p < 0.001$), and to a lesser degree, by reductions in TP concentrations ($r_s = -0.47$, $p < 0.001$), suspended sediment concentrations ($r_s = -0.37$, $p < 0.001$), and increases in discharge ($r_s = 0.14$, $p < 0.001$).

These changes translated into large areas where water quality was no longer limiting stream biological conditions (Table 3). For instance, our model simulations predicted that the percentage of the WLEB watershed stream kilometers limited by water quality would be reduced to 20% (from 31% under baseline conditions) when all acres types are treated with CPs (Table 3). This represents 2228 stream km where water quality was no longer limiting biological conditions. Despite CPs resulting in a net improvement in biological conditions, it should be noted that we observed declines in the maximum IBI score in each CP scenario. These declines occurred in 7%, 8%, and 5% of the watershed from our least to most intensive CP scenarios.

Lake Erie loads

Our modeling simulations also indicated that investment in CPs could sufficiently reduce TP loads into Lake Erie to meet loading goals (Fig. 3c). To do so, however, a large additional portion of farm acres would need to be treated. Implementing CPs to only the acres in critical need of treatment (~8% of the watershed) resulted in an average TP load of 1088 MT, with loads exceeding the target load in 13 out of 21 years. By contrast, applying CPs to acres in critical and moderate need (~48% of the watershed) resulted in an average TP load of 766 MT, with TP loads exceeding the target in only 4 of 21 years. Further benefits were possible by treating all acre types with CPs; the average TP load dropped to 575 MT and only 2 years exceeded the target load.

Improvements relative to grassland scenario

While improved biological conditions were predicted with CP implementation, even the treatment of all acre types with CPs did not reach the levels simulated in the grassland scenario. While our simulations revealed that the length of streams with improved maximum IBIs was only 1.2- to 3.6-fold higher in the grassland scenario relative to our three CP scenarios (Fig. 3a), the average change (i.e., increase) in

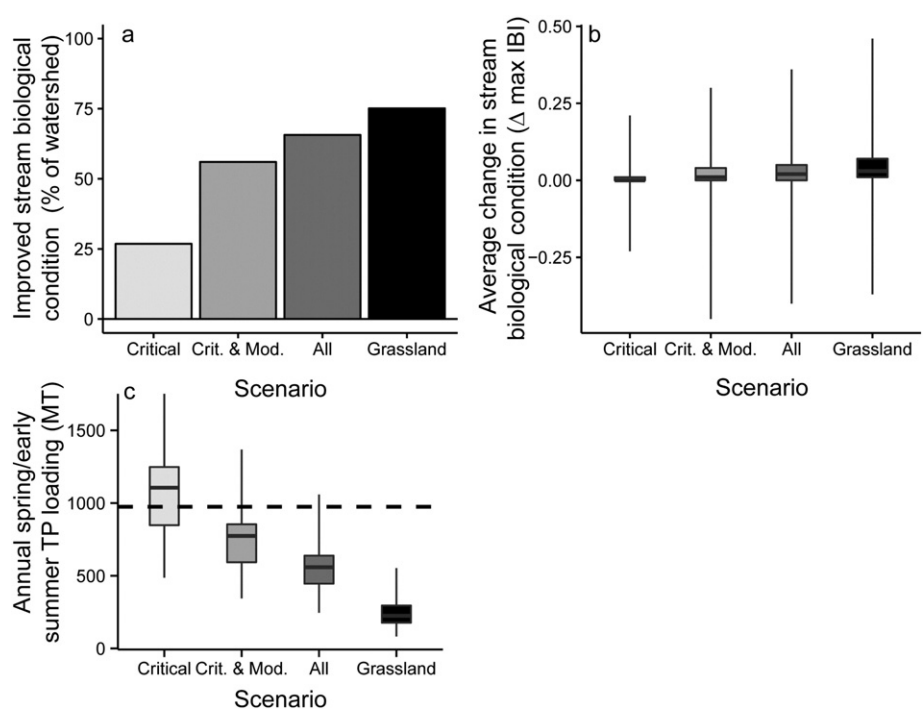


Fig. 3. Potential of agricultural conservation practices (CPs) to improve stream biological conditions and reduce phosphorus loading from the western Lake Erie basin (WLEB) watershed into Lake Erie. (a) Percentage of WLEB streams in which maximum Index of Biotic Integrity (max IBI) score improved, (b) average change in the max IBI score within a stream segment, and (c) combined annual spring/early summer total phosphorus (TP) loading from major WLEB tributaries (dashed line is baseline loading). Boxes span the interquartile range with whiskers extending to the minimum and maximum values of the change in max IBI scores or the annual spring/early summer load from 1990 to 2010.

Please cite this article as: Keitzer, S.C., et al., Thinking outside of the lake: Can controls on nutrient inputs into Lake Erie benefit stream conservation in its watershed? J. Great Lakes Res. (2016), http://dx.doi.org/10.1016/j.jglr.2016.05.012

maximum IBI was between 1.4- and 8.4-fold higher for the grassland scenario (Fig. 3b). Similarly, the percentage of the WLEB watershed limited by water quality was between 1.4- and 2.6-fold lower in the grassland scenario relative to the three CP scenarios, while the percentage of streams not limited by water quality was between 1.3- and 1.1-fold higher than in the three CP scenarios (Table 3).

Discussion

Our biophysical modeling simulations suggest that CP implementation will need to be widespread to meet nutrient management goals for Lake Erie and improve stream biological conditions. We found that only treating farm acres in critical need of CPs (~8% of the farm acres) resulted in slight reductions in spring/early summer TP loads, with TP loads exceeding the target load ~62% of the time. If a larger portion of farm acres were treated with CPs (e.g., those in critical and moderate need or about ~48% of the farm acres), however, our simulations show that it would be possible to meet nutrient management goals for TP loading in 18 of the 21 years we simulated. Other Great Lakes modeling efforts also have suggested that widespread implementation of CPs may be needed to reduce nutrient loading (Hobbs et al., 2002; USDA NRCS, 2011; Bosch et al., 2013; Scavia et al., 2016). Similarly, widespread CP implementation was needed to reduce water quality limitation on stream biological conditions. Implementing CPs on farm acres in critical need of treatment improved the maximum IBI in ~27% of the watershed, but the magnitude of these changes was relatively modest compared to more widespread CP implementation. For instance, treating farm acres in critical and moderate need not only improved stream biological conditions in nearly twice as many stream kilometers compared to only treating farm acres in critical need, but the magnitude of these changes was nearly four times as large on average. Both modeling (Einheuser et al., 2012) and empirical (Wang et al., 2002) studies indicate that widespread CP implementation is needed to improve stream biological conditions. Taken together, these studies support our conclusion that CP implementation will need to be widespread to achieve Lake Erie water quality goals and improve stream biological conditions across the WLEB watershed.

While both TP and suspended sediments certainly appear to be affecting fish community health in the WLEB watershed, our modeling suggests that reductions in stream TN concentrations would lead to the largest improvement in stream biological conditions. High nitrogen concentrations negatively affect stream fish communities (Miltner and Rankin, 1998; Wang et al., 2007; Miltner, 2010); likely because excess nitrogen feeds algal productivity (Dodds et al., 2002, 2006; Miltner, 2010) which, in turn, alters dissolved oxygen levels in streams (Miltner and Rankin, 1998; Miltner, 2010). Excess nitrogen may also alter the quality of food resources (Evans-White et al., 2009) and some forms of nitrogen, such as nitrate, may be toxic to fish at chronically high levels (Camargo et al., 2005). While CP implementation was able to reduce TN concentrations by ~40% when all acres were treated, these reductions, when combined with other water quality stressors, were unable to remove water quality limitation across the whole watershed.

Along these lines, our modeling suggests that while CPs will improve stream water quality, they are not a panacea. Even if CPs were implemented across the majority of agricultural lands in the watershed (i.e., all farm acre types treated), not all stream fish communities would become free from limitation by nutrient and sediment pollution. For instance, >3241 stream km would still be limited by poor water quality in this conservation scenario. We also found that stream biological conditions might actually decrease in some areas, a result observed in a similar study (Einheuser et al., 2012). This may reflect the fact that water quality stressors increased with CP implementation in some stream segments. It is unclear why these increases occurred, but other modeling studies suggest that CP implementation can sometimes increase NPS pollution (Einheuser et al., 2012; Scavia et al., 2016). Widespread implementation of CPs also fell short of the stream biological conditions observed in the grassland scenario, indicating that it is unrealistic to expect CPs to achieve these more "natural" conditions. Thus, additional approaches (e.g., further reducing point sources and urban and suburban runoff), emerging technologies, and more efficient CP implementation (i.e., optimized placement of CPs) may be necessary to further reduce water quality limitation of stream biological conditions throughout the watershed.

Our modeling exercise was admittedly imperfect and may underestimate the potential effectiveness of CPs for improving water quality. For example, our scenario approach used CPs that are currently in wide use across the WLEB watershed and does not apply the newest (e.g., tile biofilters, two-stage ditches, ditch plugs) or necessarily the most appropriate management options (e.g., wetlands and drainage water management were not included). Farm fields often contain acreage with various inherent vulnerabilities, including some areas that are more prone to leaching losses and some that are more prone to runoff losses. Accounting for these inherent vulnerabilities through comprehensive conservation planning would potentially allow land managers and decision makers to develop more holistic and effective conservation plans to address specific resource concerns on specific acres with the most appropriate and up to date practices. If done, achieving greater reductions in nutrient and sediment runoff might be possible by employing a more optimized management approach than we simulated in our scenarios.

Our approach also does not consider how continued climate change may affect the region. Climate change is expected to increase Maumee River discharges into Lake Erie during winter and spring, when farmland typically remains fallow and the potential for nutrient runoff is high (Cousino et al., 2015). Forecasted changes in rainfall patterns (e.g., more intense storms) also may reduce the effectiveness of CPs in controlling agricultural runoff (Hall et al., this issue). Research is needed to understand how continued climate change might alter the amount and types of CPs needed to address agricultural NPS. Such information would allow for the development of more effective nutrient management strategies to rehabilitate and protect Lake Erie and the biological communities within its watershed.

Despite these limitations, we feel that our general approach of using a biophysical model to quantify the impact of CPs on water quality within the lake, as well as within its watershed, is valuable for two reasons. First, our modeling indicates that the use of CPs to meet nutrient management goals will lead to a "win-win" for both Lake Erie and its watershed. Agriculturally-driven NPS pollution is diffuse by its nature, with numerous sources contributing unquantified amounts of nutrients and sediments across large areas (Carpenter et al., 1998). Our modeling shows that sufficiently treating this diffuse problem to address downstream concerns also may benefit biological conditions throughout the large stream network of the WLEB watershed. Second, our modeling approach offers a way to facilitate more holistic and considered conservation planning by showing that management strategies designed to rehabilitate Lake Erie—which can be hundreds of kilometers downstream from some farms (e.g., those in Michigan or the southwestern part of the basin in Ohio)—also can benefit water quality and fish communities in local streams. This knowledge may help motivate famers to adopt voluntary conservation practices, especially considering that farmers in the WLEB watershed feel a strong sense of responsibility for their local water quality (Burnett et al., 2015).

Conclusions

Our study offers four key insights that are relevant to the management of Lake Erie and other ecosystems suffering from eutrophication driven by agricultural runoff. First, efforts aimed at cleaning up a downstream ecosystem also are likely to benefit water quality and biological communities in local streams in the surrounding watershed. This notion is supported by findings from other watersheds in which

ARTICLE IN PRESS

fish communities (Wang et al., 2002; Christensen et al., 2012; Miltner, 2015) were shown to benefit from in-stream habitat and water quality associated with CPs. Second, the implementation of CPs likely will need to be widespread within the WLEB watershed to achieve Lake Erie nutrient management and watershed conservation goals. This notion is supported by other, similar studies conducted across the Great Lakes Region, including Lake Erie (Hobbs et al., 2002; USDA NRCS, 2011; Bosch et al., 2013; Einheuser et al., 2012; Scavia et al., 2016). Third, while our modeling suggests that CPs are likely to improve the health of the Lake Erie ecosystem, expectations should be realistic and it is likely that some areas of degraded conditions will persist. For instance, even if CPs were implemented across most of the watershed, stream biological conditions may still be limited by water quality in the most degraded areas of the watershed. In addition, other factors unrelated to water quality [e.g., invasive species, dispersal barriers, and habitat degradation (e.g., channelization, lack of riparian vegetation)] also likely would limit full recovery (Ohio EPA, 1999; Wang et al., 2006). Finally, because farmers can feel a strong sense of responsibility for their local water quality (Burnett et al., 2015), which likely reflects their consideration of local ecosystem services (e.g., fishing and swimming opportunities, availability of safe drinking water), more studies that consider the benefits of CPs to the surrounding stream network—not just the larger downstream ecosystem—are needed. Such knowledge could prove useful in helping promote the value of CPs to farmers and policymakers, especially when they can lead to a "win-win" for far-off water quality management and local biological conservation.

Acknowledgments

We thank the Indiana Department of Environmental Management, Michigan Department of Environmental Quality, Michigan Department of Natural Resources, and Ohio Environmental Protection Agency for access to their historical databases. We thank Maura O'Brien for her help with creating data layers used in these analyses and Jay Atwood for providing conservation scenario data. We thank Brain Zimmerman and Marc Kibbey for their help in securing, creating, and checking datasets used in these analyses. We thank Steve Davis, Gail Hesse, Ken Krieger, Bill Stanley, Carrie Vollmer-Sanders, and Chris Winslow for their discussions that helped improve the generation of ideas and their presentation in the manuscript. Chris May, Bill Stanley, John Stark, Carrie Vollmer-Sanders, and two anonymous reviewers provided helpful comments that greatly improved earlier versions of this manuscript. Funding for this project was provided by the USDA-NRCS Conservation Effects Assessment Project (cooperative agreement nos. 68-7482-12-504 and 68-7482-13-502, and interagency agreement no. 67-3A75-12-68 to JA).

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jglr.2016.05.012.

References

Angermeier, P.L., Schlosser, I.J., 1989. Species–area relationship for stream fishes. Ecology 70, 1450–1462.
Annex 4, 2015. Great Lakes water quality agreement: annex 4 objectives and targets task team. Recommended phosphorus loading targets for Lake Erie. Final Report to the Nutrients Annex Subcommittee (70 pp.).
Arnold, J.G., Srinivasan, R., Muttiah, R.S., Williams, J.R., 1998. Large area hydrologic modeling and assessment. I. Model development. J. Am. Water Resour. Assoc. 34, 73–89.
Auer, N., 1982. Identification of larval fishes of the Great Lakes basin with emphasis on the Lake Michigan drainage. Spec. Publ. No. 82–3, Great Lakes Fish. Comm., Ann Arbor, Michigan (744 pp.).
Barton, K., 2015. MuMIn: Multi-model Inference (R package version 1.15.1).
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 Lakes Res. 39, 429–436.
Burnett, E.A., Wilson, R.S., Roe, B., Howard, G., Irwin, E., Zhang, W., Martin, J., 2015. Farmers, Phosphorus and Water Quality: Part II. A Descriptive Report of Beliefs, Attitudes and Best Management Practices in the Maumee Watershed of the Western Lake Erie Basin. The Ohio State University, School of Environment & Natural Resources, Columbus, OH.
Burnham, K.P., Anderson, D.R., 1998. Model Selection and Multimodel Inference: A Practical Information-theoretic Approach. second ed. Springer, New York.
Cade, B.S., Noon, B.R., 2003. A gentle introduction to quantile regression for ecologists. Front. Ecol. Environ. 1, 412–420.
Cade, B.S., Terrell, J.W., Schroeder, R.L., 1999. Estimating effects of limiting factors with regression quantiles. Ecology 80, 311–323.
Camargo, J.A., Alonso, A., Salamanca, A., 2005. Nitrate toxicity to aquatic animals: a review with new data for freshwater invertebrates. Chemosphere 58, 1255–1267.
Carpenter, S.R., Caraco, N.F., Correll, D.L., Howarth, R.W., Sharpley, A.N., Smith, V.H., 1998. Nonpoint pollution of surface waters with phosphorus and nitrogen. Ecol. Appl. 8, 559–568.
Christensen, V.G., Lee, K.E., McLees, J.M., Niemela, S.L., 2012. Relations between retired agricultural land, water quality, and aquatic-community health, Minnesota River Basin. J. Environ. Qual. 41, 1459–1472.
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.
Daggupati, P., Yen, H., White, M.J., Srinivasan, R., Arnold, J.G., Keitzer, S.C., Sowa, S.P., 2015a. Impact of model development, calibration and validation decisions on hydrological simulations in West Lake Erie Basin. Hydrol. Process. 29, 5307–5320.
Daggupati, P., Pai, N., Ale, S., Douglas-Mankin, K.R., Zeckoski, R., Jeong, J., Parajuli, P., Saraswat, D., Youssef, M., 2015b. A recommended calibration and validation strategies for hydrological and water quality models. Trans. ASABE 58, 1705–1719.
Daloğlu, I., Cho, K.H., Scavia, D., 2012. Evaluating causes of trends in long-term dissolved reactive phosphorus loads to Lake Erie. Environ. Sci. Technol. 46, 10660–10666.
Dodds, W.K., Smith, V.H., Lohman, K., 2002. Nitrogen and phosphorus relationships to benthic algal biomass in temperate streams. Can. J. Fish. Aquat. Sci. 59, 865–874.
Dodds, W.K., Smith, V.H., Lohman, K., 2006. Erratum: nitrogen and phosphorus relationships to benthic algal biomass in temperate streams. Can. J. Fish. Aquat. Sci. 63, 1190–1191.
Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G., Gruber, B., Lafourcade, B., Leitão, P.J., Münkemüller, T., McClean, C., Osborne, P.E., Reineking, B., Schröder, B., Skidmore, A.K., Zurell, D., Lautenbach, S., 2013. Collinearity: a review of methods to with it and simulation study evaluating their performance. Ecography 36, 27–46.
Einheuser, M.D., Nejadhashemi, A.P., Sowa, S.P., Wang, L., Hamaamin, Y.A., Woznicki, S.A., 2012. Modeling the effects of conservation practices on stream health. Sci. Total Environ. 435–436, 380–391.
Esselman, P.C., Infante, D.M., Wang, L., Cooper, A.R., Wieferich, D., Tsang, Y., Thornbrugh, D.J., Taylor, W.W., 2013. Regional fish community indicators of landscape disturbance to catchments of the conterminous United States. Ecol. Indic. 26, 163–173.
Evans-White, M.A., Dodds, W.K., Huggins, D.G., Baker, D.S., 2009. Thresholds in macroinvertebrate biodiversity and stoichiometry across water-quality gradients in Central Plains (USA) streams. J. N. Am. Benthol. Soc. 28, 855–868.
Gassman, P.W., Williams, J.R., Wang, X., Saleh, A., Osei, E., Hauck, L.M., Izaurralde, R.C., Flowers, J.D., 2009. The agricultural policy environmental extender (APEX) model: an emerging tool for landscape and watershed environmental analyses. Trans. ASABE 53, 711–740.
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. Softw. 71, 121–134.
Hayhoe, K., VanDorn, J., Croley, T., Shlegal, N., Wuebbles, D., 2010. Regional climate change projections for Chicago and the US Great Lakes. J. Great Lakes Res. 36, 7–21.
Hobbs, B.F., Ludsin, S.A., Knight, R.L., Ryan, P.A., Biberhofer, J., Ciborowski, J.J.H., 2002. Fuzzy cognitive mapping as a tool to define management objectives for complex ecosystems. Ecol. Appl. 12, 1548–1565.
Horwitz, R.J., 1978. Temporal variability patterns and the distributional patterns of stream fishes. Ecol. Monogr. 48, 307–321.
Indiana DEM, 2007. Indiana Department of Environmental Management, Surface Water Quality Assessment Program. Fish Community Sampling Program. Fact Sheet B-004-OWQ-A-BS-06-0-R6 (Indianapolis, IN).
Kane, D.D., Conroy, J.D., Richards, R.P., Baker, D.B., Culver, D.A., 2014. Re-eutrophication of Lake Erie: correlations between tributary nutrient loads and phytoplankton biomass. J. Great Lakes Res. 40, 496–501.
Karr, J.R., 1981. Assessment of biotic integrity using fish communities. Fisheries 6, 21–27.
Karr, J.R., Toth, L.A., Dudley, D.R., 1985. Fish communities of Midwestern rivers: a history of degradation. Bioscience 35, 90–95.
Karr, J.R., Fausch, K.D., Angermeier, P.L., Yant, P.R., Schlosser, I.J., 1986. Assessing Biological Integrity in Running Waters: A Method and Its Rationale. Illinois Natural History Survey Special Publication 5 (28 pp.).
Koenker, R., 2015. Quantreg: Quantile Regression (R package version 5.11).
Krebs, R.A., Borden, W.C., Steiner, E.R., Lyons, M.S., Zawiski, W., Walton, B.M., 2010. Determinants of mussel diversity in Lake Erie tributaries. J. N. Am. Benthol. Soc. 29, 506–520.
Ludsin, S.A., Kershner, M.W., Blocksom, K.A., Knight, R.L., Stein, R.A., 2001. Life after death in Lake Erie: nutrient controls drive fish species richness, rehabilitation. Ecol. Appl. 11, 731–746.
Matthaei, C.D., Piggott, J.J., Townsend, C.R., 2010. Multiple stressors in agricultural streams: interactions among sediment addition, nutrient enrichment and water abstraction. J. Appl. Ecol. 47, 639–649.
Michalak, A.M., Anderson, E., Beletsky, D., Boland, S., Bosch, N.S., Bridgeman, T.B., Chaffin, J.D., Cho, K.H., Confesor, R., Daloğlu, I., DePinto, J., Evans, M.A., Fahnenstiel, G.L., He, L.,

ARTICLE IN PRESS

Ho, J.C., Jenkins, L., Johengen, T., Kuo, K.C., Laporte, E., Liu, X., McWilliams, M., 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. 110, 6448–6452.

Michigan DEQ, 1997. Michigan Department of Environmental Quality, Surface Water Quality Division. GLEAS procedure #51 survey protocols for wadable rivers Chapter 25 A In: Schneider, J.C. (Ed.), 2000. Manual of Fisheries Survey Methods II: With Periodic Updates. Michigan Department of Natural Resources, Fisheries Special Report 25, pp. 1–40 (Ann Arbor, MI).

Miltner, R.J., 2010. A method and rationale for deriving nutrient criteria for small rivers and streams in Ohio. Environ. Manag. 45, 842–855.

Miltner, R.J., 2015. Measuring the contribution of agricultural conservation practices to observed trends and recent condition in water quality indicators in Ohio, USA. J. Environ. Qual. 44, 1821–1831.

Miltner, R.J., Rankin, E.T., 1998. Primary nutrients and the biotic integrity of rivers and streams. Freshw. Biol. 40, 145–158.

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. Trans. ASABE 50, 885–900.

Ohio EPA, 1987. Biological Criteria for the Protection of Aquatic Life. Standardized Biological Field Sampling and Laboratory Methods for Assessing Fish and Macroinvertebrate Communities vol. III. Division of Water Quality Monitoring and Assessment, Ohio Environmental Protection Agency (61 pp.).

Ohio EPA, 1999. Association between nutrients, habitat, and the aquatic biota in Ohio rivers and streams. Ohio EPA Technical Bulletin MAS/1999-1-1 (70 pp.).

Ohio EPA, 2010. Ohio Lake Erie Phosphorus Task Force I. Final Report. Ohio Environmental Protection Agency (90 pp.).

Ohio EPA, 2013. Ohio Lake Erie Phosphorus Task Force II. Final Report. Ohio Environmental Protection Agency (96 pp.).

Ohio EPA, 2014. Division of surface water. Ohio 2014 integrated water quality monitoring and assessment report. Ohio Environmental Protection Agency. Ohio Lake Erie Phosphorus Task Force, 2013. Ohio Lake Erie Phosphorus Task Force II Final Report, November, 2013.

Piggott, J.J., Niyogi, D.K., Townsend, C.R., Matthaei, C.D., 2015. Multiple stressors and stream ecosystem functioning: climate warming and agricultural stressors interact to affect processing of organic matter. J. Appl. Ecol. 52, 1126–1134.

Powers, S.M., Bruslsema, T.W., Burt, T.P., Chan, N.I., Elser, J.J., Haygarth, P.M., Howden, N.J.K., Jarvie, H.P., Lyu, Y., Peterson, H., Sharpley, A.N., Shen, J., Worall, F., Zhang, F., 2016. Long-term accumulation and transport of anthropogenic phosphorus in three river basins. Nature Geosci. http://dx.doi.org/10.1038/NGEO2692.

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

Richards, R., Baker, D.B., 2002. Trends in water quality in LEASEQ rivers and streams (northwestern Ohio), 1975–1995. J. Environ. Qual. 31, 90–96.

Richards, R.P., Baker, D.B., Eckert, D.J., 2002. Trends in agriculture in the LEASEQ watersheds, 1975–1995. J. Environ. Qual. 31, 17–24.

Richards, R.P., Baker, D.B., Crumrine, J.P., 2009. Improved water quality in Ohio tributaries to Lake Erie: a consequence of conservation practices. J. Soil Water Conserv. 64, 200–211.

Richards, R.P., Baker, D.B., Crumrine, J.P., Stearns, A.M., 2010. Unusually large loads in 2007 from the Maumee and Sandusky Rivers, tributaries to Lake Erie. J. Soil Water Conserv. 65, 450–462.

Scavia, D., Allan, J.D., Arend, K.K., Bartell, S., 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, P.R., Roberts, J.J., Ruckinski, 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 Lakes Res. 40, 226–246.

Scavia, D., Kalcic, M., Muenich, R.L., Aloysius, N., Arnold, J., Boles, C., Confesor, R., DePinto, J., Gildow, M., Martin, J., Read, J., Redder, T., Robertson, D., Sowa, S., Wang, Y., White, W., Yen, H., 2016. Informing Lake Erie agriculture nutrient management via scenario evaluation. Final Report. University of Michigan Water Center (17 pp.).

Sharpley, A., Jarvie, H.P., Buda, A., May, L., Spears, B., Kleinman, P., 2013. Phosphorus legacy: overcoming the effects of past management practices to mitigate future water quality impairment. J. Environ. Qual. 42, 1308–1326.

Sheldon, A.L., 1968. Species diversity and longitudinal succession in stream fishes. Ecology 49, 193–198.

Smith, D.R., King, K.W., Williams, M.R., 2015. What is causing the harmful algal blooms in Lake Erie? J. Soil Water Conserv. 70, 27–29.

Stumpf, R.P., Wynee, T.T., Baker, D.B., Fahnenstiel, G.L., 2012. Interannual variability of cyanobacterial blooms in Lake Erie. PLoS One 7, e42444.

Taylor, J.M., King, R.S., Pease, A.A., Winemiller, K.O., 2014. Nonlinear response of stream ecosystem structure to low-level phosphorus enrichment. Freshw. Biol. 59, 969–984.

Townsend, C.R., Uhlmann, S.S., Matthaei, C.D., 2008. Individual and combined responses of stream ecosystems to multiple stressors. J. Appl. Ecol. 45, 1810–1819.

Trautman, M.B., 1939. The effects of man-made modification on the fish fauna in Lost and Gordon Creeks, Ohio, between 1887–1938. Ohio J. Sci. 39, 275–288.

Trautman, M.B., 1981. The Fishes of Ohio. Ohio State University Press, Ohio.

Trautman, M.B., Gartman, D.K., 1974. Re-evaluation of the effects of man-made modifications on Gordon Creek between 1887 and 1973 and especially as regards its fish fauna. Ohio J. Sci. 74, 162–173.

USDA NRCS, 2011. Conservation Effects Assessment Project. Assessment of the Effects of Conservation Practices on Cultivated Cropland in the Great Lakes Region. United States Department of Agricultural Natural Resources Conservation Service (172 pp.).

Vaz, S., Martin, C.S., Eastwood, P.D., Ernande, B., Carpentier, A., Meaden, G.J., Coppin, F., 2008. Modelling species distributions using regression quantiles. J. Appl. Ecol. 45, 204–217.

Wang, L., Lyons, J., Kanehl, P., 2002. Effects of watershed best management practices on habitat and fish in Wisconsin streams. J. Am. Water Resour. Assoc. 38, 663–680.

Wang, L., Lyons, J., Kanehl, P., 2006. Habitat and fish responses to multiple agricultural best management practices in a warm water stream. J. Am. Water Resour. Assoc. 42, 1047–1062 (Environ. Manage. 39, 194–212).

Wang, X., Kannan, N., Santhi, C., Potter, S.R., Williams, J.R., Arnold, J.G., 2011. Integrating APEX output for cultivated cropland with SWAT simulation for regional modelling. Trans. ASABE 54, 1281–1298.

Wang, L., Robertson, D.M., Garrison, P.J., 2007. Linkages Between Nutrients and Assemblages of Macroinvertebrates and Fish in Wadeable Streams: Implications for Nutrient Criteria Development.

Waters, T.F., 1995. Sediment in Streams: Sources, Biological Effects, and Control. AFS Monographs, Bethesda, MD.

Please cite this article as: Keitzer, S.C., et al., Thinking outside of the lake: Can controls on nutrient inputs into Lake Erie benefit stream conservation in its watershed? J. Great Lakes Res. (2016), http://dx.doi.org/10.1016/j.jglr.2016.05.012