EFFECTS OF CLIMATE, LAND USE, AND LAND MANAGEMENT ON PHOSPHORUS CYCLING AND WATER QUALITY IN THE YAHARA WATERSHED

by

Melissa M. Motew

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

(Environment & Resources)

at the

UNIVERSITY OF WISCONSIN–MADISON

2017

Date of final oral examination: May 8, 2017

The dissertation is approved by the following members of the Final Oral Committee:

Christopher J. Kucharik, Professor, Agronomy
Stephen R. Carpenter, Professor, Zoology
Steven P. Loheide II, Associate Professor, Civil & Environmental Engineering
Pete Nowak, Emeritus, Gaylord Nelson Institute for Environmental Studies
Monica G. Turner, Professor, Zoology

Abstract

EFFECTS OF CLIMATE, LAND USE, AND LAND MANAGEMENT ON PHOSPHORUS CYCLING AND WATER QUALITY IN THE YAHARA WATERSHED

Melissa M. Motew

Under the supervision of Associate Professor Christopher J. Kucharik

At the University of Wisconsin–Madison

Freshwater resources are critical to society and the biosphere, yet accelerated eutrophication due to phosphorus (P) enrichment plagues watersheds throughout the world. In agricultural regions where there is a high need for P to support crop and livestock production, mismanagement of nutrients has led to an excessive supply of P on the landscape where it is transported to waterways via surface runoff. Land use, land management, and climate all represent important drivers of the P cycle, capable of altering its supply, transport pathways, and fluxes across the landscape. As such, rapid change in these drivers promises to bring considerable uncertainty to the future sustainability of freshwater resources.

My dissertation consists of three separate inquiries into how changing climate and land use/land management (LULM) may affect P cycling and surface water quality in the future. Using a process-based modeling framework developed for the Yahara Watershed, Wisconsin, I first investigate how historical nutrient management practices, represented by "legacy P", have influenced current and future conditions of lake water quality. A model experiment that tests five alternative amounts of legacy P stored in soils and sediments indicates that legacy P has a significant and long-lasting effect on water quality in the lakes as well as in runoff and streams. The downstream Yahara lakes (Monona, Waubesa, and Kegonsa) are less sensitive to changes in legacy P than is Lake Mendota, due to the dominant influence of riverine loads from upstream lakes. Model runs show that there is a substantial overabundance of soil P in the watershed, roughly four times more than agronomic recommendations. Results also suggest that terrestrial P supply may interact synergistically with extreme rainfall in affecting lake total P concentration. Ultimately, this chapter emphasizes the important role that past land management has on building up P reserves in soils and sediments, and how that can affect water quality outcomes well into the future. Drawing down P reserves, for example through recovery and recycling of P in soils, channels, and manures, promises significant improvements in water quality and may help buffer against the effects of climate change.

The second inquiry of the dissertation examines the role that extreme precipitation plays in affecting P transport and water quality in agricultural regions. Using a 2x2 factorial experimental design, I explicitly test the models for an interaction between terrestrial P supply and precipitation intensity on surface water quality at three spatial scales. The results of this inquiry indicate that a significant synergistic interaction exists

between the two drivers in affecting dissolved P concentration at field and stream scales, as well as total P concentration at field, stream, and lake scales. The synergy arises from nonlinear dependencies between P stored in manure and daily rainfall and runoff amounts, as represented in the models. In affecting dissolved P losses, the synergy has important ecological consequences since dissolved P is highly bioavailable. Overall, the results suggest that high levels of terrestrial P supplied as manure can exacerbate water quality problems in the future as the intensity of rainfall increases along with a greater frequency of heavy rainfall events. Conversely, lowering manure P supply may help improve the resilience of freshwater ecosystems to extreme events.

The final inquiry examines the relative influences of climate and LULM in affecting outcomes of water quality, and the most important underlying biophysical mechanisms. Using long-term future scenarios developed for the Yahara Watershed (yahara2070.org), I examine the relative influences of LULM and climate on water quality over the span of six decades and at three spatial scales: field, stream, and lake. Results show that both LULM and climate have important roles in driving outcomes at all scales of the watershed. The most important LULM mechanism controlling water quality at all scales is the average field scale mass balance of P inputs and outputs, which supersedes erosion risk and percent land cover type in the watershed. Phosphorus balance is a stronger predictor of water quality at the field scale but less so at stream and lake scales, where the influence of weather variability becomes more important. This finding, which underscores the dominant role of climate in driving nutrient fluxes within the hydrologic network, suggests an inherent limitation for field scale LULM to influence water quality within streams and lakes. Nevertheless, reducing over-application of P

throughout the watershed is an effective management strategy under the four climates investigated, even during decades with wetter conditions and more frequent extreme precipitation events.

The results of my dissertation emphasize that the overabundance of P within the YW is a dominant biophysical control of surface water quality across spatial and temporal scales. Climate change will present a formidable challenge to the management of freshwater resources, yet strategies that focus on reducing legacy P and balancing P budgets on farms promise meaningful improvements in stream and lake conditions, as well as protective benefits.

Acknowledgements

I have many people to thank for their support and contributions to this work. First and foremost is my advisor, Chris Kucharik, who has also been a mentor and friend. Chris has guided and supported me throughout graduate school and the research process, continually helping me to see the big picture when I've been caught in the modeling weeds. I am grateful for the room he has given me to pursue my own technical solutions, and the encouragement to formulate my own research ideas and see them from start to finish.

Secondly, my committee has been enormously helpful in shaping my research. In addition to leading by example in their own particular fields of ecology and sociology, they have each provided me with their own unique insights into the research process as well as this fascinating system that is the Yahara Watershed. Steve Carpenter has greatly influenced my ideas and understanding about ecosystems, including ways to model, measure, and analyze them. Monica Turner has shown me the importance of defining and understanding ecosystem patterns and processes across scales. Steve Loheide has helped me understand the complexities of the hydrologic system, and to consider the assumptions and limitations of my models. Pete Nowak has taught me that balancing the needs for agriculture and clean water is indeed a wicked problem, and that understanding farmers and the challenges they face is critical to conducting meaningful research on water quality.

It has been a great honor and pleasure to be a part of the Water Sustainability and Climate Project, or as I have often referred to this group, the "Dream Team". I learned

the power of interdisciplinary collaboration firsthand, and had a blast in doing so. Each chapter of my thesis has been a collaborative undertaking, which has benefited not just the research itself but my own development as a scientist. To all on this team I am grateful.

Thanks also to Peter Vadas who has helped greatly with model implementation and interpretation, and to Shawn Essler who took me to meet farmers in the watershed and hear their stories. I am also thankful to members of the Kucharik lab group, who over the years have provided a plethora of ideas as well as great camaraderie. I also thank my friends here in Madison, who have shown me the power of community and friendship in this amazing city.

Despite not wanting me to leave home for Wisconsin, my family and friends back in New England have provided me all the support that I could have asked for, and then some. I am therefore extremely grateful to my parents, siblings, best friend, and nieces. Finally, I thank my partner Mark, who has been my champion throughout graduate school. He has been with me through each challenging stage of this adventure, and has let me lean on him at each one.

This work is dedicated to my late grandmothers, Mary Willis and Lillian Motew.

Table of Contents

Chapter 1

Introduction

1.1 Motivation

Freshwater ecosystems such as streams, rivers, lakes, and wetlands provide numerous benefits to society and the biosphere, such as water for drinking, use in agriculture and industry, and support for wildlife and biodiversity (Wilson and Carpenter 1999). For billions of people living in the world, renewable freshwater resources are limited, and water quality impairment exacerbates issues of scarcity (Oki et al. 2001). A prominent form of water quality impairment in freshwater ecosystems is eutrophication, which is often caused by phosphorus (P) over-enrichment in P-limited water bodies (Schindler 1977). Eutrophication has many undesirable symptoms which render freshwater resources unusable by humans and uninhabitable by many organisms. Consequences include an overgrowth of algae and aquatic plants, low water clarity, foul tastes and odors, oxygen depletion, fish kills, extirpation of native plants, impaired recreational opportunities and property value, as well the frequent occurrence of toxic cyanobacteria blooms (Carpenter et al. 1998, Dodds et al. 2009). The problem of excessive phosphorus loading from anthropogenic sources to aquatic ecosystems plagues many regions of the world, including North America (Ardón et al. 2010; Dubrovsky et al. 2010), Central and South America (Quiros 1991; Salas and Martino 1991), the U.K. and Europe (Withers et al. 2001, Ulén et al. 2007, Azevedo et al. 2015), India (Bhade et al. 2002), Asia (An and Jones 2000, Dai and others 2010, Chen and others 2012), Africa (Thornton 1980), New Zealand (Pridmore et al. 1985), and elsewhere (Smith 2003).

Massive alteration to the global P cycle has led to an overabundance of P in soils in some regions, where it can be transported to waterways via erosion and surface runoff (Elser and Bennett 2011). While point sources are smaller and easier to control, and have been somewhat curtailed with the passage of the Clean Water Act and similar laws, nonpoint sources remain significant contributors of P to freshwater ecosystems (Sharpley et al. 1994, Bennett et al. 2001). Most notable are agricultural lands which lose roughly 8 million tons of P annually to the environment (Cordell et al. 2009). With a rapidly growing global population, the need to balance the high demand for P to support agriculture with the need for clean water represents a "wicked" societal challenge for the 21st century (Reed and Kasprzyk 2009).

Land use and land management (LULM) are important drivers of surface water quality, capable of altering P supply, cycling, and transport pathways (Soranno et al. 1996). Whereas some landscape drivers of surface water quality are relatively constant through time, like soil type and topography, other factors can change more quickly in response to human activities (Vanni et al. 2001), such as land use type and intensity (Johnson et al. 1997, Gergel et al. 2002, Smith et al. 2013), composition and configuration of land cover (O'Neill et al. 1997, Clément et al. 2017), management practices (Sharpley et al. 1994, Crossman et al. 2016), and engineering structures (Jarvie et al. 2006, Gentry et al. 2007).

Within agricultural settings, nutrient management decisions including application rates, methods, and timing can significantly alter surface pools of P that are subject to transport via overland flow, and ultimately the fluxes of P passing from terrestrial to aquatic ecosystems. Decades of excessive P fertilizer use spurred by the Green

Revolution have led to a build up of "legacy P" in soils and sediments that can provide a steady source of P to water bodies over prolonged periods, at least years to decades, and perhaps as long as centuries (Bennett et al. 1999, Hamilton 2012, Powers et al. 2016). The precise effect of legacy P on surface water quality over a given time period is difficult to quantify based on observations alone since many interacting factors may obscure relationships (Hamilton 2012, Gillon et al. 2016).

In recent decades, efforts to control P runoff from agricultural lands including the use of best management practices (BMPs) have shown success at the field scale but little improvement in downstream ecosystems (Meals et al. 2010, Sharpley et al. 2013). Thus there exists a research gap in understanding changes in water quality due to LULM across spatial scales, including questions related to the magnitude and timing of biogeochemical and ecological responses to management practices (Wood et al. 2005, Sharpley et al. 2009). Other agricultural trends affect LULM as well, such as changes in available agricultural land (Gillon et al. 2016), intensification of livestock operations (e.g. the proliferation of concentrated animal feeding operations, CAFOs) (Smith et al. 2013), globalization and changes in agricultural trade (MacDonald et al. 2015), and adoption of new technologies (Schewe and Stuart 2015). The future sustainability of freshwater resources is thus highly uncertain since each trend may affect P cycling and water quality in harmful or beneficial ways.

In addition to LULM, climate change is expected to substantially alter freshwater quality by increasing fluxes of water and P as well as bringing warmer temperatures that favor the growth of harmful algal blooms (Gkelis et al. 2014, Michalak 2016). Nutrient transport can be significantly altered as the timing and magnitude of runoff and soil

moisture, lake levels, groundwater availability, and river discharge regimes change (Bates et al. 2008, Crossman et al. 2016). Regions where extreme rain events deliver a disproportionate amount of annual sediment and P loading to water bodies may be particularly vulnerable to increases in rainfall intensity (Haygarth and Jarvis 1997; Royer et al. 2006; Carpenter et al. 2014; Gonzalez-Hidalgo 2013). Combined with excess reserves of P stored in soils and sediments, an amplified hydrologic system could pose the risk for a "chemical time bomb" (Stigliani et al. 1991). Thus there is a great need to understand to what extent local management options can counter the risks of climate change, and how those strategies should be focused within watersheds.

While future projections of climate and LULM are fraught with uncertainty (Carpenter et al. 2006), changes in both drivers are likely to affect water quality outcomes across spatial and temporal scales in the future. In order to protect and increase the resilience of freshwater ecosystems, a better understanding is needed of the threats and opportunities posed by changing climate, land use, and land management. My dissertation focuses on several knowledge gaps related to climate and LULM on P cycling and surface water quality, using the Yahara Watershed (YW) of southern Wisconsin as a microcosm for many of the issues facing agricultural watersheds around the world. For example, the YW has an overabundance of P stored in its soils, sediments, and water bodies due to decades of P imbalance at the watershed scale (Bennett et al. 1999, Kara et al. 2011); has seen a lack of water quality improvements despite decades of management interventions (Lathrop et al. 2013); has been witness to rapid changes in land use and management including intensification of dairy agriculture (Gillon et al. 2016); exhibits high sensitivity to extreme weather events (Carpenter et al. 2014); and has

seen a rise in annual precipitation and frequency of extreme precipitation events (Kucharik et al. 2010).

Dynamic simulation models have emerged as important research tools for understanding and isolating the impacts of anthropogenic change drivers on ecosystem structure, function, and services (Turner and Carpenter 2017). Though complex models have considerable uncertainty (Oreskes 2003), when tested and used appropriately they can be useful tools in assessing many aspects of social-ecological systems that cannot be addressed with measurements alone (Seidl 2017). For example, in the study of watershed systems, simulation models provide an opportunity to investigate interactions of hydrologic fluxes, biogeochemical cycling, and human modifications of the landscape across a range of spatial and temporal scales. Process-based models, as opposed to empirical models, may be better suited for investigating systems undergoing novel transitions, such as those related to climate and/or land use, since they are less dependent on prior conditions (Gustafson 2013).

As part of the Water Sustainability and Climate Project, I helped develop a process-based watershed modeling framework that simulates dynamic responses within natural and agricultural ecosystems of the YW. Driven by inputs of climate, soil type, topography, and LULM, the modeling framework simulates cycles of carbon, water, energy, nitrogen, and phosphorus. The framework enables quantification of ecosystem services that include surface and groundwater quality and quantity, flood protection, food and forage production, soil retention, and climate regulation. The details of model development, calibration, and validation are provided in the Appendices of Chapter 2.

My dissertation uses the modeling framework to study surface water quality as indicated by P loads and concentrations at field, stream, and whole lake scales.

1.2 Objectives

Using the modeling framework, my research investigates three overarching questions related to how climate and LULM affect cycling and transport of P within the YW. These include:

(1) How have historical nutrient management practices, represented by legacy P, influenced current and future conditions of lake water quality? Here I conduct a model experiment that quantifies how different amounts of legacy P built up in soils and streambed sediments affect lake water quality over a recent 27 year time period. This study is meant to provide understanding into how sensitive the Yahara lakes are to P storage in the watershed, the prolonged impacts that nutrient management practices can have on the lakes, and whether or not management efforts need to focus on removing or reducing the overabundance of P in addition to just preventing its loss in runoff.

(2) How might changes in extreme precipitation interact with an abundant supply of terrestrial P to affect surface water quality*?* While also considered briefly in Chapter 2, Chapter 3 focuses more directly on how extreme precipitation might interact with different amounts of stored P in affecting surface water quality. I use a 2x2 factorial experimental design to test for an interaction between (high/low) terrestrial P supply and (high/low) precipitation intensity in affecting P indicator variables at field, stream, and lake scales. I consider three forms of P including dissolved, sediment, and total, and I also examine three possible terrestrial sources of P, including soils, manure, and fertilizer.

Given current trends and projections for more frequent heavy precipitation events (Kucharik et al. 2010), Chapter 3 is intended to help build understanding of the overall vulnerability of freshwater bodies to changes in extreme precipitation, and provide insight into how nutrient and watershed management interventions may help combat the challenges that climate change may bring.

(3) What are the relative influences of climate and LULM in affecting outcomes of water quality across spatial scales of a watershed? What are the most important underlying biophysical mechanisms? Using the Yahara 2070 integrated scenarios for the YW (yahara2070.org), Chapter 4 examines the relative influences of LULM and climate on water quality over the span of six decades and at field, stream, and lake scales. Because climate is driven by human activities occurring at a global scale, LULM represents an important avenue by which managers and decision makers can affect outcomes of water quality. Thus this chapter aims to provide insight into how much power exists at a local level to counteract climate change and influence the long-term fate of freshwater quality. Using linear regression, this chapter also investigates which biophysical aspects of LULM are the most important for influencing water quality at each of the three spatial scales. This chapter ultimately seeks to identify which aspects of LULM offer our best hope for improving water quality as we progress into an uncertain future.

While Chapters 2, 3, and 4 investigate the biophysical aspects of the watershed system in response to changing drivers, Chapter 5 *Conclusions* attempts to put these findings into a greater social-ecological context for the YW as well as other watersheds. First, a summary of the major findings of the dissertation are provided in Chapter 5.1

Summary. In Chapter 5.2 *Synthesis*, I explore the implications of the biophysical research findings for management within the YW. From there, Chapter 5.3 *Broader Contributions* seeks to relate the overall findings and conclusions to management and watershed modeling efforts beyond the YW. Chapter 5.4 *Limitations* discusses key shortcomings to the dissertation, including areas of future model improvement and research. Finally, in Chapter 5.5 *A Fifth Scenario*, I propose a scenario building exercise for the YW which integrates the key biophysical findings of my research, the modeling tools, and the knowledge and capacity of other actors in the watershed. This interdisciplinary exercise is intended to generate a management-focused scenario for the YW that is actionable, effective, and helps to increase the resilience of both farms and freshwater resources to future changes in climate and other pressures.

References

An K-G, Jones JR. 2000. Factors regulating bluegreen dominance in a reservoir directly influenced by the Asian monsoon. Hydrobiologia 432:37–48.

Ardón M, Morse JL, Doyle MW, Bernhardt ES. 2010. The Water Quality Consequences of Restoring Wetland Hydrology to a Large Agricultural Watershed in the Southeastern Coastal Plain. Ecosystems 13:1060–78.

Azevedo LB, van Zelm R, Leuven RSEW, Hendriks AJ, Huijbregts MAJ. 2015. Combined ecological risks of nitrogen and phosphorus in European freshwaters. Environ Pollut 200:85–92.

Bates B, Kundzeqics ZW, Wu S, Palutikof JP (eds). 2008. Climate change and water, Tech. Pap. VI Intergovermental Panel Clim. Change, IPCC Secretariat, Geneva, Switzerland. https://www.ipcc.ch/pdf/technical-papers/climate-change-water-en.pdf

Bennett EM, Carpenter SR, Caraco NF. 2001. Human impact on erodable phosphorus and eutrophication: A global perspective. Bioscience 51:227–34.

Bennett EM, Reed-Andersen T, Houser JN, Gabriel JR, Carpenter SR. 1999. A Phosphorus Budget for the Lake Mendota Watershed. Ecosystems 2:69–75.

Bhade C, Unni KS, Bhade S. 2002. Limnology and eutrophication of Tawa Reservoir, MP State, India. Verh Int Ver Theor Angew Limnol 27:3632–5.

Carpenter SR, Bennett EM, Peterson GD. 2006. Scenarios for ecosystem services: an overview. Ecol Soc 11:29.

Carpenter SR, Booth EG, Kucharik CJ, Lathrop RC. 2014. Extreme daily loads: role in annual phosphorus input to a north temperate lake. Aquat Sci 77:71–9.

Carpenter SR, Caraco NF, Correll DL, Howarth RW, Sharpley AN, Smith VH. 1998. Nonpoint pollution of surface waters with phosphorus and nitrogen. Ecol Appl 8:559–68.

Chen Y, Liu R, Sun C, Zhang P, Feng C, Shen Z. 2012. Spatial and temporal variations in nitrogen and phosphorous nutrients in the Yangtze River Estuary. Mar Pollut Bull 64:2083–9.

Clément F, Ruiz J, Rodríguez MA, Blais D, Campeau S. 2017/1. Landscape diversity and forest edge density regulate stream water quality in agricultural catchments. Ecol Indic 72:627–39.

Cordell D, Drangert J-O, White S. 2009/5. The story of phosphorus: Global food security and food for thought. Glob Environ Change 19:292–305.

Crossman J, Futter MN, Palmer M, Whitehead PG, Baulch H, Woods D, Jin L, Oni S, Dillon PJ. 2016. The effectiveness and resilience of phosphorus management practices in the Lake Simcoe Watershed, Ontario, Canada. J Geophys Res Biogeosci:2015JG003253.

Dai Z, Du J, Zhang X, Su N, Li J. 2010. Variation of Riverine Material Loads and Environmental Consequences on the Changjiang (Yangtze) Estuary in Recent Decades (1955- 2008)†. Environ Sci Technol 45:223–7.

Dodds WK, Bouska WW, Eitzmann JL, Pilger TJ, Pitts KL, Riley AJ, Schloesser JT, Thornbrugh DJ. 2009. Eutrophication of U.S. freshwaters: analysis of potential economic damages. Environ Sci Technol 43:12–9.

Dubrovsky NM, Burow KR, Clark GM, Gronberg JM, Hamilton PA, Hitt KJ, Mueller DK, Munn MD, Nolan BT, Puckett LJ, Others. 2010. The quality of our Nation's waters-Nutrients in the Nation's streams and groundwater, 1992--2004. US Geological Survey Circular 1350. http://pubs.usgs.gov/circ/1350/

Elser J, Bennett E. 2011. Phosphorus cycle: A broken biogeochemical cycle. Nature 478:29–31.

Gentry LE, David MB, Royer TV, Mitchell CA, Starks KM. 2007. Phosphorus transport pathways to streams in tile-drained agricultural watersheds. J Environ Qual 36:408–15.

Gergel SE, Turner MG, Miller JR, Melack JM, Stanley EH. 2002. Landscape indicators of human impacts to riverine systems. Aquat Sci 64:118–28.

Gillon S, Booth EG, Rissman AR. 2016. Shifting drivers and static baselines in environmental governance: challenges for improving and proving water quality outcomes. Regional Environ Change 16:759–75.

Gkelis S, Papadimitriou T, Zaoutsos N, Leonardos I. 2014. Anthropogenic and climate-induced change favors toxic cyanobacteria blooms: Evidence from monitoring a highly eutrophic, urban Mediterranean lake. Harmful Algae 39:322–33.

Gonzalez-Hidalgo JC, Batalla RJ, Cerda A. 2013/3. Catchment size and contribution of the largest daily events to suspended sediment load on a continental scale. Catena 102:40–5.

Gustafson EJ. 2013. When relationships estimated in the past cannot be used to predict the future: using mechanistic models to predict landscape ecological dynamics in a changing world. Landsc Ecol 28:1429–37.

Hamilton SK. 2012. Biogeochemical time lags may delay responses of streams to ecological restoration. Freshw Biol 57:43–57.

Haygarth PM, Jarvis SC. 1997/1. Soil derived phosphorus in surface runoff from grazed grassland lysimeters. Water Res 31:140–8.

Jarvie HP, Neal C, Withers PJA. 2006. Sewage-effluent phosphorus: a greater risk to river eutrophication than agricultural phosphorus? Sci Total Environ 360:246–53.

Johnson L, Richards C, Host G, Arthur J. 1997. Landscape influences on water chemistry in Midwestern stream ecosystems. Freshw Biol 37:193–208.

Kara EL, Heimerl C, Killpack T, Van de Bogert MC, Yoshida H, Carpenter SR. 2011. Assessing a decade of phosphorus management in the Lake Mendota, Wisconsin watershed and scenarios for enhanced phosphorus management. Aquat Sci 74:241–53.

Kucharik CJ, Serbin SP, Vavrus S, Hopkins EJ, Motew MM. 2010. Patterns of Climate Change Across Wisconsin From 1950 to 2006. Phys Geogr 31:1–28.

Lathrop RC, Carpenter SR. 2013. Water quality implications from three decades of phosphorus loads and trophic dynamics in the Yahara chain of lakes. Inland Waters 4:1–14.

MacDonald GK, Brauman KA, Sun S, Carlson KM, Cassidy ES, Gerber JS, West PC. 2015. Rethinking Agricultural Trade Relationships in an Era of Globalization. Bioscience 65:275–89.

Meals DW, Dressing SA, Davenport TE. 2010. Lag time in water quality response to best management practices: a review. J Environ Qual 39:85–96.

Michalak AM. 2016. Study role of climate change in extreme threats to water quality. Nature 535:349–50.

Oki T, Agata Y, Kanae S, Saruhashi T, Yang D, Musiake K. 2001. Global assessment of current water resources using total runoff integrating pathways. Hydrol Sci J 46:983–95.

O'Neill RV, Hunsaker CT, Jones KB, Riitters KH. 1997. Monitoring environmental quality at the landscape scale. Bioscience 47:513–9.

Oreskes N. 2003. The Role of Quantitative Models in Science. In: Canham CD, Cole JJ, Lauenroth WK, editors. Models in Ecosystem Science. Princeton and Oxford: Princeton University Press. pp 13–31.

Powers SM, Bruulsema TW, Burt TP, Chan NI, Elser JJ, Haygarth PM, Howden NJK, Jarvie HP, Lyu Y, Peterson HM, Sharpley AN, Shen J, Worrall F, Zhang F. 2016. Long-term accumulation and transport of anthropogenic phosphorus in three river basins. Nat Geosci 9:353–6.

Pridmore RD, Vant WN, Rutherford JC. 1985. Chlorophyll-nutrient relationships in North Island lakes (New Zealand). Hydrobiologia 121:181–9.

Quiros R. 1991. Empirical relationships between nutrients, phyto-and zooplankton and relative fish biomass in lakes and reservoirs of Argentina. Verh Int Ver Theor Angew Limnol 24:1198–206.

Reed J, Kasprzyk PM. 2009. Water Resources Management: The Myth, the Wicked, and the Future. Journal of Water Resources Planning and Management 135:411–3.

Royer TV, David MB, Gentry LE. 2006. Timing of riverine export of nitrate and phosphorus from agricultural watersheds in Illinois: implications for reducing nutrient loading to the Mississippi River. Environ Sci Technol 40:4126–31.

Salas HJ, Martino P. 1991. A simplified phosphorus trophic state model for warm-water tropical lakes. Water Res 25:341–50.

Schewe RL, Stuart D. 2015. Diversity in agricultural technology adoption: How are automatic milking systems used and to what end? Agric Human Values 32:199–213.

Schindler DW. 1977. Evolution of phosphorus limitation in lakes. Science 195:260–2.

Seidl R. 2017. To Model or not to Model, That is no Longer the Question for Ecologists. Ecosystems 20:222–8.

Sharpley A, Jarvie HP, 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 42:1308–26.

Sharpley AN, Chapra SC, Wedepohl R, Sims JT, Daniel TC, Reddy KR. 1994. Managing agricultural phosphorus for protection of surface waters - issues and options. J Environ Qual 23:437–51.

Sharpley AN, Kleinman PJA, Jordan P, Bergström L, Allen AL. 2009. Evaluating the success of phosphorus management from field to watershed. J Environ Qual 38:1981–8.

Smith AP, Western AW, Hannah MC. 2013. Linking water quality trends with land use intensification in dairy farming catchments. J Hydrol 476:1–12.

Smith AP, Western AW, Hannah MC. 2013. Linking water quality trends with land use intensification in dairy farming catchments. J Hydrol 476:1–12.

Smith VH. 2003. Eutrophication of freshwater and coastal marine ecosystems a global problem. Environ Sci & Pollut Res 10:126–39.

Soranno PA, Hubler SL, Carpenter SR, Lathrop RC. 1996. Phosphorus loads to surface waters: a simple model to account for spatial pattern of land use. Ecol Appl 6:865–78.

Stigliani WM, Doelman P, Salomons W, Schulin R, Smidt GRB, Van der Zee S. 1991. Chemical Time Bombs: Predicting the Unpredictable. Environment 33:4–30.

Thornton JA. 1980. Comparison of the summer phosphorus loadings to three Zimbabwean water-supply reservoirs of varying trophic states. Water S A 6:163–70.

Turner MG, Carpenter SR. 2017. Ecosystem Modeling for the 21st Century. Ecosystems 20:211–4.

Ulén B, Bechmann M, Fölster J, Jarvie HP, Tunney H. 2007. Agriculture as a phosphorus source for eutrophication in the north-west European countries, Norway, Sweden, United Kingdom and Ireland: a review. Soil Use Manage 23:5–15.

Vanni MJ, Renwick WH, Headworth JL, Auch JD, Schaus MH. 2001. Dissolved and particulate nutrient flux from three adjacent agricultural watersheds: A five-year study. Biogeochemistry 54:85–114.

Wilson MA, Carpenter SR. 1999. Economic valuation of freshwater ecosystem services in the United States: 1971–1997. Ecol Appl 9:772–83.

Withers PJA, Edwards AC, Foy RH. 2001. Phosphorus cycling in UK agriculture and implications for phosphorus loss from soil. Soil Use Manage 17:139–49.

Wood FL, Heathwaite AL, Haygarth PM. 2005. Evaluating diffuse and point phosphorus contributions to river transfers at different scales in the Taw catchment, Devon, UK. J Hydrol 304:118–38.

Chapter 2

The influence of legacy P on lake water quality in a Midwestern agricultural watershed

Motew, M., Chen, X., Booth, E.G., Carpenter, S.R., Pinkas, P., Zipper, S.C., Loheide, S.P., Donner, S.D., Tsuruta, K., Vadas, P.A., Kucharik, C.J., 2017. The Influence of Legacy P on Lake Water Quality in a Midwestern Agricultural Watershed. Ecosystems In press, 1–15. doi: 10.1007/s10021-017-0125-0

Abstract

Decades of fertilizer and manure applications have led to a buildup of phosphorus (P) in agricultural soils and sediments, commonly referred to as legacy P. Legacy P can provide a long-term source of P to surface waters where it causes eutrophication. Using a suite of numerical models we investigated the influence of legacy P on water quality in the Yahara Watershed of southern Wisconsin, USA. The suite included Agro-IBIS, a terrestrial ecosystem model; THMB, a hydrologic and nutrient routing model; and the Yahara Water Quality Model which estimates water quality indicators in the Yahara chain of lakes. Using five alternative scenarios of antecedent P storage (legacy P) in soils and channels under historical climate conditions, we simulated outcomes of P yield from the landscape, lake P loading, and three lake water quality indicators. Legacy P had a significant effect on lake loads and water quality. Across the five scenarios for Lake Mendota, the largest and most upstream lake, average P yield (kg ha^{-1}) varied by -41 to +22%, P load (kg y^{-1}) by -35 to +14%, summer total P (TP) concentration (mg L^{-1}) by -25 to +12%, Secchi depth (m) by -7 to +3%, and the probability of hypereutrophy by -67 to

+34%, relative to baseline conditions. The minimum storage scenario showed that a 35% reduction in present-day loads to Lake Mendota corresponded with a 25% reduction in summer TP and smaller reductions in the downstream lakes. Water quality was more vulnerable to heavy rainfall events at higher amounts of P storage and less so at lower amounts. Increases in heavy precipitation are expected with climate change, therefore water quality could be protected by decreasing P reserves.

2.1. Introduction

Freshwater eutrophication caused by phosphorus (P) enrichment is a serious environmental problem throughout the world, characterized by low water clarity, potentially toxic algae blooms, oxygen depletion, fish kills, and loss of economic benefits associated with clean water (MA 2005; Smith and others 2006; Bennett and Schipanski 2013). In many watersheds, a history of agricultural land use has resulted in elevated levels of P in surface soils, ditches, riparian zones, wetlands, and stream and lake sediments (Sharpley 2016). Accumulation of P can occur over finite periods when P inputs exceed agricultural demand, and this accumulation can continue to mobilize long after inputs decline (Powers and others 2016). A global P budget performed in 2001 indicated that the increase in net P storage in terrestrial and freshwater ecosystems was at least 75% greater than preindustrial levels, with a large portion occurring in agricultural soils (Bennett and others 2001). Regions with high densities of livestock operations are especially prone to buildup of soil P when manure is applied to meet nitrogen demand and subsequently supplies P in excess of crop removal (Nowak and others 1998; Staver and Brinsfield 2001; Sturgul and Bundy 2004).

Legacy P buildup can act as a long-term source of P to surface waters, in many cases delaying intended reductions in catchment P fluxes associated with best management practices (BMPs) (Meals and others 2010; Hamilton 2012; Sharpley and others 2013). The release of legacy P into surface waters can overwhelm factors favoring water quality improvement, making it difficult to distinguish the effects of current conservation measures from historical land management (Hamilton 2012; Jarvie and others 2013a; Sharpley and others 2013). Drawdown by crops may be the best mitigation

measure for legacy P, but can take decades or longer depending on how fertilizer inputs and agricultural exports are managed (McCollum 1991; Schulte and others 2010).

In addition to catchment-scale problems associated with legacy P, there is also widespread recognition of a broken global P cycle that has arguably created the legacy P problem. Finite reserves of geologic phosphate rock are being mined at unsustainable rates while also accumulating in parts of the world where it causes significant degradation to aquatic ecosystems, such as the Midwest U.S. (Elser and Bennett 2011). Proposed solutions for managing P resources, as well as legacy P, have focused on closing the global P budget, attempting to connect the surplus of legacy P to regions where P is scarce (Childers and others 2011; Sattari and others 2012; Metson and others 2015). Methods to recover P from human waste, wastewater, manure, agricultural residues, as well as soils are being developed (Elser and Bennett 2011; Kahiluoto and others 2015; Wilfert and others 2015). The recovery of legacy P could help substitute manufactured fertilizers, preserve critical reserves of phosphate rock, as well as improve water quality (Rowe and others 2015; Wolfe and others 2016).

Lake eutrophication is a major concern in the Yahara Watershed (YW) of southern Wisconsin, home to a thriving dairy industry, the city of Madison, and a chain of four highly valued lakes (Stumborg and others 2001). The largest and furthest upstream lake, Lake Mendota, receives the greatest average annual direct P loads of the four and has a large influence on river loading to the downstream lakes that include Monona, Waubesa, and Kegonsa (Lathrop and Carpenter 2013). Eutrophication has plagued the Yahara lakes since the mid-1800s, when the landscape was first transformed from native vegetation to farms and urban areas (Carpenter and others 2006). Until the

mid-20th century, sewage effluents were the greatest source of P to the lakes, but since wastewater diversion in 1971, non-point sources have dominated, most importantly agricultural runoff (Soranno and others 1996; Lathrop and others 1998; Lathrop 2007). As in other agricultural watersheds, there is strong evidence that inputs of P to the YW exceed outputs, and that soil P levels are significantly greater than needed to meet plant demand and sustain crop yields (Bennett and others 1999; Reed-Andersen and others 2000; Kara and others 2011). According to a P budget conducted in the Lake Mendota watershed in the late 1990s, only 4.6% of annual P outputs from the watershed were exported to Lake Mendota. The majority of P outputs mostly consisted of exported corn and dairy products (Bennett and others 1999). The authors estimated drawdown of soil P to 1974 levels by crops could take decades to centuries, depending on how inputs and outputs are managed (Bennett and others 1999). An updated P budget showed that inputs to the Lake Mendota watershed have likely declined since the mid-90s, but have continued to exceed outputs (Kara and others 2011).

Loads to Lake Mendota have not changed over the past three decades despite significant nutrient reduction interventions in that subwatershed (Genskow and Rumery Betz 2012; Lathrop and Carpenter 2013). Loads to Mendota vary year to year due to weather, but average loading from direct drainage sources has fluctuated around 29,600 kg y^{-1} over the 1976-2008 period, and the median summertime lake TP concentration has exceeded the mesotrophic threshold of 0.024 mg L^{-1} in most years (Lathrop and Carpenter 2013). The failure of management interventions to improve water quality has been blamed on the counteractive influence of other factors affecting water quality over this time period. These include an increase in annual precipitation and increased

frequency of extreme rainfall, continued intensification of dairy production, and a persistent legacy P problem (Gillon and others 2016; Rissman and Carpenter 2015). It can be argued that without management interventions water quality would have declined over this period.

Prior studies have identified the important role that legacy P plays in affecting water quality and the long-term challenges it poses for water resource management (Haygarth and others 2014; Powers and others 2016). However, establishing causal links between factors affecting water quality, like legacy P, and their outcomes, is challenging, especially at the catchment scale. This is because many factors may simultaneously affect water quality and/or its measurement, and are subject to change over time (Edwards and Withers 1998; Hamilton 2012; Gillon and others 2016; Rissman and Carpenter 2015). In this study, we used models to bypass this challenge. We quantified the influence of legacy P on water quality in isolation from confounding factors such as land use, land management, and climate. Specifically, we examined the sensitivity of water quality outcomes to legacy P within the YW using alternative scenarios of P storage in watershed soils and stream sediments. We asked, how much does legacy P influence lake water quality, and what would present water quality be like if past levels of stored P had been different? The study was conducted using a newly-developed numerical modeling framework to simulate the scenarios and generate outcomes of water quality.

2.2. Methods

2.2.1. Study Area

The YW is located mainly within Dane County, Wisconsin, with small portions in Rock and Columbia Counties. The 1344 km^2 watershed contains a chain of four eutrophic lakes that drain from north to south in the following order: Mendota, Monona, Waubesa, and Kegonsa. Roughly half the landscape in the YW is devoted to agriculture, with corn, soy, and dairy the principal products. The northern and southern parts of the watershed are predominantly agricultural, with dairy operations more common in the north and cash grain in the south. The Wisconsin state capital city of Madison (43°6'N, 89° 24'W) and surrounding urban area comprises roughly one quarter of the watershed and is centered on an isthmus between Lakes Mendota and Monona. The remaining quarter of the watershed is covered in natural vegetation, including forests, wetlands, and prairie.

2.2.2. Models

We developed a watershed modeling framework that can simulate an array of ecosystem services, including land-to-lake flows of water, sediment, and nutrients, and surface water quality (Fig. 1). The framework includes process-based representation of natural and managed ecosystems (Agro-IBIS), hydrologic routing of water, sediment, and nutrients through the surface hydrologic network (THMB), and prediction of Secchi disc depth (a measure of lake transparency), summertime lake total phosphorus (TP) concentration, and the probability of hypereutrophy in each lake (Yahara Water Quality Model).

The Agro-IBIS model is a spatially explicit land surface model that simulates the movement of water, energy, momentum, carbon, nitrogen, and now phosphorus, in both natural and managed ecosystems. The structure of Agro-IBIS has been described in detail (Foley and others 1996; Kucharik and others 2000; Kucharik 2003), and many components of the model have been validated across a range of ecosystems at various spatial and temporal scales (El Maayar and others 2001; Kucharik and Brye 2003; Kucharik and others 2006; Kucharik and Twine 2007; Soylu and others 2014; Zipper and others 2015). Recently, Agro-IBIS was integrated with the variably saturated soil water flow model HYDRUS-1D to enable simulation of groundwater-vegetation interactions (Soylu and others 2014). For this study several additional updates have been made to Agro-IBIS. First, biogeochemical cycling of P and loss of P to runoff were added to Agro-IBIS to enable simulation of P dynamics, including interactions between surface water quality, climate, and land management. The new terrestrial P module is largely based on SurPhos, a state-of-the-art dissolved P loss model for agricultural systems receiving manure (Vadas and others 2004, 2005, 2007, Vadas and White 2010). The new P module features P application, transformation, and loss of dissolved P to runoff; in-soil cycling of organic and inorganic forms of P; and loss of particulate-bound P with erosion (Fig. 2). Land cover categories have also been expanded in Agro-IBIS to include four classes of urban areas (high-, medium-, and low-intensity as well as open space) along with alfalfa, pasture, and wetlands. Finally, changes were made within the soil physics routines to improve representation of surface hydrology and improve model stability. Details for all changes made to Agro-IBIS are provided in Appendix A.

Agro-IBIS was linked to the Terrestrial Hydrology Model with Biogeochemistry (THMB) to enable simulation of water, sediment, and P transport at the watershed scale,

including delivery of both dissolved and particulate P loads to the YW lakes and the Rock River. The THMB model, formerly called HYDRA, is a physically-based hydrologic routing model. By linking the topographic data of a prescribed stream network and river morphological characteristics to a set of linear reservoir functions, the model simulates the temporal variability of water flow and storage in hydrologic systems (Coe 1998, 2000; Coe and others 2008). Donner and others (2002) added a nitrogen transport module to the THMB modeling framework. In this study, transport and cycling of sediment and P have been added to THMB. Using spatially explicit outputs of surface runoff, subsurface drainage, sediment yield and P yield generated for each grid cell by Agro-IBIS, THMB transports water, sediment, and P across the 2-D landscape. THMB calculates direct drainage and total loading of P and sediment to the lakes, as well as the movement of these quantities through the lake chain (Fig. 1). Further details for the new THMB capabilities are provided in Appendix B.

The Yahara Water Quality (WQ) Model predicts summer water quality variables in the four mainstem Yahara lakes given annual direct drainage loads to each lake from THMB. The model computes a mass balance for each lake given empirical relationships (Carpenter and Lathrop 2014). Total annual loads to Lake Mendota are calculated using direct drainage loads only since there are no P inputs from other lakes. For the downstream lakes, total annual loads are determined using direct drainage loads from THMB as well as estimates of P export from upstream lakes. Summer water quality is computed from terms of the P mass balance using regressions (Lathrop and Carpenter 2013; Carpenter and Lathrop 2014). Further details of the Yahara WQ Model are presented in Appendix C.

For water quality purposes, summer is defined as 30 June to 7 September, which is reliably a period of summer lake stratification. Water quality variables are TP (total P

concentration in surface water), DRP (dissolved reactive phosphorus concentration in surface water) chlorophyll (chlorophyll a concentration in surface water), and Secchi disc transparency. Lake trophic state and water quality are multivariate phenomena, and multiple indicators are commonly used to obtain different perspectives (Carlson 1977). We defined eutrophic (vs. mesotrophic conditions) by TP > 0.024 mg L^{-1} or Secchi transparency < 2 m (Carlson 1977). We define hypereutrophy by the presence of DRP at concentrations greater than 0.005 mg L^{-1}, indicating that phytoplankton are not removing all detectable DRP from the water. Empirical relationships in the Yahara WQ Model were computed using 33 full years of data for Lakes Mendota and Monona (1975-2008) and 28 full years of data for Lakes Waubesa and Kegonsa (1980-2008). Datasets and regression models are described by Lathrop and Carpenter (2013) and Carpenter and Lathrop (2014). Data are archived by the North Temperate Lakes Long-Term Ecological Research program (http://lter.limnology.wisc.edu).

2.2.3. Experimental Design

Prior to all simulations, a spin-up of C and N pools in Agro-IBIS was first performed. The dependency of P cycling on water, C, and N cycles necessitated the need for a spin-up since this is how C and N pools are brought to equilibrium within soils and vegetation (Kucharik and others 2000). Because the model was initialized without any C and N in the soil pools, the amount of time to bring the pools to equilibrium was roughly 200 years, making the spin-up period span 1786-1985. Calibration and validation results for Agro-IBIS, THMB, and the Yahara WQ Model are presented in Appendix D. Results of parameter sensitivity analyses conducted for Agro-IBIS and THMB are presented in Appendix J.

To investigate the effect of legacy P on surface water quality, we ran five simulations over the recent historical period (1986-2013) in which the initial amount of P stored in soils and river channels was varied (Fig. 3). The simulations were named according to their ranking in initial P content: low (LO), medium-low (ML), medium (ME), medium-high (MH), and high (HI). Using output from the Yahara WQ Model, we analyzed three water quality indicators for each simulation including summer lake TP concentration, Secchi depth, and the probability of hypereutrophy. Model inputs of climate, soil, groundwater, LULC, and nutrients are described in Appendices E-I.

The simulations were designed to span a range of plausible values for P storage in surface soils and streams based on available data and past studies. The third simulation, ME, was determined through the calibration process and was used to represent the best estimates of historical soil P, P yield, and P loading to the lakes. The LO and ML simulations were initialized with less soil and channel P storage than ME, and likewise the MH and HI simulations were initialized with more soil and channel P storage. All Agro-IBIS simulations were preceded by a 25 y soil P spin-up (1961-1985) to allow soil P pools in croplands to build up gradually to five contrasting levels in 1986. The levels were chosen to span a plausible range of soil test P observed in surface soil layers in croplands (Sharpley and others 1994; Andraski and Bundy 2003; Smith and Warnemuende-Pappas 2015). These values were approximately 20, 95, 175, 255, and 350 ppm for LO, ML, ME, MH, and HI, respectively. An upper limit of 350 ppm was chosen because a simulated equilibrium was observed near this value for silt loam soil within the watershed. The spin-ups were conducted by first initializing the labile P pools to 6, 20, 30, 37.5, and 50 ppm, respectively, for the five simulations. Assuming the labile pool is equal to half of Bray-1 soil test P (Vadas and White 2010), these values would correspond to 12, 40, 60, 75, and

100 ppm soil test P, respectively, in the surface layer. Manure was applied annually throughout each spin-up at rates of 15, 35, 50, 62, and 80 kg ha^{-1}, respectively, for the five simulations, in order to build up soil P to the desired levels. The soil P spin-up was conducted only in cropland grid cells, where corn, soy, wheat, or alfalfa was grown. This was done because croplands represent the primary source for non-point P pollution in the YW, and also because P interventions are mostly targeted at croplands, where soil P dynamics are heavily influenced by management. Labile P in non-crop grid cells was initialized at 15 ppm (30 ppm soil test P), which is consistent with local observations of soil test P in lawns and remnant prairies (Bennett and others 2004).

Channel P storage, or P stored in stream sediments, is an important contributor to in-stream P concentration (Walling and others 2008) and so channel P storage in THMB was also varied across the five simulations to represent a range of conditions consistent with low storage (0.2 kg ha^{-1}, LO), medium-low storage (2.0 kg ha^{-1}, ML), calibrated storage (21 kg ha^{-1}, ME), medium-high storage (103 kg ha^{-1}, MH) and high storage (207 kg ha^{-1}, HI). Levels of in-stream P storage vary significantly among watersheds and actual in-stream P storage data are highly limited. Owens and Walling (2002) reported in-stream total P storage at different locations of a rural watershed in the UK, which had similar sediment P concentration levels as the YW (Water Resources Management Practicum, 2015). Owens and Walling (2002) presented a range of in-stream P storage levels from 0.4 kg ha^{-1} to 63.2 kg ha^{-1}, which was consistent with the range of storage levels used in our experiment. The higher cases of MH and HI allowed us to also explore P load variability under more extreme conditions of legacy P buildup. Lake P storage was not adjusted in the experiment. Land cover and meteorological input data as well as manure and fertilizer P applications were the same for all five simulations, identical to those used in the

calibrated historical run (ME). This was done in order to isolate the effect of legacy P storage on lake loads and water quality indicators.

2.2.4. Analysis

To analyze P transport through the watershed and assess water quality outcomes for each simulation, the following quantities were analyzed: average annual P yield calculated by Agro-IBIS; total annual direct drainage load to each lake calculated by THMB; and summertime lake TP, Secchi depth, and probability of hypereutrophy in each lake calculated by the Yahara WQ Model. Phosphorus yields represented the average amount of P coming off of land-based grid-cells within the total drainage area for each lake (Appendix Fig. B1). These were calculated by summing both dissolved and particulate forms of P in runoff. Loads calculated in THMB represented the annual amount of total P reaching each lake via direct drainage (i.e. did not include riverine contributions from upstream lakes since those are determined by the Yahara WQ Model). The loads calculated by THMB were used as inputs to the Yahara WQ Model where calculations of summer TP concentration, Secchi depth, and the probability of hypereutrophy were made for each lake. All annual variables were calculated for the water year, defined as November 1st - October 31st. For each variable analyzed, we performed a randomized block design ANOVA with year as the random effect and simulation as the fixed effect to test the null hypothesis that the means of each P storage treatment in the five simulations were the same. Standard deviations in error were also calculated according to the randomized block ANOVA and were reported in the results along with means and used as error bars in figures. All analyses were performed using the MATLAB software package and statistics toolbox (The MathWorks, Inc. 2015).

2.3. Results

Simulated surface soil P concentration varied among the five experimental runs (Fig. 4) as well as among lake drainage areas (Appendix Fig. B1). The variations across drainage areas could be explained by differences in the spatial distribution of land cover types and the corresponding applications of fertilizer and manure. Soil P was greatest in the Lake Mendota drainage area in simulations ML through HI because of the prevalence of dairy operations and high rates of manure application north of Lake Mendota. Average surface soil P in croplands changed over time for each simulation as well (Fig. 5) due to climatic variability as well as equilibration processes within the soil system. For example, in the case of LO and ML, soil P increased between 1986 and 2013 as applications of fertilizer and manure accumulated in the surface layer. In contrast, for the high starting concentrations in MH and HI, surface soil P decreased through time as rainfall leaching and physical mixing caused excess P to transport to lower soil layers.

In general, the magnitude of P fluxes within the watershed increased and water quality degraded with each simulation as P storage increased. This was the case for mean P yield which increased across the simulations (Fig. 6). Mean yields were statistically different according to ANOVA ($p < 0.01$), and ranged from 0.70-1.43 $\pm$ 0.12 (s.d.) kg ha^{-1} for LO and HI, respectively, in the Lake Mendota drainage area. Similarly, mean P yield spanned 0.67-1.39 $\pm$ 0.12 kg ha^{-1} for Lake Monona, 0.64-1.36 $\pm$ 0.12 kg ha^{-1} for Lake Waubesa, and 0.61-1.33 $\pm$ 0.12 kg ha^{-1} for Lake Kegonsa.

Direct drainage loads into the four lakes also increased with P storage (Fig. 7A). For Lake Mendota, mean loads for the five simulations were statistically different ($p < 0.05$) and

ranged from 18,662-32,500 ± 4,276 kg y^{-1}. Mean loads to the other three lakes were also statistically different across the simulations ($p < 0.01$) and ranged from 1,389-4,016 ± 953 kg y^{-1} in Lake Monona, 773-2,382 ± 495 kg y^{-1} in Lake Waubesa, and 2,350-5,783 ± 1,133 kg y^{-1} in Lake Kegonsa. Lake Mendota received the highest direct drainage loads of any lakes followed by Kegonsa, Monona, and Waubesa in descending order. Direct drainage loads to Lake Mendota were more sensitive to P storage than the other lakes.

Summer TP concentration in each of the lakes was positively related to P storage (Fig. 7B). Mean concentrations of TP were statistically different ($p < 0.05$ for Lake Mendota; $p < 0.01$ for Lakes Monona, Waubesa, and Kegonsa) and ranged from 0.026-0.039 ± 0.006 mg L^{-1} in Lake Mendota, 0.028-0.033 ± 0.002 mg L^{-1} in Lake Monona, 0.047-0.059 ± 0.004 mg L^{-1} in Lake Waubesa, and 0.056-0.074 ± 0.008 mg L^{-1} in Lake Kegonsa. Among all four lakes, Kegonsa experienced the highest TP concentrations followed by Waubesa, Mendota, and Monona in descending order.

Consistent with the patterns observed in load and TP, Secchi depth decreased with increasing soil P (Fig. 7C). Secchi depths were significantly different ($p < 0.05$ for Lake Mendota; $p < 0.01$ for Lakes Monona, Waubesa, and Kegonsa) and ranged from 2.05-2.26 ± 0.054 m in Lake Mendota, 1.75-1.81 ± 0.16 m in Lake Monona, 1.05-1.05 ± 0.002 m in Lake Waubesa, and 0.96-1.02 ± 0.009 m in Lake Kegonsa. Lake Mendota had the highest observed Secchi depths, with all five simulations exceeding the 2 m index for mesotrophy (Carlson 1977). None of the other lakes reached a 2 m depth in any simulation. The probability of hypereutrophy increased across the simulations (Fig. 7D). Probabilities were significantly different ($p < 0.05$ for Lake Mendota; $p < 0.01$ for Lakes Monona, Waubesa, and Kegonsa) and ranged from 0.04-0.15 ± 0.074 in Lake Mendota, 0.04-0.06 ± 0.013 in Lake Monona, 0.18-0.32 ± 0.050 in Lake

Waubesa, and 0.28-0.46 ± 0.052 in Lake Kegonsa. Lake Kegonsa had the highest probability of hypereutrophy followed by Lakes Waubesa, Mendota, and Monona, in descending order. The lowest probability of hypereutrophy for any lake was 0.008, observed for Lake Mendota in LO.

We examined the percentage change in P yield, direct drainage load, summer TP, Secchi depth, and probability of hypereutrophy compared to the ME case, which best captured the recent historical record of P loads (Fig. 8). In general, the deviations from ME were different among the five variables and the four lakes. For Lake Mendota, the deviations in P yield matched the relative spacing among soil P trajectories (Fig. 5), with a greater absolute deviation occurring in LO (-40%) than in HI (+22%). The relative deviations in P load and summer TP were similar to the relative deviations in P yield although they spanned a smaller range. Secchi depth was the least responsive of the five variables, spanning the greatest range in Lake Mendota (-7% to +3%) and the smallest range in Lake Waubesa (-0.6% to +0.2%). The probability of hypereutrophy varied the most among all five variables, spanning the greatest range for Lake Mendota (-67% to +34%), and the smallest range in Lake Kegonsa (-33% to +11%). For the downstream three lakes, the deviations in P load for LO and ML were greater than the deviations in P yield. The responses of summer TP, Secchi depth, and the probability of hypereutrophy were muted, however, compared with the responses of these indicators in Lake Mendota. This suggests that load reductions in the lower lakes do not result in the same degree of water quality improvement that would be expected in Lake Mendota.

We also analyzed the response of summer TP concentration to the frequency of extreme precipitation events for each simulation (Fig. 9). Linear regression analysis of log(TP) versus the number of weeks per year with rainfall exceeding 100 mm (r100) was performed for each simulation and lake combination. The analysis indicated that the relationship between r100 and

log(TP) was significant in all cases ($p < 0.01$), with R^2 values for all regressions near 0.66. After verifying that plots of the residuals lacked discernible pattern, we used ANCOVA to confirm that the slopes of the regression lines were significantly different ($p < 0.01$) among the simulations for each lake. In general the slopes increased across the simulations, indicating a higher sensitivity of TP to extreme precipitation events as P storage increased. This suggests that lake water quality may be more vulnerable to extreme precipitation when P storage is high, and conversely less vulnerable when P storage is low.

2.4. Discussion

We tested the effect that different levels of legacy P at the onset of a 27-year simulation period would have on the magnitude and trajectory of P loading to the Yahara lakes and a suite of water quality variables. Our results showed that legacy P was an important determinant of future outcomes; the amount of P stored in soils and channel beds had a significant impact on lake loads and water quality. For Lake Mendota, the drop in initial P storage between ME and LO resulted in a 35% reduction in direct drainage P loads (Fig. 8). This reduction led to a 25% decrease in mean annual summer TP concentration in that lake, and increased the frequency of reaching mesotrophic status in summer (TP < 0.024 mg L^{-1}) by 25% (12 to 15 of 27 years). The water quality response to the LO legacy P reduction was muted in the lower lakes. Direct drainage load reductions of 60, 60 and 52% in Monona, Waubesa and Kegonsa, respectively, did not cause TP concentrations to reach the mesotrophic threshold in any year despite exceeding the current management target for the Yahara lakes of a 50% reduction in lake loading (Lathrop and others 1998). Our results suggest that lower direct drainage loads than those simulated in LO would be needed to achieve the target TP concentration in all four lakes. It should also be noted

that total loads to the lower lakes are generally dominated by riverine loads from upstream lakes (Carpenter and Lathrop 2014). Reductions in direct drainage loads achieved through changes in land use therefore represent only a partial reduction in total loads to the lower lakes. It may be necessary to also remove P stored in upstream lake sediments in order to decrease total loads to the downstream lakes sufficiently to alter their trophic status. The effect of lake sediment P storage was not examined in this analysis since it is implicitly handled by the Yahara WQ Model.

By using the same historical nutrient applications for all simulations, we isolated the effect of antecedent legacy P storage on water quality. This approach also showed the effect of historical nutrient applications on mean soil P concentration in the watershed. Specifically, the steady accumulation of soil P in LO over the 1986-2013 period indicated a build-up of P in the surface soil layer occurring as a direct result from historical nutrient applications and land management (Fig. 5). Given that the starting soil P concentration for LO in 1986 corresponded roughly with the recommended level for crops (20 ppm), the buildup between 1986 and 2013 demonstrated that historical nutrient applications exceeded crop needs and were great enough to increase soil P above agronomic recommendations within a 27-year period. Additionally, ME represented the best estimate of soil P, P yield, and lake loads over the historical period and indicated that surface soil P was approximately in equilibrium at 176 ppm (Fig. 5), and the second soil layer (2.5 – 15 cm) was roughly in equilibrium around 73 ppm (Appendix D). This implied that historical application rates were high enough to sustain a soil P concentration roughly nine times the recommended level for crops in the surface layer, and roughly four times the recommended level in the second soil layer. Overall, these findings suggest that there is

currently a substantial excess of P stored in watershed soils that is likely to influence outcomes of P cycling and lake water quality well into the future.

Among the four lakes, Mendota had the greatest absolute reduction in direct drainage P loading when comparing ME and LO (Fig. 7A). This is likely due to its large direct drainage area and close proximity to dairy operations in the northern YW. Additionally, all loads entering Mendota come from direct drainage sources, whereas the majority of total loads to the lower lakes come from the upstream lakes. As a result, water quality is generally worse in the lower lakes and less variable year-to-year. Water quality in the Yahara lakes is strongly influenced by the presence of *Daphnia pulicaria*, an effective grazer of algae (Lathrop and Carpenter 2013). Among zooplankton species, *Daphnia pulicaria* is exceptionally effective in reducing lake TP concentrations because of its rapid grazing rate, broad diet, high metabolic demand for P, and sedimentation of P in its feces (Carpenter and Kitchell 1993). Our results were calculated assuming the presence of *Daphnia*. Without the presence of *Daphnia*, summer TP concentrations in Lake Mendota would be approximately 30% higher in Fig. 7b for each simulation (results not shown). Populations of *Daphnia* were relatively steady over the time period examined in this study, however the recent invasion of spiny water flea (*Bythotrephes longimanus*) within the lakes has resulted in a collapse of *Daphnia* populations, making the future presence of *Daphnia* in the lakes uncertain (Walsh and others 2016). Given its important role, managing the Yahara lakes to maintain *Daphnia* populations should be a priority.

Prior studies have shown that heavy precipitation events can mobilize large amounts of soil P in runoff (Kleinman and others 2006; Shigaki and others 2007). Carpenter and others (2014) showed that heavy runoff events are responsible for the majority of annual total P loading to Lake Mendota. They found that on average 29 precipitation days per year (roughly one quarter

of all days with measurable precipitation) accounted for 74% of annual loading to the lake over the 1976-2008 period. Our findings were consistent with these observations, however we also observed that the response of summer TP to high precipitation (and thus high load) events was greater at higher levels of P storage (Fig. 9). This effect is due to greater amounts of stored P being mobilized during heavy rain events. Supporting analysis (results not shown) did not find this effect to be more or less pronounced for dissolved or sediment forms of P yield, or for different sources of P yield, such as manure, fertilizer, or soil. The implication of this finding is that water quality may be more vulnerable to heavy rainfall events at higher levels of P storage and conversely less vulnerable at lower levels. Given anticipated increases in annual and extreme precipitation with climate change (WICCI 2011, Trenberth 2011), this provides a two-fold incentive for reducing P reserves in the watershed. Not only will the consequences of legacy P be exacerbated if accumulation continues, but if drawdown occurs the benefits may include additional protection from extreme precipitation.

Gillon and others (2016) suggested that the water quality benefits of BMPs within the YW have likely been stymied by changes in precipitation, land use, and the problem of legacy P. Adoption of P control strategies within the YW has been fueled by collaborations among researchers, NGOs, stakeholder groups, resource managers, and local government. These include a variety of policies, regulations, and incentives geared toward the use of BMPs. Formalized nutrient management plans, conservation tillage, vegetative buffers, and manure digesters represent some of the BMPs that have been emphasized and implemented in the YW. However, adoption of these measures has not always been effectively targeted to areas of the landscape where they are most needed (Wardropper and others 2015). Additionally, many BMPs that target soil erosion and loss of particulate P will not fully address the problem since loads to Lake

Mendota are roughly equally comprised of dissolved and particulate forms. Dissolved P loading from manured plots, particularly in late-winter and early spring, represents a significant source of bioavailable P to surface waters. No-till conservation practices, which have been widely adopted in the YW, are effective at preserving topsoil and reducing sediment P loads but also have the unintended consequence of increasing DRP loads, particularly in systems receiving broadcast manure applications (Bundy and others 2001; Kleinman and others 2008). Conservation practices aimed at preventing P runoff can potentially worsen legacy P buildup, creating a "chemical time-bomb" that remains vulnerable to runoff and other transport mechanisms (Stigliani and others 1991). Conversion to perennial systems, including grasslands and forests, may not necessarily reduce loads since legacy P effects have been observed to be persistent in such systems even in the absence of high-intensity rainfall (Scott and others 2001; Bilotta and others 2007; Horrocks and others 2014).

Research elsewhere has shown that while BMPs can effectively stem P runoff at the field-scale, there has been disappointingly little improvement in downstream water quality and ecology (Jarvie and others 2013a, 2013b). Phosphorus losses from soil and fluvial sediment stores still occur in places of high P buildup and can effectively mask the water quality benefits of conservation efforts (Sharpley and others 2013; Sharpley 2016). By quantifying the direct effect that legacy P exerted on water quality in the Yahara, our study supports the general hypothesis that BMPs and conservation efforts are counteracted by the slow release of legacy P from watersheds and water bodies (Jarvie and others 2013). Our results may also provide some support to observations of the legacy effects of other non-point source contaminants, such as nitrogen, sediment, and chloride (Bester and others 2006; Meals and others 2010; James 2013; Van Meter and others 2016).

With its strong agricultural history and growing population, the YW is an exemplar of many of the world's watersheds that support intensive agriculture. As such, the results of our study have important implications for other regions where there is intensive agricultural production and high P abundance, such as China (Chen and others 2012; Dai and others 2011), U.K. and other European countries (Withers and others 2001; Azevedo and others 2015), and elsewhere in the U.S. (Ardón and others 2010; Dubrovsky and others 2010). P yield, loading, and water quality will vary across landscapes according to differences in land use, management, topography, soil type, climate, etc. However, the basic principles governing the supply and transport of P apply to all regions. Using physically-based models, we demonstrated a causal connection between legacy P and surface water quality. The strength of this connection may differ across regions, but its important role within the Yahara implies a similarly important role in other landscapes having a surplus of P and poor freshwater quality.

Our findings suggest that drawing down soil P reserves, for example through recovery and recycling of P in soils, channels, and manures, would help buffer against the unintended effects of management as well as climate change. For example, numerous conservation efforts were enacted over the past decade in an attempt to decrease P loading to Lake Erie. Yet, the actual changes in management and climate over this time period likely hindered those efforts. The adoption of conservation tillage and the subsequent need for broadcast P applications increased P accumulation in surface soils, thereby increasing P available to runoff (Sharpley 2016). Additionally, widespread adoption of tile drainage created an easy path for nutrient rich subsurface drainage to enter waterways and ultimately Lake Erie. On top of that, extreme precipitation events increased over the time period, and according to our results may have exacerbated the effect of increasing surface soil P on P losses. An alternative strategy for

management that focuses on drawing down the P surplus, and not just preventing its loss in runoff, would address a root cause of the eutrophication problem. Removing P from the system precludes much of the uncertainty in management and climate, and how those factors may affect P cycling and transport. Our results suggest that successful removal of P from saturated watersheds promises improvements in water quality, as well as protective benefits.

The Yahara is a well-studied watershed having long-term observations and many prior studies from which to draw. Despite the wealth of data available, challenges still exist in performing simulations at the watershed scale. One challenge is insufficient data to initialize the model and represent the diversity of management practices that affect hydrology and nutrient cycling. For example, soil P may vary widely within watersheds (Wang and others 2009) and even within individual fields as a result of past management (Page and others 2005). In this study the starting soil P concentration in 1986 across the watershed at a 220 m spatial resolution was not known and could only be approximated for different land cover types based on prior independent observations (e.g. Bennett and others 2004). We relied on small-scale validation studies to verify the correct sensitivity of the model (i.e. ability to simulate long-term soil P dynamics over a range of soil types, rainfall regimes, and management conditions), and calibrated the starting concentration to be within a reasonable range while also optimizing performance of P loading in THMB. While simulated soil P and lake loads were plausible for each simulation, further research is needed to understand how spatial diversity in soil P is related to P loads at the watershed scale, and to what extent input data must be improved.

Another challenge stemmed from the fact that in this study, the Yahara WQ Model was used to predict water quality outcomes using P loads that sometimes extended above and below the range of observations used to build the model (1975-2008). Therein lies a drawback of

statistical models when used to predict outcomes under anomalous conditions, which stands in contrast to their relative ease of development and use. One motivation of process-based model design is to be robustly parameterized so as to respond appropriately to a wide range in inputs and therefore be well-suited to simulate novel conditions of change. With the exception of the Yahara WQ Model, our modeling framework employed a process-based approach. This approach has helped us advance our understanding of the processes that affect water quality, such as the mechanisms driving soil P imbalance at the field-scale, and rates at which accumulation and depletion may occur at the watershed scale. We demonstrated that soil P may change rapidly over several years, or remain unchanged for decades or longer. Our model is thus poised to investigate timescales of watershed accumulation and depletion and the resulting impacts to water quality, an important knowledge gap in P research (Jarvie and others 2013; Haygarth and others 2014).

Further studies using the modeling framework will focus on the effects of climate, land use, and land management to better understand the roles that these drivers play in P cycling and water quality, and the timescales over which they act. The framework also includes representation of other ecosystem services, including groundwater quality and recharge, flood protection, carbon storage, and food production. Representation of these will enable study of the interactions, trade-offs and synergies among multiple ecosystem services. The framework thus represents a comprehensive watershed modeling tool that can simulate the dynamic effects of major change drivers on water quality and other services key for human well-being.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. DEB-1038759.

References

Andraski TW, Bundy LG. 2003. Relationships between phosphorus levels in soil and in runoff from corn production systems. J Environ Qual 32:310–6.

Ardón M, Morse JL, Doyle MW, Bernhardt ES. 2010. The Water Quality Consequences of Restoring Wetland Hydrology to a Large Agricultural Watershed in the Southeastern Coastal Plain. Ecosystems 13:1060–78.

Azevedo LB, van Zelm R, Leuven RSEW, Hendriks AJ, Huijbregts MAJ. 2015. Combined ecological risks of nitrogen and phosphorus in European freshwaters. Environ Pollut 200:85–92.

Bennett EM, Carpenter SR, Caraco NF. 2001. Human impact on erodable phosphorus and eutrophication: A global perspective. Bioscience 51:227–34.

Bennett EM, Carpenter SR, Clayton MK. 2004. Soil phosphorus variability: scale-dependence in an urbanizing agricultural landscape. Landsc Ecol 20:389–400.

Bennett EM, Reed-Andersen T, Houser JN, Gabriel JR, Carpenter SR. 1999. A phosphorus budget for the Lake Mendota Watershed. Ecosystems 2:69–75.

Bennett EM, Schipanski ME. 2013. The Phosphorus Cycle. In: Fundamentals of Ecosystem Science. Elsevier. pp 159–78.

Bilotta GS, Brazier RE, Haygarth PM. 2007. Processes affecting transfer of sediment and colloids, with associated phosphorus, from intensively farmed grasslands: erosion. Hydrol Process 21:135–9.

Bundy LG, Andraski TW, Powell JM. 2001. Management practice effects on phosphorus losses in runoff in corn production systems. J Environ Qual 30:1822–8.

Carlson RE. 1977. A trophic state index for lakes. Limnol Oceanogr 22:361–9.

Carpenter SR, Booth EG, Kucharik CJ, Lathrop RC. 2014. Extreme daily loads: role in annual phosphorus input to a north temperate lake. Aquat Sci 77:71–9.

Carpenter SR, Kitchell JF, editors. 1993. The Trophic Cascade in Lakes. Cambridge University Press, Cambridge, England. 385 pp.

Carpenter SR, Lathrop RC. 2014. Phosphorus loading, transport and concentrations in a lake chain: a probabilistic model to compare management options. Aquat Sci 76:145–54.

Carpenter SR, Lathrop RC, Nowak P, Bennett EM, Reed T, Soranno PA. 2006. The ongoing experiment: restoration of Lake Mendota and its watershed. Long-term dynamics of lakes in the landscape Oxford Univ Press, London, UK:236–56.

Chen Y, Liu R, Sun C, Zhang P, Feng C, Shen Z. 2012. Spatial and temporal variations in nitrogen and phosphorous nutrients in the Yangtze River Estuary. Mar Pollut Bull 64:2083–9.

Childers DL, Corman J, Edwards M, Elser JJ. 2011. Sustainability Challenges of Phosphorus and Food: Solutions from Closing the Human Phosphorus Cycle. Bioscience 61:117–24.

Coe MT. 1998. A linked global model of terrestrial hydrologic processes: Simulation of modern rivers, lakes, and wetlands. J Geophys Res 103:8885–99.

Coe MT. 2000. Modeling Terrestrial Hydrological Systems at the Continental Scale: Testing the Accuracy of an Atmospheric GCM. J Clim 13:686–704.

Coe MT, Costa MH, Howard EA. 2008. Simulating the surface waters of the Amazon River basin: impacts of new river geomorphic and flow parameterizations. Hydrol Process 22:2542–53.

Dai Z, Du J, Zhang X, Su N, Li J. 2010. Variation of Riverine Material Loads and Environmental Consequences on the Changjiang (Yangtze) Estuary in Recent Decades (1955- 2008)†. Environ Sci Technol 45:223–7.

Donner SD, Coe MT, Lenters JD, Twine TE, Foley JA. 2002. Modeling the impact of hydrological changes on nitrate transport in the Mississippi River Basin from 1955 to 1994. Global Biogeochem Cycles 16:1–19.

Dubrovsky NM, Burow KR, Clark GM, Gronberg JM, Hamilton PA, Hitt KJ, Mueller DK, Munn MD, Nolan BT, Puckett LJ, Others. 2010. The quality of our Nation's waters-Nutrients in the Nation's streams and groundwater, 1992--2004. US Geological Survey Circular 1350. http://pubs.usgs.gov/circ/1350/

Edwards AC, Withers PJA. 1998. Soil phosphorus management and water quality: a UK perspective. Soil Use Manage 14:124–30.

El Maayar M, Price DT, Delire C, Foley JA, Black TA, Bessemoulin P. 2001. Validation of the Integrated Biosphere Simulator over Canadian deciduous and coniferous boreal forest stands. J Geophys Res 106:14339.

Elser J, Bennett E. 2011. Phosphorus cycle: A broken biogeochemical cycle. Nature 478:29–31.

Foley JA, Prentice IC, Ramankutty N, Levis S, Pollard D, Sitch S, Haxeltine A. 1996. An integrated biosphere model of land surface processes, terrestrial carbon balance, and vegetation dynamics. Global Biogeochem Cycles 10:603–28.

Genskow K, Rumery Betz C. 2012. Farm Practices in the Lake Mendota Watershed: A Comparative Analysis of 1996 and 2011.

Gillon S, Booth EG, Rissman AR. 2016. Shifting drivers and static baselines in environmental governance: challenges for improving and proving water quality outcomes. Regional Environ Change 16:759–75.

Hamilton SK. 2012. Biogeochemical time lags may delay responses of streams to ecological restoration. Freshw Biol 57:43–57.

Haygarth PM, Jarvie HP, Powers SM, Sharpley AN, Elser JJ, Shen J, Peterson HM, Chan N-I, Howden NJK, Burt T, Worrall F, Zhang F, Liu X. 2014. Sustainable phosphorus management and the need for a long-term perspective: the legacy hypothesis. Environ Sci Technol 48:8417–9.

Horrocks CA, Dungait JAJ, Cardenas LM, Heal KV. 2014. Does extensification lead to enhanced provision of ecosystems services from soils in UK agriculture? Land Use Policy 38:123–8.

Jarvie HP, Sharpley AN, Spears B, Buda AR, May L, Kleinman PJA. 2013a. Water quality remediation faces unprecedented challenges from 'legacy phosphorus'. Environ Sci Technol 47:8997–8.

Jarvie HP, Sharpley AN, Withers PJA, Scott JT, Haggard BE, Neal C. 2013b. Phosphorus mitigation to control river eutrophication: Murky waters, inconvenient truths, and 'postnormal' science. J Environ Qual 42:295–304.

Kahiluoto H, Kuisma M, Ketoja E, Salo T, Heikkinen J. 2015. Phosphorus in manure and sewage sludge more recyclable than in soluble inorganic fertilizer. Environ Sci Technol 49:2115–22.

Kara EL, Heimerl C, Killpack T, Van de Bogert MC, Yoshida H, Carpenter SR. 2011. Assessing a decade of phosphorus management in the Lake Mendota, Wisconsin watershed and scenarios for enhanced phosphorus management. Aquat Sci 74:241–53.

Kleinman PJA, Sharpley AN, Saporito LS, Buda AR, Bryant RB. 2008. Application of manure to no-till soils: phosphorus losses by sub-surface and surface pathways. Nutr Cycling Agroecosyst 84:215–27.

Kleinman PJA, Srinivasan MS, Dell CJ, Schmidt JP, Sharpley AN, Bryant RB. 2006. Role of rainfall intensity and hydrology in nutrient transport via surface runoff. J Environ Qual 35:1248–59.

Kucharik CJ. 2003. Evaluation of a process-based agro-ecosystem model (Agro-IBIS) across the U.S. Corn Belt: Simulations of the interannual variability in maize yield. Earth Interact 7:1–33.

Kucharik CJ, Barford CC, Maayar ME, Wofsy SC, Monson RK, Baldocchi DD. 2006. A multiyear evaluation of a dynamic global vegetation model at three AmeriFlux forest sites: Vegetation structure, phenology, soil temperature, and CO2 and H2O vapor exchange. Ecol Modell 196:1–31.

Kucharik CJ, Brye KR. 2003. Integrated BIosphere Simulator (IBIS) yield and nitrate loss predictions for Wisconsin maize receiving varied amounts of nitrogen fertilizer. J Environ Qual 32:247–68.

Kucharik CJ, Foley JA, Delire C, Fisher VA, Coe MT, Lenters JD, Young-Molling C, Ramankutty N, Norman JM, Gower ST. 2000. Testing the performance of a dynamic global ecosystem model: Water balance, carbon balance, and vegetation structure. Global Biogeochem Cycles 14:795–825.

Kucharik CJ, Twine TE. 2007. Residue, respiration, and residuals: Evaluation of a dynamic agroecosystem model using eddy flux measurements and biometric data. Agric For Meteorol 146:134–58.

Lathrop RC. 2007. Perspectives on the eutrophication of the Yahara lakes. Lake Reserv Manag 23:345–65.

Lathrop RC, Carpenter SR. 2013. Water quality implications from three decades of phosphorus loads and trophic dynamics in the Yahara chain of lakes. Inland Waters 4:1–14.

Lathrop RC, Carpenter SR, Stow CA, Soranno PA, Panuska JC. 1998. Phosphorus loading reductions needed to control blue-green algal blooms in Lake Mendota. Can J Fish Aquat Sci 55:1169–78.

MA (Millenium Ecosystem Assessment). 2005. Ecosystems and human well-being: summary for decision makers. Washington D.C.: Island Press

McCollum RE. 1991. Buildup and decline in soil phosphorus: 30-year trends on a typical Umprabuult. Agron J.

Meals DW, Dressing SA, Davenport TE. 2010. Lag time in water quality response to best management practices: a review. J Environ Qual 39:85–96.

Metson GS, MacDonald GK, Haberman D, Nesme T, Bennett EM. 2015. Feeding the Corn Belt: Opportunities for phosphorus recycling in U.S. agriculture. Sci Total Environ. http://dx.doi.org/10.1016/j.scitotenv.2015.08.047

Nowak P, Shepard R, Madison F, Hatfield JL, Stewart BA, Others. 1998. Farmers and manure management: a critical analysis. In: Hatfield JL, Stewart BA, editors. Animal Waste Utilization: Effective use of manure as a soil resource. Lewis Publishers. pp 1–32.

Owens PN, Walling DE. 2002. The phosphorus content of fluvial sediment in rural and industrialized river basins. Water Res 36:685–701.

Page T, Haygarth PM, Beven KJ, Joynes A, Butler T, Keeler C, Freer J, Owens PN, Wood GA. 2005. Spatial variability of soil phosphorus in relation to the topographic index and critical source areas. J Environ Qual 34:2263–77.

Powers SM, Bruulsema TW, Burt TP, Chan NI, Elser JJ, Haygarth PM, Howden NJK, Jarvie HP, Lyu Y, Peterson HM, Sharpley AN, Shen J, Worrall F, Zhang F. 2016. Long-term accumulation and transport of anthropogenic phosphorus in three river basins. Nat Geosci.

Reed-Andersen T, Carpenter SR, Lathrop RC. 2000. Phosphorus flow in a watershed-lake ecosystem. Ecosystems 3:561–73.

Rissman AR, Carpenter SR. 2015. Progress on nonpoint pollution: Barriers & opportunities. Daedalus 144:35–47.

Rowe H, Withers PJA, Baas P, Chan NI, Doody D, Holiman J, Jacobs B, Li H, MacDonald GK, McDowell R, Sharpley AN, Shen J, Taheri W, Wallenstein M, Weintraub MN. 2015. Integrating legacy soil phosphorus into sustainable nutrient management strategies for future food, bioenergy and water security. Nutr Cycling Agroecosyst:1–20.

Sattari SZ, Bouwman AF, Giller KE, van Ittersum MK. 2012. Residual soil phosphorus as the missing piece in the global phosphorus crisis puzzle. Proc Natl Acad Sci U S A 109:6348–53.

Schulte RPO, Melland AR, Fenton O, Herlihy M, Richards K, Jordan P. 2010. Modelling soil phosphorus decline: Expectations of Water Framework Directive policies. Environ Sci Policy 13:472–84.

Scott CA, Walter MF, Nagle GN, Todd Walter M, Sierra NV, Brooks ES. 2001. Residual phosphorus in runoff from successional forest on abandoned agricultural land: 1. Biogeochemical and hydrological processes. Biogeochemistry 55:293–310.

Sharpley A. 2016. Managing agricultural phosphorus to minimize water quality impacts. Sci Agric 73:1–8.

Sharpley A, Jarvie HP, 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 42:1308–26.

Sharpley AN, Chapra SC, Wedepohl R, Sims JT, Daniel TC, Reddy KR. 1994. Managing agricultural phosphorus for protection of surface waters - issues and options. J Environ Qual 23:437–51.

Shigaki F, Sharpley A, Prochnow LI. 2007. Rainfall intensity and phosphorus source effects on phosphorus transport in surface runoff from soil trays. Sci Total Environ 373:334–43.

Smith DR, Warnemuende-Pappas EA. 2015. Vertical tillage impacts on water quality derived from rainfall simulations. Soil Tillage Res 153:155–60.

Smith VH, Joye SB, Howarth RW. 2006. Eutrophication of freshwater and marine ecosystems. Limnol Oceanogr 51:351–5.

Soranno PA, Hubler SL, Carpenter SR, Lathrop RC. 1996. Phosphorus loads to surface waters: a simple model to account for spatial pattern of land use. Ecol Appl 6:865–78.

Soylu ME, Kucharik CJ, Loheide, II, Steven P. 2014. Influence of groundwater on plant water use and productivity: Development of an integrated ecosystem – Variably saturated soil water flow model. Agric For Meteorol 189-190:198–210.

Staver KW, Brinsfield RB. 2001. Agriculture and water quality on the Maryland eastern shore: Where do we go from here? Bioscience 51:859–68.

Stigliani WM, Doelman P, Salomons W, Schulin R, Smidt GRB, der Zee SEATMV. 1991. Chemical time bombs: Predicting the unpredictable. Environment 33:4–30.

Stumborg BE, Baerenklau KA, Bishop RC. 2001. Nonpoint source pollution and present values: A contingent valuation study of Lake Mendota. Rev Agric Econ 23:120–32.

Sturgul SJ, Bundy LG. 2004. Understanding Soil Phosphorus. University of Wisconsin-Extension http://ipcm.wisc.edu/download/pubsNM/UnderstandingSoilP04.pdf

The MathWorks, Inc. 2015. MATLAB and Statistics Toolbox. Natick Massachusetts, USA: The MathWorks, Inc.

Trenberth KE. 2011. Changes in precipitation with climate change. Clim Res 47:123–38.

Vadas PA, Gburek WJ, Sharpley AN, Kleinman PJA, Moore PA Jr, Cabrera ML, Harmel RD. 2007. A model for phosphorus transformation and runoff loss for surface-applied manures. J Environ Qual 36:324–32.

Vadas PA, Haggard BE, Gburek WJ. 2005. Predicting dissolved phosphorus in runoff from manured field plots. J Environ Qual 34:1347–53.

Vadas PA, Kleinman PJA, Sharpley AN. 2004. A simple method to predict dissolved phosphorus in runoff from surface-applied manures. J Environ Qual 33:749–56.

Vadas PA, White MJ. 2010. Validating Soil Phosphorus Routines in the SWAT Model. Transactions of the ASABE 53:1469–76.

Walling DE, Collins AL, Stroud RW. 2008. Tracing suspended sediment and particulate phosphorus sources in catchments. J Hydrol 350:274–89.

Walsh JR, Carpenter SR, Vander Zanden MJ. 2016. Invasive species triggers a massive loss of ecosystem services through a trophic cascade. Proc Natl Acad Sci U S A. http://dx.doi.org/10.1073/pnas.1600366113

Wang Y, Zhang X, Huang C. 2009. Spatial variability of soil total nitrogen and soil total phosphorus under different land uses in a small watershed on the Loess Plateau, China. Geoderma 150:141–9.

Wardropper CB, Chang C, Rissman AR. 2015. Fragmented water quality governance: Constraints to spatial targeting for nutrient reduction in a Midwestern USA watershed. Landsc Urban Plan 137:64–75.

Water Resources Management Practicum. 2015. Assessment of transient sediment in the six-mile creek watershed. https://www.nelson.wisc.edu/docs/WRM2013_report.pdf

WICCI. 2011. Wisconsin's Changing Climate: Impacts and Adaptation. Wisconsin Initiative on Climate Change Impacts. Nelson Institute for Environmental Studies, University of Wisconsin-Madison and the Wisconsin Department of Natural Resources, Madison, Wisconsin

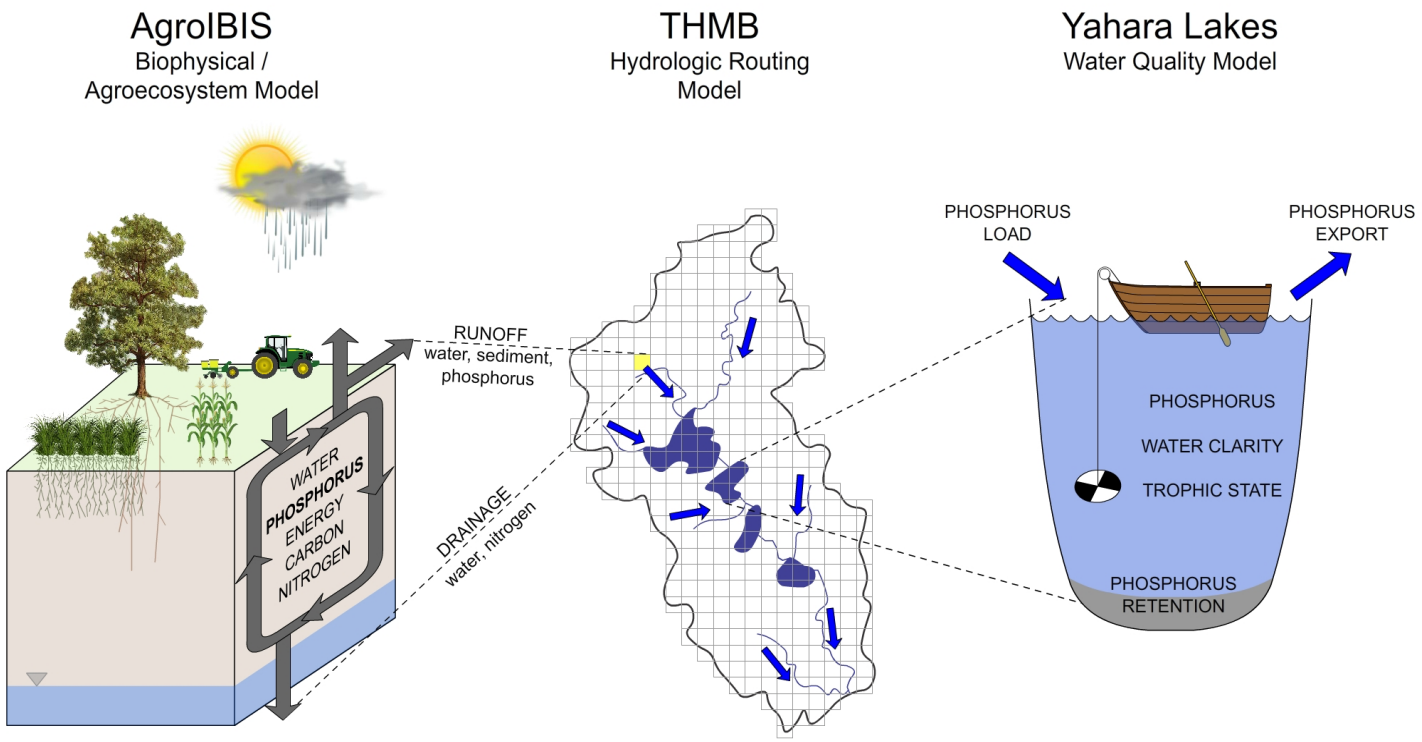
Wilfert P, Kumar PS, Korving L, Witkamp G-J, van Loosdrecht MCM. 2015. The relevance of phosphorus and iron chemistry to the recovery of phosphorus from wastewater: A review. Environ Sci Technol 49:9400–14.

Withers PJA, Edwards AC, Foy RH. 2001. Phosphorus cycling in UK agriculture and implications for phosphorus loss from soil. Soil Use Manage 17:139–49.

Wolfe ML, Ting KC, Scott N, Sharpley AN, Jones JW, Verma L. 2016. Engineering solutions for food-energy-water systems: it is more than engineering. Journal of Environmental Studies and Sciences. http://dx.doi.org/10.1007/s13412-016-0363-z. Last accessed 04/03/2016

Zipper SC, Soylu ME, Booth EG, Loheide, II, Steven P. 2015. Untangling the effects of shallow groundwater and soil texture as drivers of subfield-scale yield variability. Water Resour Res 51:6338–58.

Figure 1. Suite of three models used in this study.



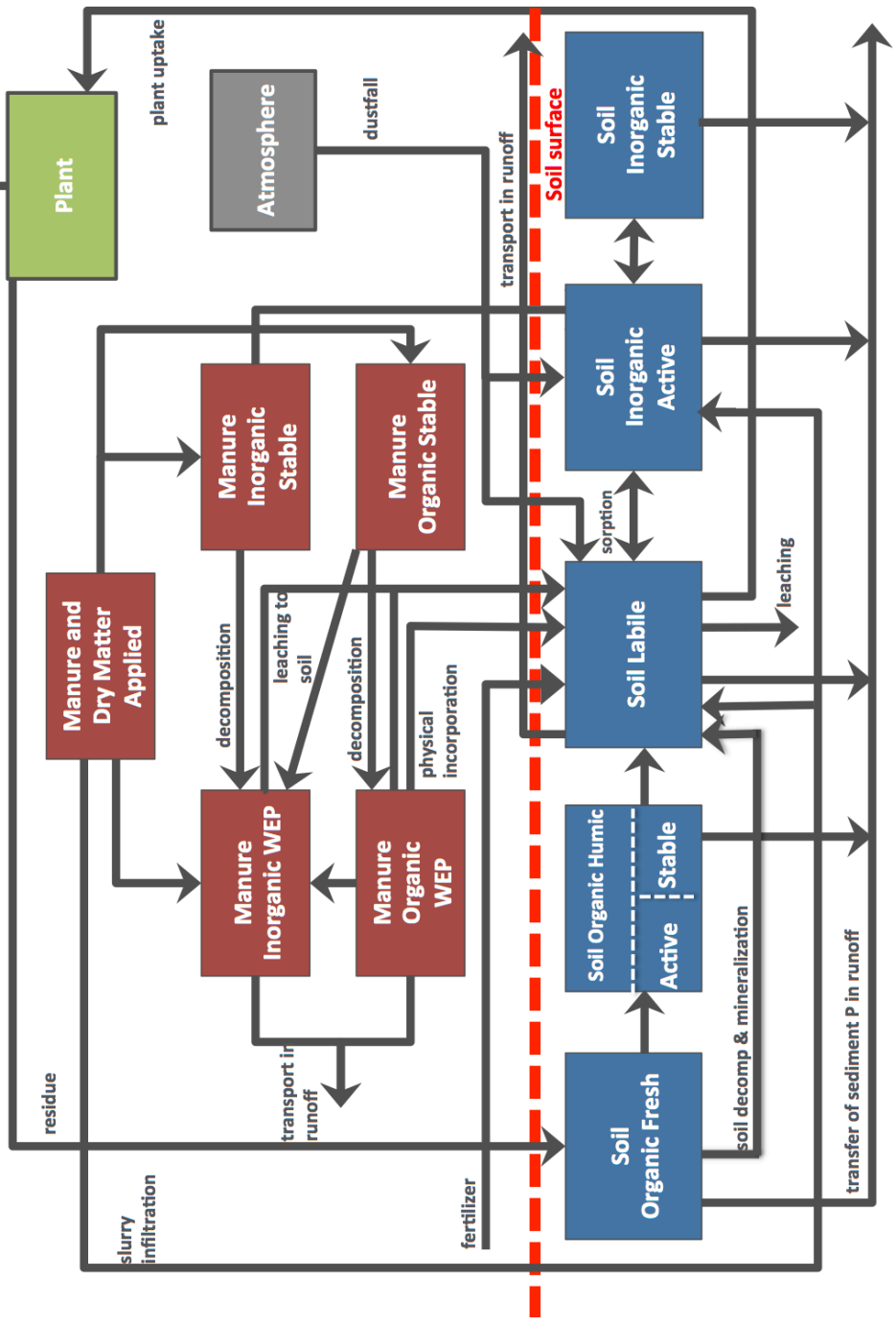


Figure 2. The terrestrial P module in Agro-IBIS, largely based on the SurPhos model of Vadas and others (2007).

Figure 3. Surface soil P concentration (ppm) and in-channel P storage (kg ha^{-1}) used to initialize Agro-IBIS and THMB in the five simulations. Simulation names correspond to their ranking of initial P: low (LO), medium-low (ML), medium (ME), medium-high (MH), and high (HI).

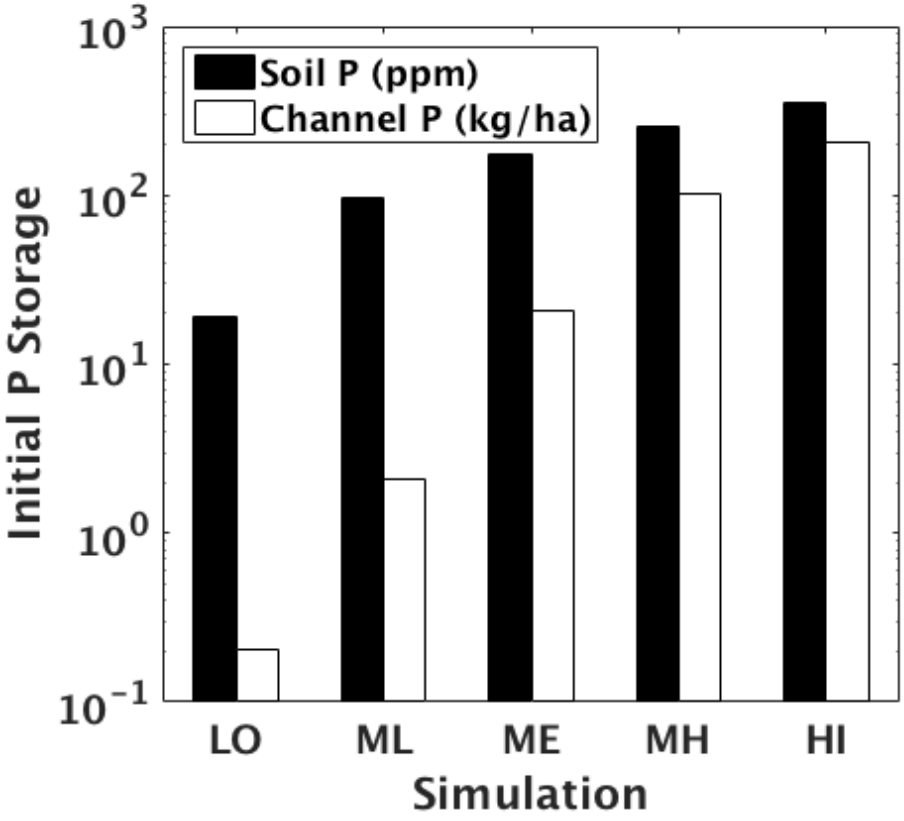


Figure 4. Surface soil P concentration in croplands (ppm) averaged over each lake's drainage area and the 1986-2013 time period.

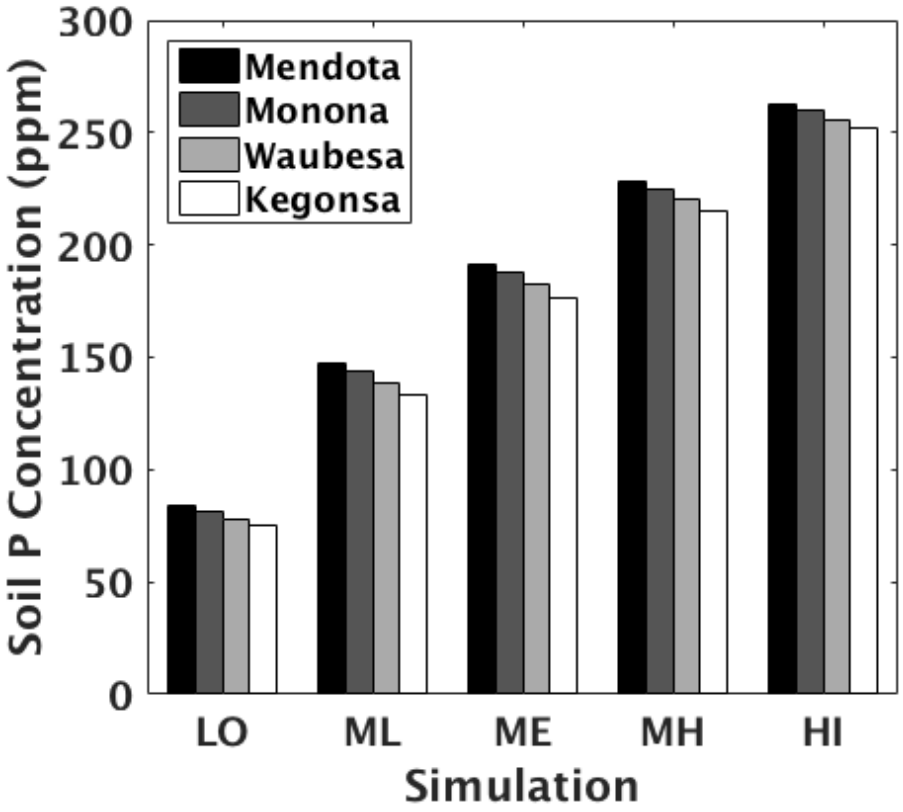


Figure 5. Watershed-average surface soil P concentration in croplands for both the spin-up period (1961-1985) and the simulation period (1986-2013).

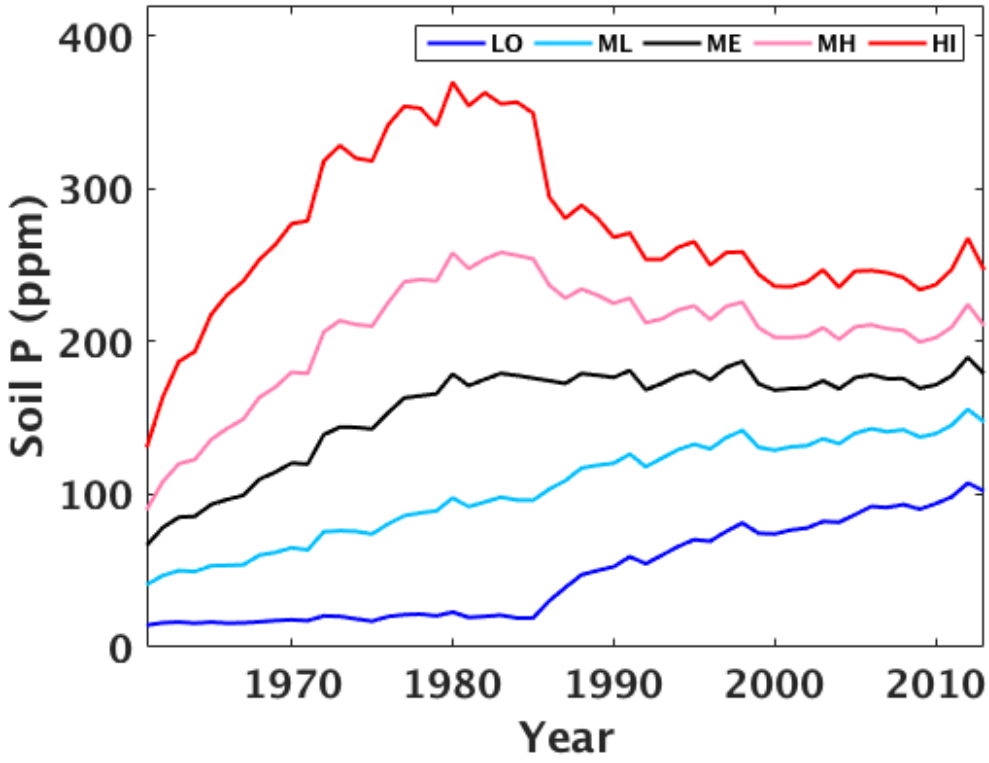


Figure 6. Mean P yield (kg ha^{-1}) in each lake's drainage area for the five simulations, 1986-2013.

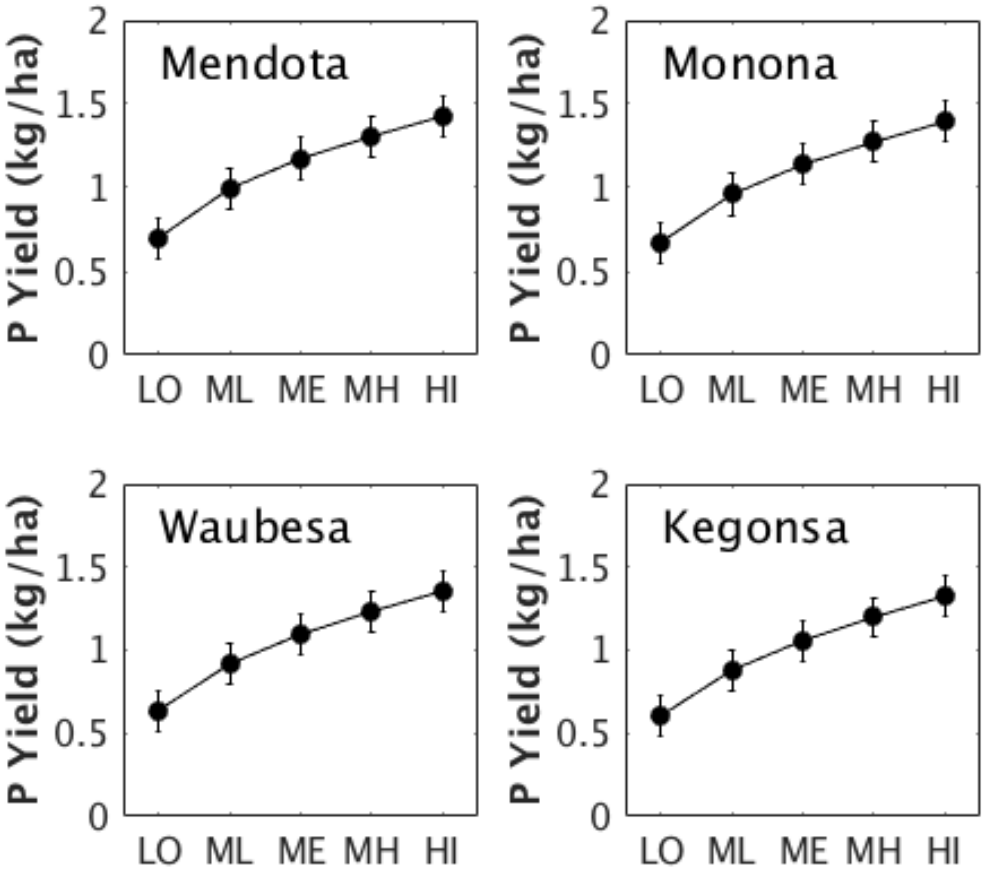


Figure 7. Water quality indicators for each simulation including (A) direct drainage P load (kg y^{-1}), (B) in-lake summer TP concentration (mg L^{-1}), (C) Secchi depth (m), and (D) probability of hypereutrophy. Dashed lines indicate the mesotrophic boundary.

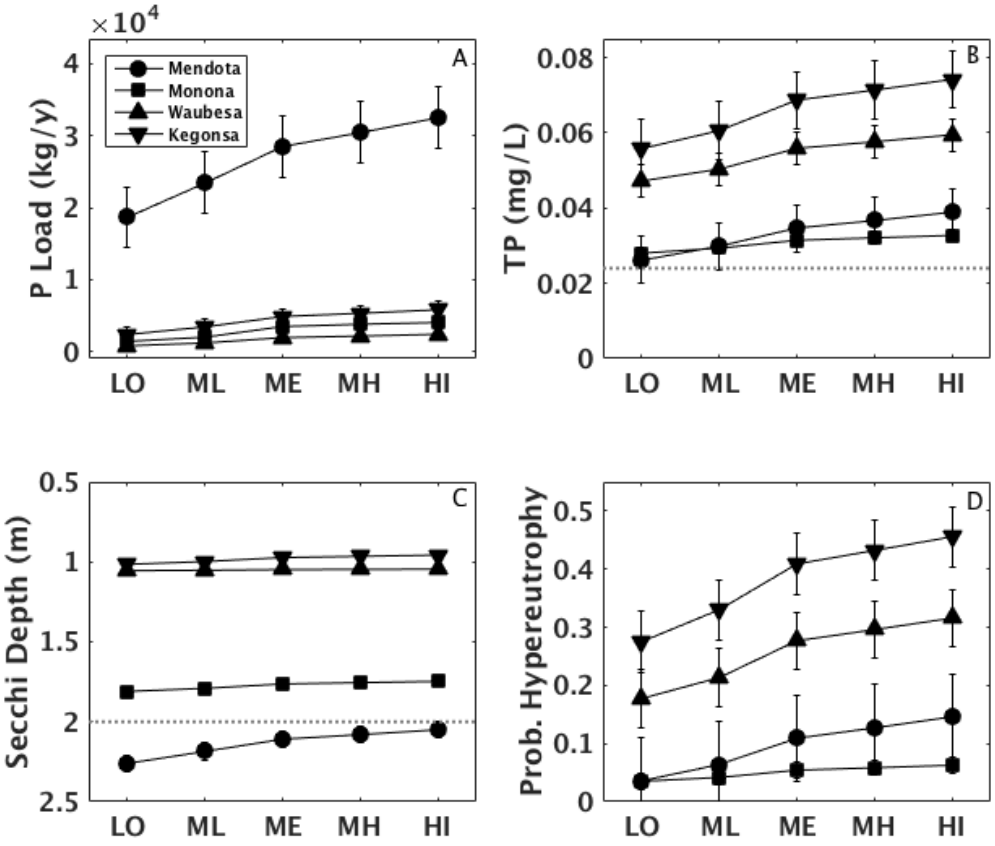


Figure 8. Percent deviation in P yield, direct drainage P load, Secchi depth, and probability of hypereutrophy for each simulation compared with recent historical conditions (ME).

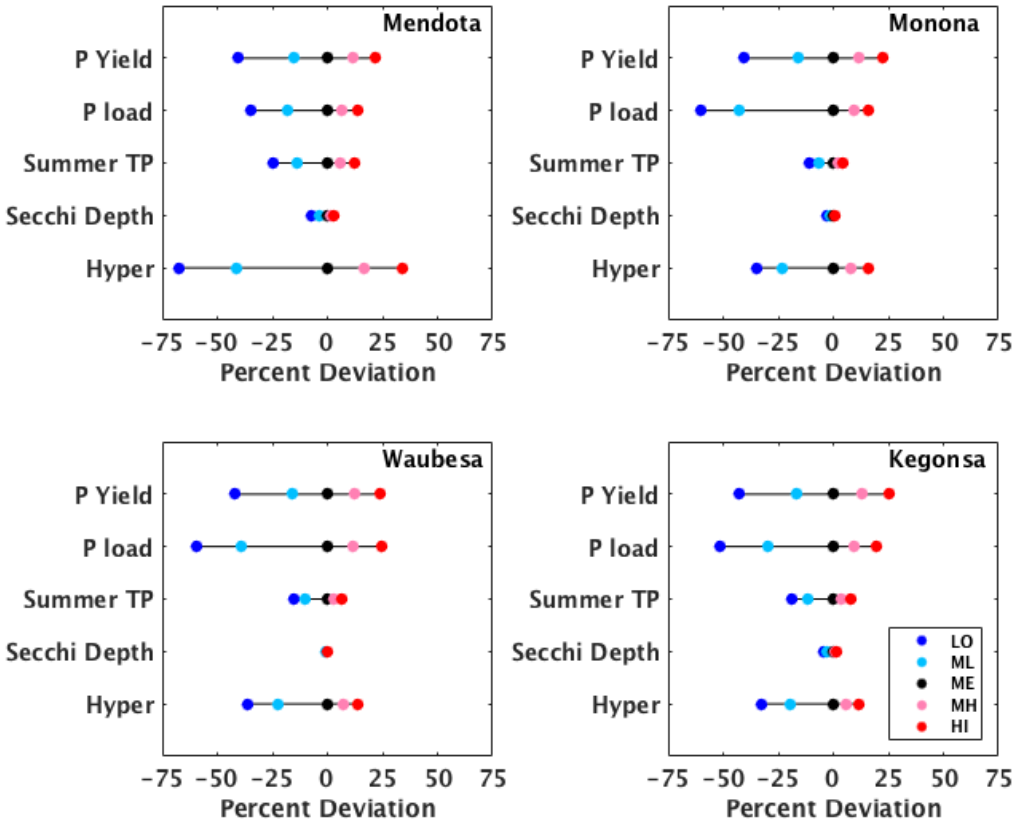
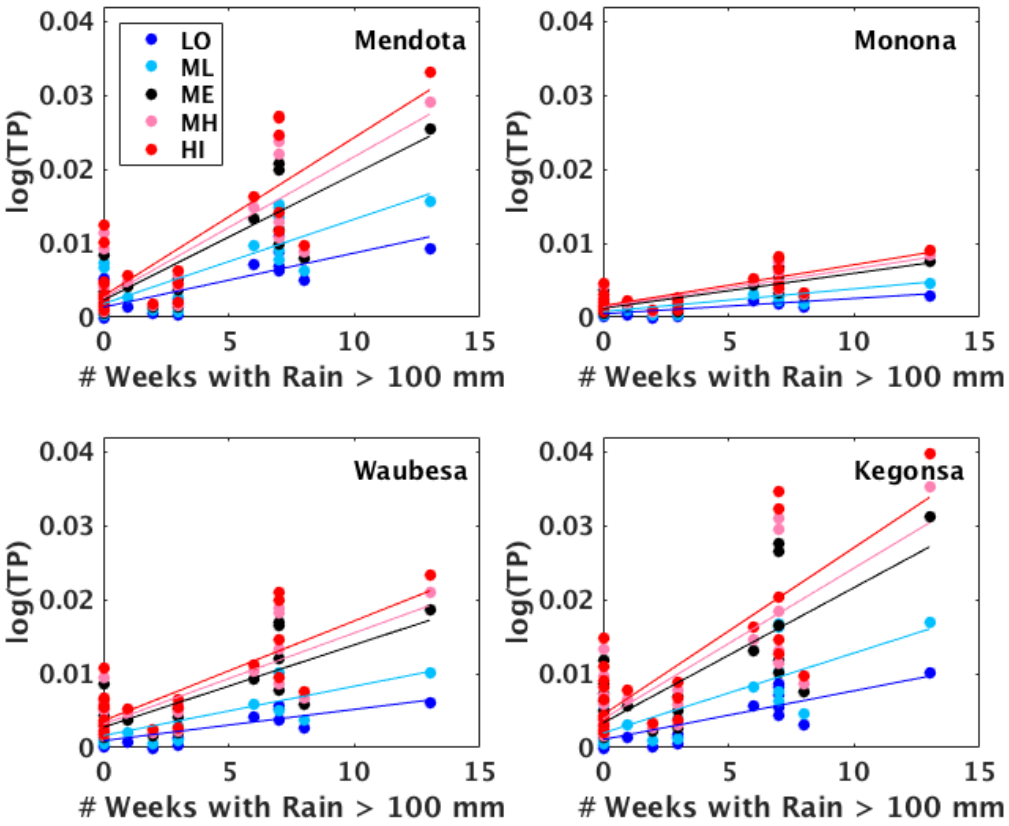


Figure 9. Log of summer TP concentration (mg L^{-1}) versus the number of weeks per year with rainfall exceeding 100 mm.



Appendix A: Overview and updates to Agro-IBIS

The Agro-IBIS model simulates the movement of water, energy, momentum, carbon, nitrogen, and now phosphorus, in both natural and managed ecosystems. The structure of Agro-IBIS has been described in detail (Foley and others 1996; Kucharik and others 2000; Kucharik 2003), and many components of the model have been validated across a range of ecosystems at various spatial and temporal scales (El Maayar and others 2001; Kucharik and Brye 2003; Kucharik and others 2006; Kucharik and Twine 2007; Soylu and others 2014; Zipper and others 2015). Recently, Agro-IBIS was integrated with the variably saturated soil water flow model HYDRUS-1D in order to more accurately simulate the subsurface using the physically based Richards Equation (Soylu et al. 2014, Richards 1931). Soylu and others (2014) validated the new Agro-IBIS model with crop net primary production (NPP), leaf area index (LAI), and soil moisture and temperature performance with plot-level observations on silt loam soil in the YW. Subsequent work by Zipper and others (2015) validated the model's soil moisture and temperature, LAI, and corn yield results across a field with variable soil texture and groundwater depths (0-7 m), also in the YW. Since integration with HYDRUS-1D, the infiltration reduction function was modified to generate saturation-excess overland flow when the head of the surface soil layer exceeds a user-defined puddling depth, allowing for more accurate representation of surface hydrology (Dunne and Black 1970a, 1970b). Additionally, a minimum head threshold parameter from HYDRUS-1D was introduced in order to prevent extreme head gradients, which improved model stability, particularly in sandy soils.

For this study, biogeochemical cycling of P and loss of P to runoff have been added to Agro-IBIS to enable simulation of P dynamics, including interactions between surface water quality, climate, and land management. The new terrestrial P module in Agro-IBIS features P

application, transformation, and loss of dissolved P to runoff; in-soil cycling of organic and inorganic forms of P; and loss of particulate-bound P with erosion (Fig. 2). The SurPhos model was integrated into Agro-IBIS to handle the simulation of inorganic P cycling in manure and soils, and loss of dissolved P to runoff (Vadas and others 2004, 2005, 2007a). SurPhos represents state-of-the-art dissolved P loss modeling for agricultural systems receiving manure. The ability of SurPhos to simulate P release has been validated in laboratory, soil-box, and field-scale settings for a variety of manure types and rainfall regimes (Vadas and others 2004, 2007a, 2007b; Sen and others 2012, Collick and others 2016). SurPhos simulates interconnected pools of dry matter and P on the soil surface, flows of manure P between pools, leaching and assimilation of manure P into soil, and loss of manure P to runoff (Fig. 2). Inorganic P fertilizer application as well as plowing are also simulated by SurPhos. Since SurPhos simulates only the soil P dynamics that are most important for dissolved P loss to runoff, it does not include organic P transformations, dynamic plant growth and uptake, or erosion. However, the model is designed to be incorporated into more complex models that simulate these processes, as is done here. Phosphorus estimation tools including APLE (Vadas and others 2012) and SnapPlus (Good and others 2012) are also based on SurPhos, making the basic methodology employed in Agro-IBIS consistent with tools currently used by managers and researchers in Wisconsin. SurPhos is described in further detail by Vadas and others (2007a) and is summarized in the following section.

SurPhos

Within SurPhos, surface manure P is divided into four pools: inorganic water-extractable P (WEP), organic WEP, inorganic stable WEP, and organic stable WEP. For manure

applications containing less than 15% solids, 60% of manure P slurry is allowed to infiltrate into the soil on the day of application. Stable forms of manure P are transformed to WEP by a decomposition process that is a function of mean daily air temperature and manure moisture and age factors. SurPhos also simulates slow, physical assimilation of manure dry matter and P into soil, representing incorporation by macroinvertebrates, or rain. Required soil, climate, and hydrology input data to SurPhos include initial labile P content, soil textural properties, daily average temperature, rainfall, and runoff. For manure applications, inputs include dates of application, application area, manure dry matter and P content.

For manure slurry applications, slurry WEP that infiltrates into the soil is added to the labile P pool, where P can then be gradually transformed to soil active inorganic P. Slurry inorganic stable P that infiltrates is added to soil active inorganic P. During rain events, inorganic WEP is leached into soil, with 80% added to the surface labile pool where it may interact with runoff via desorption, and 20% infiltrating to deeper layers. For slow, physical assimilation, manure inorganic WEP is added to the soil labile pool, and manure inorganic stable P is added to the soil active inorganic pool. Loss of dissolved inorganic P in runoff from the surface soil labile pool is calculated by multiplying the concentration of labile P by 0.004. Labile P is assumed to be half of Mehlich-3 and Bray-1 soil P (Vadas and White 2010).

Soil P Pools and Organic P Cycling

Soil P pools in Agro-IBIS are represented based on the soil-plant model of Jones and others (1984) which was designed to simulate long-term dynamics of inorganic and organic P for a variety of soils under variable management practices. Other models that simulate agricultural P runoff are based on the model of Jones and others (1984), including SWAT (Neitsch and others

2011), EPIC (Williams 1995), and AGNPS (Young and others 1989). Soil P pools in Agro-IBIS contain three layers with depths of 2.5, 15, and 75 cm, respectively. The three inorganic pools include stable, active, and labile forms which are largely handled by SurPhos. Labile P represents easily desorbed P that is available to both runoff and plant uptake. The stable and active inorganic pools are not easily desorbed to runoff and maintain equilibrium based on a function of soil properties (Vadas and others 2006, 2007a).

Organic soil P pools are also based on the model of Jones and others (1984) with further modifications found in SWAT (Neitsch and others 2011) (Fig. 2). The fresh organic pool is associated with plant residue and microbial biomass. The soil humic pool is partitioned into active and stable components to account for the variation in availability of humic substances to mineralization (Neitsch and others 2011). The soil decomposition equations controlling flows of P between the organic pools and the labile pool are net decomposition equations since they account for both immobilization as well as mineralization. These transformations depend on temperature, soil moisture, and N availability (Dalal 1977; Anderson 1980). For crops, 90% of plant P is assumed to be harvested at the end of each growing season. The remaining 10%, accounting for plant residue, is transferred to the organic fresh pool. For grass PFTs, 100% of plant P is returned to the fresh organic pool on December 31st each year. For trees, 5% of plant P is returned each year (assumed litterfall), with 95% allocated to plant storage. Phosphorus taken up by plants is removed from the surface soil layer labile pool. During model development uptake from additional soil depths was attempted but failed in non-crop systems due to the slow flow of litter-based P through the soil profile. This caused P to build up disproportionately in the surface layer. Restricting uptake to the top layer alleviates this problem in non-crop systems, but because crops may switch to non-crops and vice-versa during simulations, uptake is limited to

the surface layer in all cases to maintain a consistent approach for crops and non-crops. Losses of P in runoff and soil P concentration in croplands are relatively insensitive to single versus multi-layer uptake. Future model enhancements however will focus on representing plant uptake as a function of root distribution as well as plant limitation. In the YW, soil P is generally high enough where plants are not P limited.

Daily plant uptake is simulated using a rate constant that scales daily increments of biomass for crops, and NPP for non-crops. Rate constants were calibrated to achieve typical plant P removal rates in crops (Laboski and Peters 2012). For natural ecosystems, uptake rates were calibrated to achieve realistic plant P fractions in grasses (Ryser and others 1997) and trees (Ovington and Madgwick 1958) as well as to maintain approximate soil P equilibrium, since P fluxes to and from unfertilized soils are relatively small (Walker and Syers 1976), and uptake rates in natural ecosystems are lower than in croplands (Hobbie 1992; Horrocks and others 2014). Dustfall loading of P is represented following Lathrop (1979) and Amy and others (1974), with 0.62 kg ha^{-1} yr^{-1} added to the inorganic soil P pools (Fig. 2).

Erosion, loss of particulate P, and surface runoff

Sediment yield is calculated in each grid cell using the Modified Universal Soil Loss Equation (MUSLE) (Williams 1975). The MUSLE calculates daily sediment yield as:

$$Y = 11.8 \times (Q_{surf} \times Q_{peak})^{0.56} \times K \times C \times LS \times P \quad [1]$$

where Y is daily sediment yield (tons), Q_{surf} is surface runoff (m^3), Q_{peak} is the peak runoff rate (m^3 s^{-1}), K is the soil "erodibility" factor, C is the cover and management factor, LS is the

topographic factor, and *P* is the support practice factor. *K*, *C*, *LS*, and *P* are all unitless factors ranging from 0-1. The coefficient and power constants for the MUSLE, given as 11.8 and 0.56 respectively for metric units, are location-specific parameters that are often calibrated (Sadeghi and others 2014). The soil erodibility factor, *K*, is computed at the beginning of each year using the sand, silt, and clay fractions of the soil as well as the organic carbon content, as done in EPIC (Williams 1995). The cover and management factor, *C*, is assigned based on land cover type. For corn, the *C* factor is calibrated to achieve reasonable rates of sediment yield at the grid-cell level when compared with local observations (Stuntebeck and others 2011). For the remaining biomes, *C* factors are assigned based on the relative differences between *C* factors observed by Panagos and others (2015) for a range of crop and non-crop types. Barren, corn, soy, wheat, alfalfa, pasture, grassland, forest, and wetland are assigned *C* values of 0.9, 0.8, 0.5, 0.45, 0.40, 0.20, 0.10, 0.05, and 0.01, respectively. The topographic factor *LS* is calculated using slope and slope length of each grid cell, as done in the USLE (Wischmeier and Smith 1978) and EPIC (Williams 1995). The support practice factor *P* represents conservation practices for sloping soils vulnerable to erosive rains, like contour tillage, strip-cropping, and terracing. Because slopes are generally small in the YW, and these practices are typically rare, the *P* factor is set to 1.0 throughout the watershed. Daily loss of sediment-bound P is calculated using the total amount of P in the surface layer, daily sediment yield, and a P enrichment ratio (PER) (McElroy and others 1976; Williams and Hann 1978, Sharpley 1985). Total P in the surface layer is the sum of P in all five soil pools including the labile pool, as recommended by Vadas and White (2010). The PER is defined as the ratio of the concentration of P transported with sediment to the concentration in the surface soil layer. It is calculated using a relationship developed by Sharpley (1985) in which there is a linear relationship between the logarithms of soil loss and the PER:

$$\ln(PER) = 1.21 - 0.16 \ln(sed) \quad [2]$$

where *sed* is daily soil loss in kg ha^{-1}.

Surface runoff is calculated based on the following governing equation for water balance in Agro-IBIS:

$$P = Q_{surf} + ET + D + \Delta S \quad [3]$$

where P is precipitation (mm), ET is evapotranspiration (mm), D is drainage exiting the soil column (mm), and ΔS is the change in storage (mm). The model solves the mixed-based Richards equation to simulate water flow in the unsaturated zone (Soylu and others 2014). Surface runoff accumulates as a puddle on the soil surface when the upper soil layer reaches saturation, or the infiltration rate exceeds the maximum infiltration rate allowed. Peak runoff rate, Q_{peak}, is calculated for each grid-cell following the modified rational method which assumes that for a given rainfall event having a constant and uniform intensity over a basin, the rate of runoff will increase until the time of concentration when the entire subbasin is contributing to flow at the outlet. Because hydrologic transport is handled separately by THMB (Terrestrial Hydrology Model with Biogeochemistry, Appendix B), no channel is assumed in land grid cells simulated by Agro-IBIS and time of concentration is computed for overland flow only.

Representation of land cover

Agro-IBIS has been modified to allow simulation of annually changing land cover grids that can contain both crops and non-crops. Additionally, biomes have been modified to accommodate the range of land cover classes found in the YW. Pre-existing biomes in the model that can be readily used in the YW include corn, soy, wheat, temperate deciduous forest, and grassland. No new plant functional types (PFTs) have been developed but new biomes have been added based on pre-existing PFTs. New biomes include alfalfa, hay, wetlands, and urban areas. To simulate alfalfa, C3 grasses are grown. To simulate hay, grassland is grown, consisting of both C3 and C4 grasses. Ninety percent of the aboveground biomass in both alfalfa and hay is harvested at the end of each year, with corresponding adjustments made to both carbon and phosphorus pools. To represent areas of the watershed growing small grains, spring wheat is grown as a proxy. For fruits and vegetables, soy is grown in order to best capture plant phenological trends. To represent wetlands, C4 grasses are grown.

Urban areas are represented as combinations of impervious area and well-maintained turf cover following Schneider and others (2012). Turf cover is represented by growing C4 grasses subject to a maximum canopy height of 11 cm and normal vegetation albedo. Soil albedo is set to 0.28 in all urban cells. Four classes of urban areas are simulated: high-, medium-, and low-intensity, and open space. Except for open space, each of the classes is assigned a fraction of pervious area, assumed to be growing turf cover. These fractions are 10, 35, and 65% respectively for high-, medium-, and low-intensity urban cells. The remainder of each cell is assumed to be impervious. Each urban cell is first simulated as 100% turf cover, and an adjustment is made to water balance terms to account for fraction of the impervious area. For runoff, the Curve Number approach (NRCS 2004) is used to calculate the amount of runoff generated from the impervious area. The pervious and impervious fractions are then used to scale

runoff from turf cover and impervious area, respectively. Additional model quantities are scaled according to the pervious fraction in each urban cell, including ET, drainage, biomass, NPP, LAI, net ecosystem exchange (NEE), sediment yield, and nitrate leached from the soil column. Open space grid cells are left unadjusted, with no assumed impervious cover.

Code Parallelization

Agro-IBIS simulations are costly in terms of CPU time, however the fact that there is no data exchange between land points allows for efficient process level parallelization. To accomplish this, the geographic simulation domain is divided into several smaller subdomains and Agro-IBIS is executed for each of the subdomains as a separate process on its own core. After all subdomain simulations are completed, the subdomain results are assembled to cover the original domain. Domain division, process launch and monitoring, and results assembly are performed automatically by a custom written parallelization software. A 60-core/120-thread machine with four Intel Xeon E7-4890 processors is used to run all simulations. A 250 year simulation (including spin-up) of the Yahara Watershed using a 220 m grid and 25,881 grid cells takes approximately 28 hours of real time.

Appendix B: THMB Overview and Updates

Agro-IBIS is coupled to the THMB model to enable simulation of water, sediment, and P transport at the watershed scale, including delivery of P loads to the YW lakes and the Rock River. The THMB model, formerly called HYDRA, is a physically-based hydrologic routing model that has been coupled to Agro-IBIS and validated in several previous studies (Donner and Kucharik 2003, Donner and others 2004, Donner and others 2008). By linking the topographic data of a prescribed stream network and river morphological characteristics to a set of linear reservoir functions, the model simulates the temporal variability of water flow and storage in hydrologic systems (Coe 1998, 2000; Coe and others 2008). Donner and others (2002) added a nitrogen transport module to the THMB modeling framework. In this study, sediment and P transport have been added to THMB.

Soils in the YW are mostly silt loam, indicating a fine particle size. Therefore, the sediment transport functions in the THMB model have been developed based on a fine sediment transport model from Patil and others (2012). The mass balance of sediment loads in the stream network S_r is described by the following equation:

$$\frac{d(S_r)}{dt} = \Sigma\, S_{in} + S_g\,(1 - A_w) \;+\; E \;+\; I \;-\; L \;-\; \left(\frac{S_r}{T_r}\right) \qquad [4]$$

where S_{in} is the sediment load from upstream cells; S_g is the sediment load from land; A_w is the fractional standing water area in the cell; E is the in-channel erosion rate; I is the sediment load from point sources; L is the sediment removal due to in-channel processes; and T_r is the river

water residence time (Donner and others 2002; Coe and others 2008). Equations for in-channel erosion rate (*E*) and in-channel deposition (*L*) can be found in Patil et al. (2012).

The P loads are divided into two major forms, dissolved P and sediment P. These two forms have different transport mechanisms. Dissolved P is considered a conservative solute with the mass balance of dissolved P in the stream network, P_d^r, described by:

$$\frac{d(P_r^d)}{dt} = \Sigma\, P_{in}^d + \left(\frac{P_s^d}{T_s} + \frac{P_d^d}{T_d}\right)(1 - A_w) + I^d - \left(\frac{P_r^d}{T_r}\right) \quad [5]$$

where P_{in}^d is the dissolved P flux from upstream cells; P_s^d and P_d^d are the dissolved P loads in surface runoff and subsurface drainage from the land, respectively; T_s and T_d are the residence times of cells for surface runoff and subsurface drainage, respectively; and I^d is the dissolved P load from point sources. The transport function of dissolved P is similar to the N transport function developed by Donner and others (2002), which has the capability to be adapted to simulate the transport of different kinds of semi-conservative chemicals. The transport process for sediment P is related to sediment processes. The erosion and enrichment of sediment P on the land are handled by Agro-IBIS, while the in-channel erosion and deposition of sediment P are handled by THMB. The mass balance of sediment P in the stream network, P_r^s, is described by the following equation:

$$\frac{d(P_r^s)}{dt} = \sum P_{in}^s + \left(\frac{P_s^s}{T_s} + \frac{P_d^s}{T_d}\right)(1 - A_w) + E^s + I^s - L^s - \left(\frac{P_r^s}{T_r}\right) \quad [6]$$

where P_{in}^{s},P_{s}^{s} ,P_{d}^{s}, and $I^{s}$are different sources of sediment P loads with similar descriptions as the dissolved P loads; E^{s} is the sediment P load from in-channel erosion; and L^{s} is the sediment P removal via in-channel deposition. In-channel erosion is calculated as:

$$X_p = \frac{k_{XP}}{Y_E^b} \quad [7]$$

$$E = k_E X_p Y_E \quad [8]$$

where $X_p$is the in-channel enrichment ratio; $k_{xp}$is the enrichment coefficient; Y_E is the sediment erosion rate; *b* is the enrichment optimization parameter; and k_E is the sediment P erosion coefficient. Enrichment functions of this form are commonly used in process-based models for P transport simulation (Viney and others 2000). In-channel sediment P deposition is calculated with the following equation:

$$L = D_S \left(\frac{C_{pS}}{C_s}\right) \quad [9]$$

where D_s is the amount of sediment deposition; C_{Ps} is sediment P concentration; and C_s is sediment concentration. In general, the amount of sediment P deposition is considered to have a proportional relationship with the sediment deposition in the stream network. In addition to erosion and deposition of sediment P, the in-channel storage of sediment P is also represented in the model. Erosion of sediment P can be limited depending on the in-channel storage of sediment P.

Internally-drained basins – areas with no contribution to surface runoff but are connected on the subsurface level – are common in the Yahara Watershed (Figure B1) due to the influence of glaciation as recently as 10,000 years ago. As a result, THMB only simulates the subsurface transport processes in these areas. Surface runoff and sediment/nutrient transport are considered to be isolated from the main watershed.

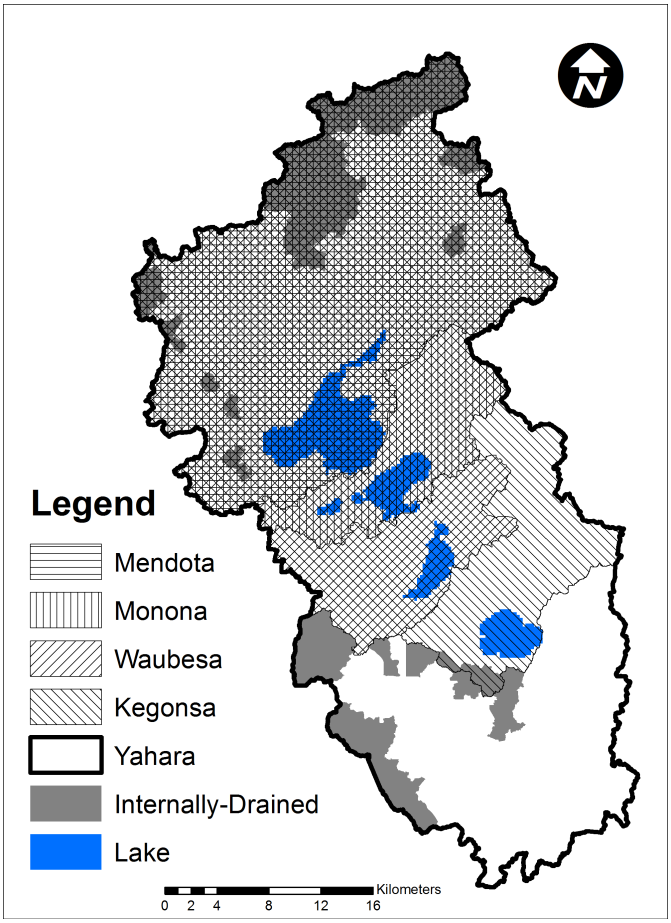


Figure B1. Map showing total drainage areas for each lake as well as internally-drained basins of the Yahara Watershed.

Appendix C: Yahara Water Quality Model

The Yahara Water Quality Model predicts summer water quality variables in the four mainstem Yahara lakes given annual direct drainage loads to each lake (Carpenter and Lathrop 2014). The model first computes P mass balances for the four lakes using mass balance principles and empirical relationships among terms of the mass balance. The mass balance is computed between 1 November and 31 October to correspond with fall-overturn estimates of P mass in the water column of each lake. Then, summer water quality variables are computed using empirical regression models. These models predict summer water quality from terms of the P mass balance. Summer is defined as 30 June to 7 September, which is reliably a period of summer stratification. Water quality variables are TP (total P concentration in surface water), DRP (dissolved reactive phosphorus concentration in surface water), chlorophyll (chlorophyll a concentration in surface water), and Secchi disc transparency. Empirical relationships are computed using 33 full years of data for lakes Mendota and Monona (1975-2008) and 28 full years of data for lakes Waubesa and Kegonsa (1980-2008). Datasets and regression models are described by Lathrop and Carpenter (2013) and Carpenter and Lathrop (2014). Here we explain how the Yahara Water Quality Model is used to compute water quality in the four lakes given direct drainage loads computed by THMB.

Annual loads and exports

The four mainstem lakes, from upstream to downstream, are Mendota (Me), Monona (Mo), Waubesa (Wa), and Kegonsa (Ke). For Lake Mendota, which has phosphorus inputs from

land but no inflow of phosphorus from other lakes, the annual load for year *k*, $L_{k,Me}$ in kg yr^{-1} is the same as the direct drainage load, $M_{k,Me}$:

$$L_{k,Me} = M_{k,Me} \quad [10]$$

For the other lakes, annual loads also include transfers from upstream lakes.

$$L_{k,Mo} = M_{k,Mo} + E_{k,Me}$$
$$L_{k,Wa} = M_{k,Wa} + E_{k,Mo} \quad [11]$$
$$L_{k,Ke} = M_{k,Ke} + E_{k,Wa}$$

In equations [11], *E* is the annual export (kg yr^{-1}). *E* is estimated from *M* using models fitted for each lake as explained below.

Computing P dynamics over each year

For each lake *l*, P dynamics during each year *k* from 1 November to 31 October are modeled as

$$\frac{dP_{k,l}}{dt} = (1 - w_l)\, L_{k,l} - (s_l + h_l)P_{k,l} \quad [12]$$

where *P* is the mass of P phosphorus (kg) in year *k* in lake *l*, *s* and *h* are sedimentation and outflow coefficients for *P,* and *w* is the outflow coefficient for *L*. Comparisons using the Akaike information criterion (AIC) show that model fits are improved significantly by using two outflow

coefficients. From equation 12 it follows that cumulative exports from a lake during each year are given by

$$\frac{dE_{k,l}}{dt} = wL_{k,l} + h_{k,l}\,P_{k,l} \qquad [13]$$

For integration within each year, the initial values of *P* and *E* are the 1 November observation of P mass in the water and 0, respectively. Values of *P* and *E* after 1 year (i.e. 31 October of the next year) are obtained by integrating equations 12 and 13 by the Euler method using 30 time steps.

Appendix D: Model Calibration and Evaluation

Agro-IBIS

The goals for model performance were to generate reasonable values of sediment and P loss at the field scale in Agro-IBIS and also to generate values of in-stream sediment and P load within THMB that agreed with the long-term USGS gage record. To achieve the model performance goals with minimal calibration in Agro-IBIS, calibration was limited to three parameters. These included the MUSLE coefficient and power constants (Eq. 1) as well as the PER (Eq. 2). To achieve satisfactory simulation of sediment yield at both the field scale and in the river network, we calibrated the location-specific MUSLE coefficients, originally given as 11.8 and 0.56 for metric units (Williams 1975). The final calibrated values of these parameters were 5.9 and 0.35, respectively.

To guide P cycling and the magnitude of P loss in runoff, the soil P pools were initialized by first setting the labile pools to roughly correspond with observed values of soil test P for local croplands and prairies (Bennett 2004). As in SurPhos, the labile P pool was assumed to be half of Bray-1 soil test P (Vadas and White 2010). For the calibrated historical run, ME, the P pools were allowed to spin up for 25 years prior to the start of the simulation time period. During the spin-up, manure was applied to all cropland cells at a rate of 50 kg ha^{-1} y^{-1}. By the start of the simulation time period (1986), average surface soil P concentration in croplands was approximately 176 ppm.

To achieve reasonable values of P loss in both the dissolved and sediment form, the PER was calibrated. This was required since without calibration of the PER, simulated sediment P in the river network was significantly higher than the observed gage record and also exceeded

typical edge-of-field loss rates observed locally (Stuntebeck and others 2011). The PER was the optimal parameter to calibrate since simulated sediment yield and dissolved P loads/losses performed well without calibration, but sediment P losses did not. The coefficient used in the equation for the PER (Eq. 2) was calibrated from its original value of 0.16 to 0.02, and an additional multiplier was included in the equation having a value of 0.09.

SurPhos and the soil P routines used in Agro-IBIS have been validated previously (Vadas and others 2004, 2007a, 2007b; Vadas and White 2010; Sen and others 2012, Collick and others 2016). However, we wished to verify that the calibrated version of Agro-IBIS simulated a reasonable range of soil P values within the YW. To do this, we compared simulated soil test P from the calibrated historical run (ME) with three separate sources of local observations as well as observations reported in the literature. Simulated soil P in the second soil layer (2.5-15 cm) (Fig. D1) was compared with a set of soil P observations made in three subwatersheds of the YW in the years 2011-2014 (Kyle Minks, Dane County Land and Water Resources Department, unpublished data). The measurements were made on 1,869 anonymous crop fields spanning a variety of cropping systems and management practices. The data were collected using either the Bray-1 or Mehlich-3 methods. The average soil test P measured across all three subwatersheds was 68 ppm. Simulated soil test P in the second soil layer of all YW croplands, averaged over the 1986-2013 historical period, was 73 ppm. This suggested that the model simulated a mean value of soil P in croplands consistent with local observations.

We also compared mean simulated soil test P from ME with observations made from 2003-2008 at the Discovery Farms and Pioneer Farm (DPF) sites located throughout Wisconsin (Stuntebeck and others 2011). At these sites, (at least) 39 separate soil P measurements were made on 18 fields that spanned a range of landscapes, soil textures, and farming systems typical

of livestock farms in southern Wisconsin. Based on the reported values, we calculated a mean soil test P of 79 ppm, with a standard deviation of 37 ppm. Compared to the average simulated soil P value of 73 ppm, this result further verified reasonable model performance.

Andraski and Bundy (2003) measured soil test P in the 0-2 cm layer at three sites in Wisconsin, with Mehlich-3 measurements ranging from 7-325 ppm, and Bray-1 measurements ranging from 5-274 ppm. Simulated soil test P in the 0-2.5 cm layer for croplands averaged 176 ppm. falling within the ranges observed by Andraski and Bundy. Their observations across soil layers indicated stratification, which they attributed to unincorporated manure treatments in conjunction with no-till. This observation is consistent with other studies showing that soils receiving long-term broadcast application of fertilizer and manure without physical incorporation can accumulate P in the soil surface, increasing the potential for P loss in overland flow (Oloya and Logan 1980; Sharpley and others 1994; Andraski and Bundy 2003; Sharpley 2003; Smith and Warnemuende-Pappas 2015). Sharpley (2003) reported Mehlich-3 values for corn/soy rotations receiving long-term manure applications ranging between 128-961 ppm in the 0-5 cm layer which declined rapidly with depth. In a study by Smith and Warnemuende-Pappas (2015), stratification was observed in corn/soy rotations managed with both no-till and vertical tillage (227 ppm in the 0-5 cm layer compared with 102 ppm in the 5-15 cm layer for no-till; 154 ppm in the 0-5 cm layer compared with 114 ppm in the 5-15 cm layer for vertical tillage.) For our simulations, tillage to a 10 cm depth occurred each year at the time of planting, immediately following the springtime manure application. Both of the remaining manure applications (winter and fall) were unincorporated. Stratification was evident in our simulations (Fig. D1), with cropland soil P in the second soil layer averaging 59% less than the first layer (176 ppm versus

73 ppm). This helps confirm the ability of the model to capture soil P dynamics across soil layers.

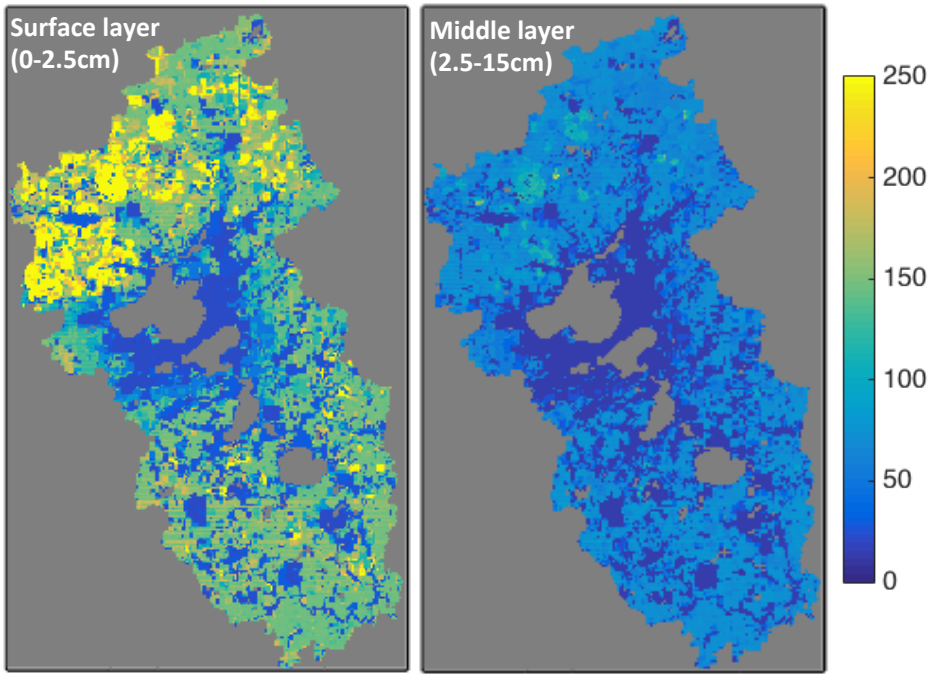


Figure D1. Mean simulated soil test P (ppm) for the calibrated historical simulation (ME) in the top two soil layers, averaged from 1986 to 2013.

Edge-of-field observations on local farms and elsewhere indicate that typical values of sediment and P yield span a wide range of values. At the Discovery and Pioneer farm (DPF) sites, annual sediment yield ranged from 3-5,604 kg ha^{-1}, with a mean of 748 kg ha^{-1} (Stuntebeck and others 2011). Observed average annual total P in runoff ranged from 0.03-7.8 kg ha^{-1}, with a mean of 2.2 kg ha^{-1}. Sediment and dissolved forms of P yield were roughly in proportion (50:50). Simulated values of sediment yield (ME) ranged from 670-2,310 kg ha^{-1}, with a mean of 1,325 kg ha^{-1}. Mean annual simulated TP was 1.0 kg ha^{-1}, and spanned a range of 0.5-2.1 kg ha^{-1}, with the ratio of sediment and dissolved forms roughly 53:47. Modeled values of P yield were generally lower than measurements made at DPF. Results from a modeling study conducted for

grazing-based dairy farms in Wisconsin also suggest that simulated values of P yield from Agro-IBIS tend to be low yet reasonable. In that study, average annual whole-farm P losses were estimated to be 0.5-1.8 kg ha^{-1} (Vadas and others 2015). These values were consistent with values of P yield simulated by Agro-IBIS yet were cited as being low due to low erosion rates on vegetated pastures. Local observations made for the Wisconsin Buffer Initiative observed a wide range in P yield on a single field between just two years, with 22.5 kg ha^{-1} in the first year and <0.1 kg ha^{-1} in the following year (UW-CALS 2005).

The Agro-IBIS simulation (ME) was not designed to mimic the cropping systems and management practices employed at DPF or any other validation site, thus these comparisons serve only as a general check for simulated magnitudes. Further studies from agricultural fields in the U.S. and beyond indicate a wide range in observed sediment and P yield observations, sometimes explained by variation in local environmental or management-based conditions (Wischmeier and Smith 1978; Laflen and others 2004)**,** but often due to unexplained sources of variation (Nearing and others 1999).

THMB

Observations of streamflow, sediment load and P load collected from six USGS gages were used to calibrate and validate THMB (Fig. D2). Continuous monthly streamflow data were split into the periods 2004-2013 for calibration and 1994-2003 for validation. Calibration and validation of sediment load was performed using data from three of the six gages (05427718, 05427965 and 05427948), and for P load from two of the six gages (05427718 and 05427948). THMB was calibrated using a stepwise procedure. First, the simulation time step was chosen to optimize model performance while considering computational expense. The time step was

chosen to be twelve minutes, shorter than the time step of one hour used in previous THMB studies (Coe, 1998; Coe et al., 2008). Second, the streamflow was calibrated based on historical stream gauge data. Major parameters adjusted in this step included: u_{o1}, the minimum effective river flow velocity (u_{o1}=0.35 m s^{-1}); i_o, the reference gradient (i_o=0.0001 m m^{-1}); and p_o, the reference wetted perimeter (p_o=25 m). Physical explanations of these parameters have been previously given by Coe and others (2008). Third, we calibrated sediment transport using the calibrated streamflow. Major parameters adjusted in this step included: e_0, the reference erosion rate (e_0=2.0x10^{-4} kg m^{-2} s^{-1}); and n_p, the sediment porosity (n_p=0.5). Finally, P transport was calibrated based on the calibrated streamflow and sediment loads. Major parameters adjusted in this step included b, a phosphorus enrichment optimization parameter (b=0.27), and k_E, a sediment P erosion coefficient (k_E=0.0002).

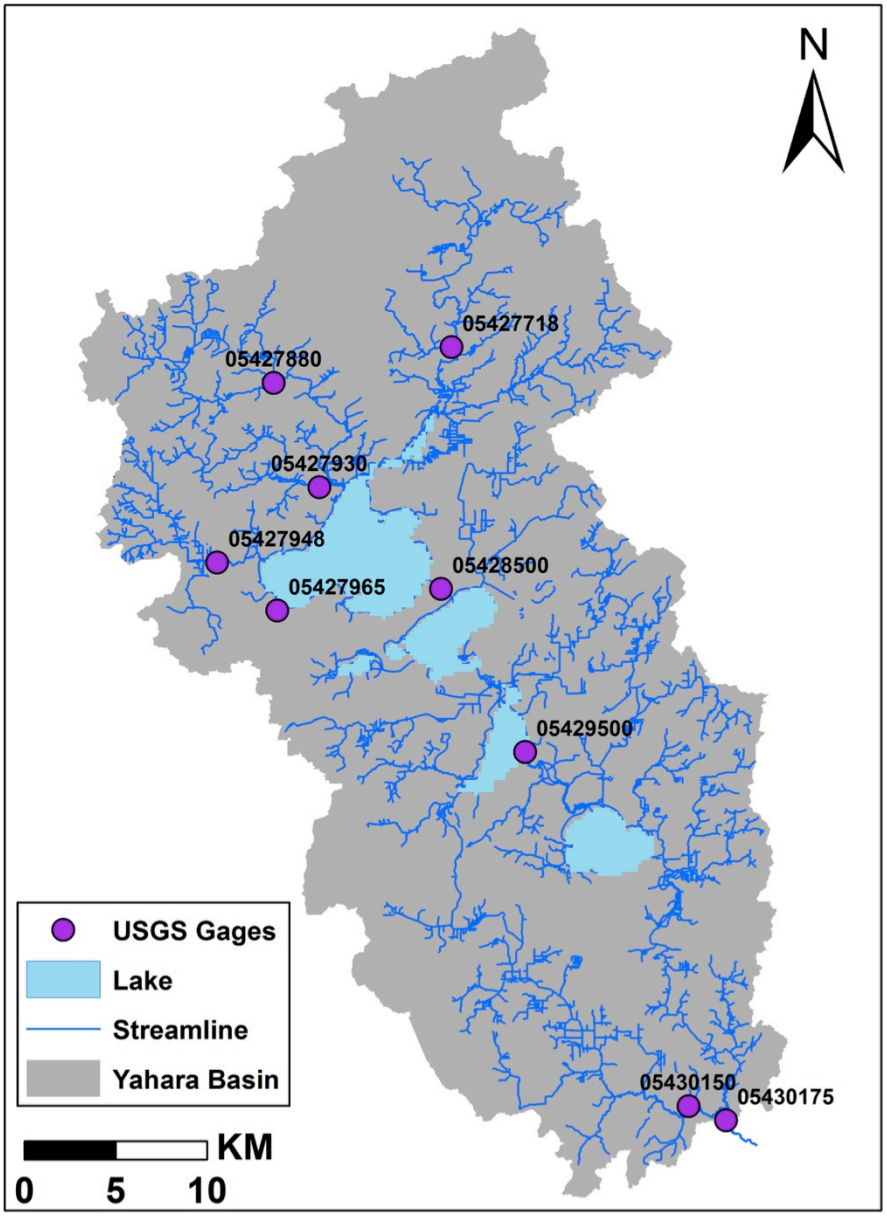


Figure D2. Map of USGS gages in the Yahara Watershed

USGS Gage ID	Calibration			Validation		
	NSE	R^2	RMSE ($m^3 s^{-1}$)	NSE	R^2	RMSE ($m^3 s^{-1}$)
05427718	0.55	0.55	0.46	0.09	0.46	0.31
05427965	0.63	0.78	0.03	0.59	0.77	0.03
05428500	0.49	0.53	1.93	-	-	-
05429500	0.51	0.53	2.75	0.26	0.45	2.40
05430150	0.73	0.82	0.61	0.29	0.60	0.65
05430175	0.66	0.67	3.55	0.44	0.56	3.27

Table D1. Monthly streamflow simulation performance in the Yahara Watershed using Nash-Sutcliffe efficiency (NSE), coefficient of determination (R^2), and root-mean-squared-error.

Table D1 lists monthly streamflow simulation performance at different locations in the Yahara Watershed. The performance varied across locations and was generally acceptable. Simulations of sediment and P loads were closely related to streamflow. As shown in Table D2, the performance of sediment and P transport was worse than streamflow due to the accumulation of errors from streamflow to sediment and P loading. Future model improvements should focus on increasing the accuracy of surface runoff simulation since both sediment and P loads are highly sensitive to changes in surface runoff.

We evaluated simulated P transport using records of direct drainage P loads to Lake Mendota (Lathrop and others 1998; Carpenter and Lathrop 2008; Lathrop and Carpenter 2013). The comparison showed that the model is able to capture high and low P loads with no significant bias (Figure D3). Overall performance was deemed acceptable because annual TP loads to the lakes were required to evaluate water quality outcomes using the Yahara WQ Model. Given that the magnitude of annual loads was well-captured by the model, this satisfied the minimum performance criteria required for our study.

USGS Gage ID	Sediment Mass Flux				Phosphorus Mass Flux			
	Calibration		Validation		Calibration		Validation	
	R^2	RMSE (kg s^{-1})	R^2	RMSE (kg s^{-1})	R^2	RMSE (kg/s)	R^2	RMSE (kg s^{-1})
05427718	0.24	0.27	0.35	0.30	0.29	4.51e-4	0.16	3.99e-4
05427965	0.29	0.16	0.31	0.25	0.24	1.93e-4	0.43	2.07e-4
05427948	0.58	0.012	0.61	0.025	-	-	-	-

Table D2. Monthly sediment and phosphorus transport simulation performance.

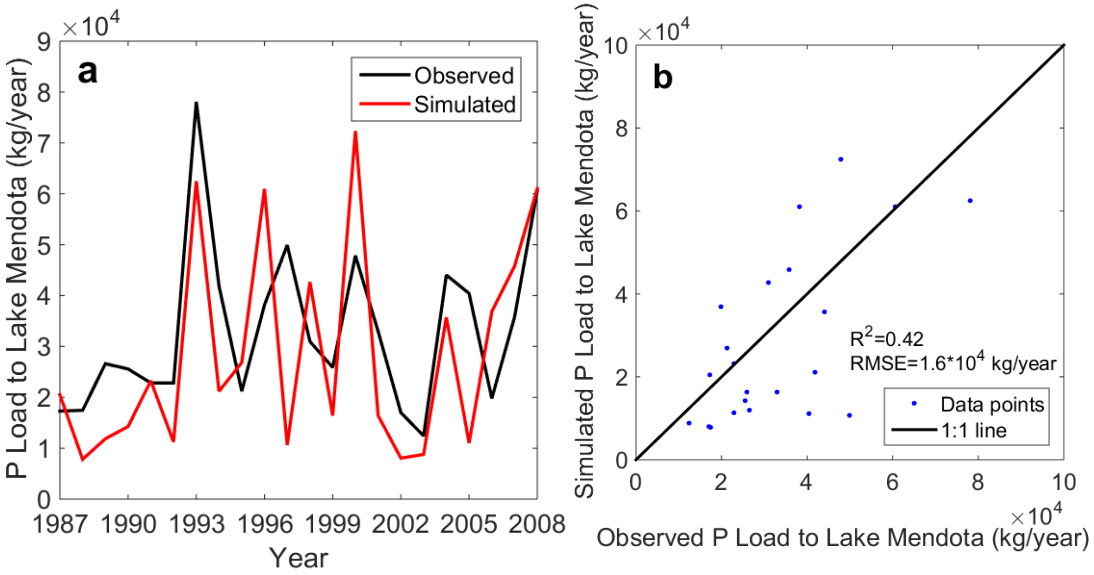


Figure D3. Comparison of modeled annual P load to Lake Mendota and records from previous studies.

Yahara WQ Model

Observations of the P budget and water quality in the four lakes were used to estimate the empirical coefficients used in the Yahara WQ Model. *P*, the mass of phosphorus in each lake at fall overturn, was inventoried for each lake on 1 November of each year (Lathrop and Carpenter 2013; Carpenter and Lathrop 2014). Annual values of *M* and *E* for each lake were measured between 1 November and 31 October for each year (Carpenter and Lathrop 2014). Summer water quality (TP, DRP and Secchi transparency) were measured between 30 June and 7 September each year (Lathrop and Carpenter 2013). Years of observation were 1975-2008 for

lakes Mendota and Monona and 1980-2008 for lakes Waubesa and Kegonsa. Datasets are archived by the North Temperate Lakes Long-Term Ecological Research site (http://lter.limnology.wisc.edu).

Summer chlorophyll was computed from TP using the model of Filstrup and others (2014). Multiple regressions for each lake were used to predict summer water quality (TP, Secchi transparency and chlorophyll) from terms of the annual phosphorus budget (Lathrop and Carpenter 2013). Summer DRP data were used to estimate the probability of hypereutrophy, defined as DRP > 0.005 mg/L. Probability of hypereutrophy was predicted from summer TP using logistic regression.

Regressions to predict summer water quality from terms of the annual phosphorus budget were computed with the lm() package in R. Results of these models are presented in Lathrop and Carpenter (2013) and Carpenter and Lathrop (2014).

Two new model fits were carried out for this paper. Logistic regression to predict the probability of hypereutrophy from summer TP was computed with the glm() package in R using the binomial error distribution. Parameters of equations 12 and 13 (Appendix C) were estimated by fitting the one-year ahead projections of the two equations to the observed annual time series of E, L, and P using maximum likelihood (Hilborn and Mangel 1997).

Comparisons based on AIC showed that the probability of hypereutrophy was best predicted using a single model fit to data from all four lakes, rather than individual models for each lake. Parameter estimates for the logistic regression were intercept = 8.89 (s.e. = 1.91) and log[TP] effect 3.42 (s.e. = 0.69).

Maximum likelihood estimates of *s*, *h*, and *w* obtained by fitting equations 12 and 13 to data are presented in Table D3.

Lake	**s (y^{-1})**	**h (y^{-1})**	**w (y^{-1})**
Mendota	0.342	0.688	0.220
Monona	0.893	0.021	0.538
Waubesa	0.216	0.466	0.879
Kegonsa	0.557	0	0.892

Table D3. Maximum likelihood estimates for parameters *s*, *h*, and *w* of the Yahara WQ Model.

All computations for the YWQM were performed with the base package of R 3.2.3 (R core team 2015).

Appendix E: Climate Inputs

Inputs of daily precipitation, maximum and minimum air temperature, solar radiation, relative humidity, and wind speed are required by Agro-IBIS. Daily precipitation observations from the official NOAA automated weather observation station at the Madison, Wisconsin airport in the north-central part of the watershed were used for 1986-2000 and applied across the model domain (NCDC 2015). Spatially-variable radar-derived daily precipitation estimates (4 km x 4 km resolution) compiled by the National Centers for Environmental Prediction were used for 2001-2013 (NCEP 2015). Daily maximum and minimum air temperature and wind speed from the Madison airport were used for 1986-2013 and applied across the model domain (NCDC 2015). Daily solar radiation and relative humidity data were used from a weather station in Arlington, Wisconsin, representing the closest geographic source of daily solar radiation measurements. These data were used for the 1986-2013 period and applied across the model domain (UW-Extension 2014). Finally, random years between 1986 and 2013 were selected to create the meteorological variables used during the model spin-up period (1786-1985).

Appendix F: Soil Inputs and Groundwater Specifications

Soil layers simulated by Agro-IBIS included a 0.5 cm surface layer followed by 99 deeper layers of linearly increasing thickness. Soil textural class in each grid cell was assigned based on 11 possible categories in the USDA Soil Survey Geographic Database (SSURGO) (USDA 2013). Associated physical properties including percent sand, silt, and clay were then obtained for each class based on data in Campbell and Norman (1998). For each soil class, Van Genuchten soil water characteristic curve relationships (Van Genuchten 1980) were used with retention parameters from Carsel and Parrish (1988). Soils in the YW are primarily silt loam, but to minimize convergence issues within the soil physics routines of HYDRUS-1D, the soil texture was assumed constant throughout the soil profile using properties of the top 2.5 cm surface layer, and clay and sandy soils were substituted with silt loam and loam, respectively, which account for about 15% of grid cells in the watershed. To set the pressure head bottom boundary condition in Agro-IBIS, a constant soil profile depth of 10 m was set across the model domain. The bottom boundary condition for the variably-saturated groundwater flow equation was either (1) a constant head equal to the difference between the soil column depth and the water table depth (WTD) if the WTD was less than 10 m; or (2) free drainage if the WTD was greater than or equal to 10 m. Values for the WTD were determined based on a steady-state, present day (2014) groundwater flow simulation for Dane County (Parsen and others 2016).

Appendix G: LULC Determination

To generate land cover grids for the watershed in each year of simulation, land-use/land-cover (LULC) was divided into 17 biophysically distinct categories and interpolated to a 220 m by 220 m spatial resolution. For recent years (2001-2013), maps were based on land use data from Dane, Columbia, and Rock counties, and national-level datasets including the National Land Cover Database (Homer and others 2015), and the Cropland Data Layer (USDA 2014) (see Gillon and others 2016 for more details). We estimated LULC in 1786 using a pre-settlement vegetation map of Wisconsin that was created using U.S. General Land Office Notes from the mid-1800s, compiled by Finley in 1976 and later digitized (WDNR 2011). The landscape at the time of the land surveys was dominated by oak forest, oak openings, prairie, and wetlands. While Euro-American settlement of the watershed began in the 1830s, we assumed, based on historical U.S. agricultural census data for Dane County (UVGSDC 2004), that LULC changed from the pre-settlement condition to peak agricultural development in 1870. This LULC condition remains until 1890 when urban areas begin to emerge. We digitized urban areas from Dane County plat maps for 1890, 1911, and 1953 as well as from Wisconsin Land Economic Inventory for 1929. We also used digitized tax parcel maps from 1962, 1973, and 1981 to estimate urban areas based on a maximum parcel area threshold of 6 hectares. Between 1992 and 2000, urban areas were assigned using the Wisconsin Land Cover database derived from satellite imagery captured in 1992 (WDNR 2013). Prior to 1920, all non-urban areas except open water and wetlands were classified as corn. From 1920 to 2000, LULC for these non-urban areas was rotated using LULC from 2008 to 2014. While non-urban LULC was likely different between these two periods, we made this assumption in order to institute crop rotations within the model and capture its important effect on the spin-up of carbon and nitrogen.

Appendix H: Determination of manure and fertilizer rates

For nutrient inputs, historical manure P and N production was determined based on estimates of watershed animal units and milk production. First, livestock numbers (cattle, swine, poultry, horses, and sheep) and milk production were determined and interpolated for each year between 1870 and 2014 based on historical agricultural census data for Dane County (Hibbard 1904; USDA-NASS 2013, 2014). Next, county animal units were calculated for each livestock type based on the interpolated animal numbers and historical estimates of animal production weights (USDA-ERS 2015).

In order to scale the county estimates to the Yahara watershed and spatially distribute manure production, we developed a livestock operation inventory for 2001-2014 based on unpublished spatial data from the Dane County Land & Water Resources Department, Concentrated Animal Feeding Operation data from the Wisconsin Department of Natural Resources (WDNR 2015), and milk producer locations from the Wisconsin Department of Agriculture, Trade, and Consumer Protection (WDATCP 2015). This inventory includes livestock type, estimated number of animal units, and estimated manure hauling distances. In addition, we used historical aerial imagery from Google Earth to document whether operations were active or inactive in a given year. Due to data limitations, we assumed that all operations in 2001 existed from 1870 to 2000 even though the total number of farms has decreased through time (UVGSDC 2004). Based on the 2001-2014 analysis, it was determined that the number of animal units in the Yahara watershed is approximately 38% of that for Dane County. We assumed that this percentage is constant from 1870 to 2000.

Manure P and N production (as excreted) per animal unit for each livestock type (excluding dairy cows) was then calculated for each livestock operation based on guidelines

from the (USDA 2008). Dairy cow manure P and N production were separately calculated using regression equations based on milk production (Nennich and others 2005), which has in addition to manure steadily increased per cow over the past half-century (Gillon and others 2016). The estimated manure N was reduced by 70% at all operations to account for ammonia volatilization and first year crop availability (Laboski and Peters 2012) which can vary substantially due to storage/collection characteristics and weather conditions (Rotz and others 2014; Powell and Rotz 2015). Land available for spreading manure P and N from a single operation included all crop and pasture land within a radius equal to the estimated manure hauling distance. More details on the method for land application including treatment of overlapping hauling areas can be found in Appendix I.

Following the determination of manure P and N application rates, fertilizer P and N application rates were calculated for the following LULC types: alfalfa, corn, fruits/vegetables, small grains, and soybeans. Recent (2007-2013) fertilizer application rates were determined based on University of Wisconsin – Extension nutrient application guidelines (Laboski and Peters 2012) assuming high yield potential soils, optimal soil nutrient status, and high yield goals. If manure had already been applied in a grid cell, additional fertilizer was applied at a lower rate only if the LULC type was corn. Application rates for 1945 to 2006 were determined based on county-level fertilizer sales data compiled by Alexander and Smith (1990) and Gronberg and Spahr (2012). These values were used to create a ratio of total fertilizer sold in a given year to that in 2006; this ratio was then multiplied by the 2007-2013 rates.

For all simulations, annual manure rates were divided into three applications per year. In the first application, 10% was applied on February 15th. In the second application, 45% of the annual rate was applied just prior to spring planting, a date that is determined in Agro-IBIS based

on climate. In the third application, the remaining 45% was applied on October 1st. Fertilizer was applied in a single application also just prior to planting. Plowing to a 10 cm depth also occurred at the time of planting, immediately following the springtime manure application. The winter and fall manure applications were unincorporated.

Appendix I: Algorithm for determining spatial application of manure and fertilizer rates

This appendix documents the logic used to determine manure and fertilizer applications of nitrogen and phosphorus to grid cells in the YW. Manure is estimated for each livestock operation based on animal units and livestock type. This manure is spread on available land within the estimated hauling distance radius. Overlapping of manure by different operations is accounted for and explained. Fertilizer application rates are based on University of Wisconsin-Extension recommendations.

Step-by-step description of MATLAB script

- The modeling grid is loaded with a grid cell size of 219.456 m by 219.456 m.
- The historical land-use, land cover (LULC) data are loaded.

The following categories of LULC are available for manure and/or fertilizer:

Code	LULC Category
11	Alfalfa
12	Corn
13	Fruits/Veg
14	Small Grains
15	Soybeans
16	Hay
17	Pasture

- The livestock operation characteristics dataset is loaded.
- Data on total number of animal units in YW and total milk production per year for the historical time period (1870-2014) is loaded. The baseline estimate for 2013 uses data from USDA-Agricultural Census and USDA-NASS surveys on Dane County livestock numbers and milk production (USDA-NASS 2013). The conversion from Dane County animal units to Yahara watershed animal units is based on the 2010 ratio, which was calculated to be 38%.
- Historical animal production weights – based on data from the USDA-Economic Research Service (USDA-ERS 2015) – are loaded.
- Minimum manure (wet) application rate is defined to be 23,000 kg ha^{-1}. This variable is based on estimates from surveys of south-central Wisconsin dairy farmers by (Powell and others 2005).
- The number of livestock operations for each year is loaded. The active livestock operations are determined for each year.
- The first nested *for* loop calculates the sum of all animal units (the 2010 estimate) at all active livestock operations in the watershed. This value is then used to calculate the animal unit ratio which is the ratio of the current year Yahara animal units to the 2010 Yahara animal units. This ratio later scales the animal units at each livestock operation to account for changes through time at the county scale.
- The next nested *for* loop is started and loops through each active livestock operation. First, manure (wet, dry, phosphorus, and nitrogen) production is estimated for each operation. Second, the manure is spread according to the estimated hauling distance and

the LULC available with some modifications based on the minimum manure application rate threshold and overlapping manure hauling areas (explained further below).

- The production of manure (wet, dry, phosphorus, nitrogen) from livestock (except lactating cows) is calculated based on the USDA-Agricultural Waste Management Field Handbook (USDA 2008). The production of lactating cow manure is estimated based on a regression model using milk production as a predictor from Nennich and others (2005). Note that manure is calculated on an as-excreted basis.
- The spatial distribution of manure is done over the following steps:
    - The grid cells that 1) fall within the circular area with radius equal to the estimated manure hauling distance; and 2) are cropland or pasture are identified ('haulCells')
    - If, for some reason, no cells are identified, then the closest cell that is cropland or pasture is identified (a 'flag' variable is used to identify whether this ever happens)
    - The available haul area is calculated
    - A series of if statements handle the different cases…
    - if no manure has been spread in the haul area yet, then the manure is spread everywhere evenly; but if the minimum manure application rate threshold is not reached, then the available haul area is iteratively "shrunk" from the outside-in until the threshold is reached
    - else, some manure has been spread somewhere in the haul area (overlapping case)
        - if at least half of the available haul area is still "dry", then spread manure evenly to just the dry cells (but the minimum manure application rate must

be satisfied…if not then the dry cells are iteratively "shrunk" from outside-in until threshold is reached)

- elseif there are no dry cells at all (all available haul cells have at least some manure applied already), first sort the cells by increasing manure N, then increasing distance. Manure is added proportionally more to the beginning of the list compared to the end so that the first gets twice as much as evenly spread and the last gets nothing.
- else (if some dry cells exist but they sum to less than half of the available haul area), manure is applied to the dry cells at twice the rate as if all the haul cells were dry and the remaining amount is added to the wet cells (but first, the rate for the dry cells needs to be checked to see if it is above the minimum rate threshold…if not, then the available wet and dry areas are iteratively "shrunk" from outside-in).

- the livestock operation loop is closed

- Next, the fertilizer N and P application rate is determined
  - The next nested *for* loop is started that loops through every grid cell
  - if the cell has manure applied, then fertilizer is only applied if it is corn
  - else (no manure has been applied), fertilizer is applied to all crops
  - Application baseline rates are taken from UW-Extension recommendations (A2809) and determined based on the crop type and an assumption of "optimum" soil conditions. These rates are then multiplied by the relative fertilizer application rate for each year

Appendix J: Parameter Sensitivity Analysis

Agro-IBIS

To identify parameters that exert a strong influence on P yield, a simple, one-at-a-time parameter sensitivity analysis was conducted for Agro-IBIS. Thirty-three parameters used to calculate sediment, dissolved, and total P yield were examined (Table J1). Two single-cell simulations were conducted for each parameter, both using a grid-cell located in the northern region of the YW. This cell was chosen to represent soil texture and cropland management practices typical to the northern YW. Continuous corn was simulated from 1986-2013 using nominal annual fertilizer and manure applications of 12.9 and 28.9 kg ha^{-1}, respectively. Historical climate data were used and silt loam was the simulated soil texture. The two simulations were conducted by varying the nominal value of each parameter by +1 and -1%, for a total of 66 simulations. Changes in the response variables (sediment, dissolved, and total P yield) greater than 1% could indicate over-sensitivity to a particular parameter with the influence of such a parameter becoming amplified within the model. Changes in the response variables were analyzed for each simulation, and those parameters inducing a response greater than 1% were identified.

The results showed that no parameters induced a change in total P yield greater than 1% (Table J1), implying that none of the parameters exerted a disproportionate influence on modeled P losses. This was also the case for dissolved P yield, for which the amount of manure applied (MANPAPP) exerted the largest influence (0.70%). For sediment P yield, two parameters induced a change slightly greater than 1%. These were the puddle normalization constant (1.03%) and the MUSLE power (1.18%). The puddle normalization constant represents a

maximum puddle depth for which standing water is allowed to collect. Water collecting in excess of the puddle depth is treated as runoff; therefore the puddle depth is an important control on the amount of runoff generated. The influence of runoff on the transport of P from soil, in both dissolved and sediment forms, explains why this parameter had a relatively large impact on total P yield (0.90%), dissolved P yield (0.64%), and sediment P yield (1.03%). Within the MUSLE, daily runoff as well the peak runoff rate are used to calculated sediment yield (Eq. 1). The original value of the MUSLE power given by Williams (1975) was 0.56. For this study, calibration of this location-specific parameter yielded a value of 0.35, a value that helped generate both reasonable amounts of sediment yield at the field scale and sediment load within streams (Appendix D). The sensitivity analysis showed that this parameter exerted an important influence on sediment P yield (1.18%) as well as total P yield (0.79%). As expected, it had no influence on dissolved P yield (0.01%). The other MUSLE parameters, including the coefficient (originally given as 11.8 by Williams (1975) for metric units, and calibrated here to 5.9), and the *C*, *K*, and *LS* parameters, also exerted a relatively strong influence on sediment P yield (0.97%) and total P yield (0.65%). Since these parameters all act as simple multipliers in the MUSLE (Eq. 1), their influence was identical in this experiment.

Table J1. Parameter sensitivity results for Agro-IBIS showing the maximum absolute percent change induced in P yield for a 1% change in each parameter.

Agro-IBIS Parameter	Description	Nominal Value	Total P Yield (%)	Sediment P Yield (%)	Dissolved P Yield (%)
INITLAB	Initial labile P concentration, top layer	40 ppm	0.19	0.24	0.09
CLAY	Clay percentage in soil	20%	0.01	0.01	0.00
SOM	Organic matter percentage in soil	7%	0.00	0.00	0.00
BD	Soil bulk density	1.5 g cm^{-3}	0.60	0.75	0.30
FERTPKG	Annual fertilizer application.	12.9 kg ha^{-1}	0.17	0.18	0.14
MANPAPP	Annual manure application.	28.9 kg ha^{-1}	0.47	0.35	0.71
MWIPFRAC	Manure inorganic water extractable P fraction	0.50	0.10	0.05	0.19
MWOPFRAC	Manure organic water extractable P fraction	0.05	0.01	0.00	0.02
DRYMAT	Dry matter content of manure applied.	9%	0.00	0.00	0.00
PCTCOV	Percent of field covered by manure application	99%	0.01	0.01	0.04
EFFINC	Surface incorporation efficiency	30%	0.09	0.02	0.25
FMIX	Soil mixing efficiency	30%	0.38	0.48	0.18
DEPTIL	Depth of tillage	10 cm	0.00	0.00	0.00
betares	Fresh P mineralization coefficient	0.06	0.00	0.00	0.00
betamin	Humic P mineralization coefficient	0.002	0.00	0.00	0.00
cnrres	C:N ratio of residue	25	0.00	0.00	0.00
fracnsta	Fraction of soil stable N.	1%	0.00	0.00	0.00
fracnact	Fraction of soil active N.	99%	0.01	0.01	0.00
pctpharv	Fraction of plant P returned to organic fresh pool at harvest.	10%	0.03	0.04	0.02
uptake	Plant uptake rate coefficient.	0.004	0.34	0.42	0.17
INFIL	Slurry P infiltration fraction	0.40	0.07	0.14	0.50
PER	P enrichment ratio	1.0	0.66	0.99	0.00
K	MUSLE K factor	0.30	0.65	0.97	0.00
LS	MUSLE LS factor	0.35	0.65	0.97	0.00
C	MUSLE C factor	0.80	0.65	0.97	0.00
manning	Mannings roughness coefficient	0.05	0.00	0.00	0.00
hydcond	Saturated hydraulic conductivity	1.89e-6 m s^{-1}	0.68	0.73	0.71
puddle	Puddle normalization constant	30 mm	0.90	1.03	0.64
zmin	Runoff threshold	0.03	0.74	0.75	0.70
zmax	Runoff threshold	0.30	0.44	0.65	0.08
PSP	P sorption parameter	0.17	0.13	0.35	0.32
POW	MUSLE power	0.35	0.79	1.18	0.01
COEFF	MUSLE coefficient	5.9	0.65	0.97	0.00

THMB

The sensitivity analysis conducted for THMB used streamflow, sediment load, and P load as the output variables. For the THMB sensitivity simulations, the entire watershed was simulated over the historical time period using climate and land use inputs identical to those used in ME. Seven parameters that have major influence on streamflow, sediment, and P processes were examined. For streamflow, parameters included u_{o1}, the minimum effective river flow velocity (u_{o1}=0.35 m s^{-1}); i_o, the reference gradient (i_o=0.0001 m m^{-1}); and p_o, the reference wetted perimeter (p_o=25 m). For sediment load, parameters included e_0, the reference erosion rate (e_0=1.0x10^{-5} kg $m^{-2}s^{-1}$) and n_p, the sediment porosity (n_p=0.1). For P load, parameters included b, a phosphorus enrichment optimization parameter (b=0.27), and k_E, a sediment P erosion coefficient (k_E=0.0005). Each parameter was varied by ± 1% and the percentage change in each output variable was calculated. For streamflow and sediment load, the changes were calculated based on mean annual watershed values. For P load, the changes were based on mean annual P loads to Lake Mendota. The results are shown in Table J2.

Table J2. Parameter sensitivity results for THMB showing the percent change induced in streamflow, sediment load, and direct drainage P load for a 1% change in each parameter.

THMB Parameter	Description	Nominal Value	Streamflow (%)	Sediment Load (%)	Phosphorus Load (%)
u_{o1}	velocity	0.35 m s^{-1}	0.0012	0.22	-0.11
i_o	Reference gradient	1e-4 m m^{-1}	-0.0002	0.25	-0.11
p_o	Reference wetted perimeter	25 m	-0.0002	0.08	-0.13
e_0	Erosion rate	2e-4 kg $m^{-2}s^{-1}$	0	0.25	0.01
n_p	Sediment porosity	0.5	0	2.72	0.27
b	Phosphorus enrichment optimization parameter	0.27	0	0	-0.08
k_E	Sediment P erosion coefficient	2e-4	0	0	0.07

All parameters showed a reasonable amount of influence on their related variables, with most of them lower than 1%. The influence of streamflow parameters on sediment load and P load, and the influence of sediment load parameters on streamflow and P load, were included only as a matter of interest. Among the three output variables examined, changes in sediment and P loads were larger than changes in streamflow, especially when the parameters related to streamflow were changed. Sediment loads in the stream network have power law relationships with streamflow. Additionally, sediment P transport processes follow sediment transport, in terms of erosion and deposition. Therefore, changes in streamflow will be magnified in sediment and P loads. Parameter adjustments did not always cause same-sign changes among the output variables. This highlighted the nonlinearity of the relationships between the parameters and the variables.

References

Alexander RB, Smith RA. 1990. County-level estimates of nitrogen and phosphorus fertilizer use in the United States, 1945 to 1985. Open-file report 90-130. Reston, VA: U.S. Geological Survey http://pubs.usgs.gov/of/1990/ofr90130/

Amy G, Pitt R, Rameshawar-Singh, Bradford WL, LaGraff MB. 1974. Water quality management planning for urban runoff. Washington, D.C.: U.S. Environmental Protection Agency

Anderson G. 1980. Assessing organic phosphorus in soils. In: The Role of Phosphorus in Agriculture. American Society of Agronomy, Crop Science Society of America, Soil Science Society of America. pp 411–31.

Andraski TW, Bundy LG. 2003. Relationships between phosphorus levels in soil and in runoff from corn production systems. J Environ Qual 32:310–6.

Bennett EM, Carpenter SR, Clayton MK. 2004. Soil phosphorus variability: scale-dependence in an urbanizing agricultural landscape. Landsc Ecol 20:389–400.

Campbell GS, Norman JM. 1998. An Introduction to Environmental Biophysics. Second Edition. Springer Science & Business Media

Carpenter SR, Lathrop RC. 2008. Probabilistic estimate of a threshold for eutrophication. Ecosystems 11:601–13.

Carpenter SR, Lathrop RC. 2014. Phosphorus loading, transport and concentrations in a lake chain: a probabilistic model to compare management options. Aquat Sci 76:145–54.

Carsel RF, Parrish RS. 1988. Developing joint probability distributions of soil water retention characteristics. Water Resour Res 24:755–69.

Coe MT. 1998. A linked global model of terrestrial hydrologic processes: Simulation of modern rivers, lakes, and wetlands. J Geophys Res 103:8885–99.

Coe MT. 2000. Modeling Terrestrial Hydrological Systems at the Continental Scale: Testing the Accuracy of an Atmospheric GCM. J Clim 13:686–704.

Coe MT, Costa MH, Howard EA. 2008. Simulating the surface waters of the Amazon River basin: impacts of new river geomorphic and flow parameterizations. Hydrol Process 22:2542–53.

Collick AS, Veith TL, Fuka DR, Kleinman PJA, Buda AR, Weld JL, Bryant RB, Vadas PA, White MJ, Harmel RD, Easton ZM. 2016. Improved Simulation of Edaphic and Manure Phosphorus Loss in SWAT. J Environ Qual. https://dl.sciencesocieties.org/publications/jeq/abstracts/0/0/jeq2015.03.0135

Dalal RC. 1977. Soil organic phosphorus. Adv Agron 29:83–117.

Donner SD, Coe MT, Lenters JD, Twine TE, Foley JA. 2002. Modeling the impact of hydrological changes on nitrate transport in the Mississippi River Basin from 1955 to 1994. Global Biogeochem Cycles 16:1–19.

Donner SD, Kucharik CJ. 2003. Evaluating the impacts of land management and climate variability on crop production and nitrate export across the Upper Mississippi Basin. Global Biogeochem Cycles 17:1–16.

Donner SD, Kucharik CJ. 2008. Corn-based ethanol production compromises goal of reducing nitrogen export by the Mississippi River. Proc Natl Acad Sci U S A 105:4513–8.

Donner SD, Kucharik CJ, Foley JA. 2004. Impact of changing land use practices on nitrate export by the Mississippi River. Global Biogeochem Cycles 18:1–21.

Dunne T, Black RD. 1970a. An experimental investigation of runoff production in permeable soils. Water Resour Res 6:478–90.

Dunne T, Black RD. 1970b. Partial area contributions to storm runoff in a small New England watershed. Water Resour Res 6:1296–311.

El Maayar M, Price DT, Delire C, Foley JA, Black TA, Bessemoulin P. 2001. Validation of the Integrated Biosphere Simulator over Canadian deciduous and coniferous boreal forest stands. J Geophys Res 106:14339.

Filstrup CT, Wagner T, Soranno PA, Stanley EH, Stow CA, Webster KE, Downing JA. 2014. Regional variability among nonlinear chlorophyll—phosphorus relationships in lakes. Limnol Oceanogr 59:1691–703.

Foley JA, Prentice IC, Ramankutty N, Levis S, Pollard D, Sitch S, Haxeltine A. 1996. An integrated biosphere model of land surface processes, terrestrial carbon balance, and vegetation dynamics. Global Biogeochem Cycles 10:603–28.

Gillon S, Booth EG, Rissman AR. 2016. Shifting drivers and static baselines in environmental governance: challenges for improving and proving water quality outcomes. Regional Environ Change 16:759–75.

Good LW, Vadas P, Panuska JC, Bonilla CA, Jokela WE. 2012. Testing the Wisconsin Phosphorus Index with year-round, field-scale runoff monitoring. J Environ Qual 41:1730–40.

Gronberg JAM, Spahr NE. 2012. County-level estimates of nitrogen and phosphorus from commercial fertilizer for the conterminous United States, 1987–2006. Scientific Investigations Report 2012-5207. Reston, VA: U.S. Geological Survey http://pubs.usgs.gov/sir/2012/5207/

Hibbard BH. 1904. The history of agriculture in Dane County, Wisconsin. Madison, WI: University of Wisconsin http://digital.library.wisc.edu/1711.dl/WI.HistAgDane

Hilborn R, Mangel M. 1997. The ecological detective: confronting models with data. Princeton, N.J.: Princeton University Press

Hobbie SE. 1992. Effects of plant species on nutrient cycling. Trends Ecol Evol 7:336–9.

Homer CG, Dewitz JA, Yang L, Jin S, Danielson P, Xian G, Coulston J, Herold ND, Wickham JD, Megown K. 2015. Completion of the 2011 National Land Cover Database for the conterminous United States-Representing a decade of land cover change information. Photogramm Eng Remote Sens 81:345–54.

Horrocks CA, Dungait JAJ, Cardenas LM, Heal KV. 2014. Does extensification lead to enhanced provision of ecosystems services from soils in UK agriculture? Land Use Policy 38:123–8.

Jones CA, Cole CV, Sharpley AN, R WJ. 1984. A simplified soil and plant phosphorus model: I. Documentation. Soil Sci Soc Am J 48:1–3.

Kucharik CJ. 2003. Evaluation of a Process-Based Agro-Ecosystem Model (Agro-IBIS) across the U.S. Corn Belt: Simulations of the Interannual Variability in Maize Yield. Earth Interact 7:1–33.

Kucharik CJ, Barford CC, Maayar ME, Wofsy SC, Monson RK, Baldocchi DD. 2006. A multiyear evaluation of a Dynamic Global Vegetation Model at three AmeriFlux forest sites: Vegetation structure, phenology, soil temperature, and CO2 and H2O vapor exchange. Ecol Modell 196:1–31.

Kucharik CJ, Brye KR. 2003. Integrated BIosphere Simulator (IBIS) yield and nitrate loss predictions for Wisconsin maize receiving varied amounts of nitrogen fertilizer. J Environ Qual 32:247–68.

Kucharik CJ, Foley JA, Delire C, Fisher VA, Coe MT, Lenters JD, Young-Molling C, Ramankutty N, Norman JM, Gower ST. 2000. Testing the performance of a dynamic global ecosystem model: Water balance, carbon balance, and vegetation structure. Global Biogeochem Cycles 14:795–825.

Kucharik CJ, Twine TE. 2007. Residue, respiration, and residuals: Evaluation of a dynamic agroecosystem model using eddy flux measurements and biometric data. Agric For Meteorol 146:134–58.

Laboski CAM, Peters JB. 2012. Nutrient application guidelines for field, vegetable, and fruit crops in Wisconsin. University of Wisconsin Extension

Laflen JM, Flanagan DC, Engel BA. 2004. Soil erosion and sediment yield prediction accuracy using WEPP. JAWRA 40:289–97.

Lathrop RC. 1979. Dane County Water Quality Plan. Dane County Regional Planning Commission

Lathrop RC, Carpenter SR. 2013. Water quality implications from three decades of phosphorus loads and trophic dynamics in the Yahara chain of lakes. Inland Waters 4:1–14.

Lathrop RC, Carpenter SR, Stow CA, Soranno PA, Panuska JC. 1998. Phosphorus loading reductions needed to control blue-green algal blooms in Lake Mendota. Can J Fish Aquat Sci 55:1169–78.

McElroy AD, Chiu SY, Nebgen JW, Aleti A, Bennett FW. 1976. Loading functions for assessment of water pollution from nonpoint sources. Environ Prot Tech Serv, EPA 600/2-76-151.

NCDC. 2015. Global Historical Climate Network. National Climatic Data Center, National Oceanic and Atmospheric Administration http://www.ncdc.noaa.gov/data-access

NCEP. 2015. National Stage IV Quantitative Precipitation Estimate Mosaic, National Centers for Environmental Prediction. Center for Data Analytics, Office of Water Information, U.S. Geological Survey http://cida.usgs.gov/thredds/catalog.html?dataset=cida.usgs.gov/thredds/ncep_stageiv

Nearing MA, Govers G, Norton LD. 1999. Variability in soil erosion data from replicated plots. Soil Sci Soc Am J 63:1829–35.

Neitsch SL, Arnold JG, Kiniry JR, Williams Grassland JR, Labora WR. 2011. Soil and Water Assessment Tool Theoretical Documentation Version 2009.

Nennich TD, Harrison JH, VanWieringen LM, Meyer D, Heinrichs AJ, Weiss WP, St-Pierre NR, Kincaid RL, Davidson DL, Block E. 2005. Prediction of manure and nutrient excretion from dairy cattle. J Dairy Sci 88:3721–33.

NRCS. 2004. National Engineering Handbook - Hydrology Chapters: Chapter 10 Estimation of Direct Runoff from Storm Rainfall. USDA-NRCS

Oloya TO, Logan TJ. 1980. Phosphate desorption from soils and sediments with varying levels of extractable phosphate. J Environ Qual 9:526–31.

Ovington JD, Madgwick HAI. 1958. The sodium, potassium and phosphorus contents of tree species grown in close stands. New Phytol 57:273–84.

Panagos P, Borrelli P, Meusburger K, Alewell C, Lugato E, Montanarella L. 2015. Estimating the soil erosion cover-management factor at the European scale. Land Use Policy 48:38–50.

Parsen, MJ, Bradbury, KR, Hunt, RJ, and Feinstein, DT, 2016. The 2016 groundwater flow model for Dane County, Wisconsin: Wisconsin Geological and Natural History Survey Bulletin 110, 56 p.

Patil S, Sivapalan M, Hassan MA, Ye S, Harman CJ, Xu X. 2012. A network model for prediction and diagnosis of sediment dynamics at the watershed scale. J Geophys Res 117:F00A04.

Powell JM, McCrory DF, Jackson-Smith DB, Saam H. 2005. Manure Collection and Distribution on Wisconsin Dairy Farms. J Environ Qual 34:2036.

Powell JM, Rotz CA. 2015. Measures of nitrogen use efficiency and nitrogen loss from dairy production systems. J Environ Qual 44:336–44.

R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Richards LA. 1931. Capillary conduction of liquids through porous mediums. J Appl Phys 1:318–33.

Rotz CA, Montes F, Hafner SD, Heber AJ, Grant RH. 2014. Ammonia emission model for whole farm evaluation of dairy production systems. J Environ Qual 43:1143–58.

Ryser P, Verduyn B, Lambers H. 1997. Phosphorus allocation and utilization in three grass species with contrasting response to N and P supply. New Phytol 137:293–302.

Sadeghi SHR, Gholami L, Darvishan AK, Saeidi P. 2014. A review of the application of the MUSLE model worldwide. Hydrol Sci J 59:365–75.

Schneider A, Logan KE, Kucharik CJ. 2012. Impacts of Urbanization on Ecosystem Goods and Services in the U.S. Corn Belt. Ecosystems 15:519–41.

Sen S, Srivastava P, Vadas PA, Kalin L. 2012. Watershed-level comparison of predictability and sensitivity of two phosphorus models. J Environ Qual 41:1642–52.

Sharpley AN. 1985. The selection erosion of plant nutrients in runoff. Soil Sci Soc Am J 49:1527–34.

Sharpley AN. 2003. Soil mixing to decrease surface stratification of phosphorus in manured soils. J Environ Qual 32:1375–84.

Sharpley AN, Chapra SC, Wedepohl R, Sims JT, Daniel TC, Reddy KR. 1994. Managing agricultural phosphorus for protection of surface waters - issues and options. J Environ Qual 23:437–51.

Smith DR, Warnemuende-Pappas EA. 2015. Vertical tillage impacts on water quality derived from rainfall simulations. Soil Tillage Res 153:155–60.

Soylu ME, Kucharik CJ, Loheide, II, Steven P. 2014. Influence of groundwater on plant water use and productivity: Development of an integrated ecosystem – Variably saturated soil water flow model. Agric For Meteorol 189-190:198–210.

Stuntebeck T, Komiskey M, Peppler M, Owens D, Frame D. 2011. Precipitation-runoff relations and water-quality characteristics at edge-of-field stations, Discovery Farms and Pioneer Farm, Wisconsin, 2003–08.

USDA. 2008. Agricultural Waste Characteristics. In: Agricultural Waste Management Field Handbook, 210-VI-AWMFH, March 2008. Washington D.C.: Natural Resources Conservation Service, U.S. Department of Agriculture

USDA. 2013. Soil Survey Geographic (SSURGO) Database for Columbia, Dane, and Rock counties, Wisconsin. http://websoilsurvey.nrcs.usda.gov/. Last accessed 22/09/2011

USDA. 2014. United States Department of Agriculture, National Agricultural Statistics Service Cropland Data Layers. http://nassgeodata.gmu.edu/CropScape/

USDA-ERS. 2015. Livestock and poultry live and dressed weights. Economic Research Service, U.S. Department of Agriculture http://www.ers.usda.gov/data-products/livestock-meat-domestic-data.aspx#26070

USDA-NASS. 2013. USDA Census of Agriculture Historical Archive. Albert R. Mann Library at Cornell University and the National Agricultural Statistics Service, U.S. Department of Agriculture http://agcensus.mannlib.cornell.edu/

USDA-NASS. 2014. Census of Agriculture. National Agricultural Statistics Service, United States Department of Agriculture http://www.agcensus.usda.gov/

UVGSDC. 2004. Historical Census Browser. University of Virginia, Geospatial and Statistical Data Center http://mapserver.lib.virginia.edu/

UW-CALS. 2005. The Wisconsin Buffer Initiative. A report to the Natural Resources Board of the Wisconsin Department of Natural Resources by the University of Wisconsin-Madison, College of Agricultural and Life Sciences.

UW-Extension. 2014. Automated Weather Observation Network Data. University of Wisconsin-Extension Agricultural Weather http://agwx.soils.wisc.edu/uwex_agwx/awon

Vadas PA, Gburek WJ, Sharpley AN, Kleinman PJA, Moore PA Jr, Cabrera ML, Harmel RD. 2007a. A model for phosphorus transformation and runoff loss for surface-applied manures. J Environ Qual 36:324–32.

Vadas PA, Haggard BE, Gburek WJ. 2005. Predicting dissolved phosphorus in runoff from manured field plots. J Environ Qual 34:1347–53.

Vadas PA, Harmel RD, Kleinman PJA. 2007b. Transformations of soil and manure phosphorus after surface application of manure to field plots. Nutr Cycling Agroecosyst 77:83–99.

Vadas PA, Joern BC, Moore PA. 2012. Simulating soil phosphorus dynamics for a phosphorus loss quantification tool. J Environ Qual 41:1750–7.

Vadas PA, Kleinman PJA, Sharpley AN. 2004. A simple method to predict dissolved phosphorus in runoff from surface-applied manures. J Environ Qual 33:749–56.

Vadas PA, Krogstad T, Sharpley AN. 2006. Modeling Phosphorus Transfer between Labile and Nonlabile Soil Pools. Soil Sci Soc Am J 70:736.

Vadas PA, Mark Powell J, Brink GE, Busch DL, Good LW. 2015. Whole-farm phosphorus loss from grazing-based dairy farms. Agric Syst 140:40–7.

Vadas PA, White MJ. 2010. Validating Soil Phosphorus Routines in the SWAT Model. Transactions of the ASABE 53:1469–76.

Van Genuchten MT. 1980. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. Soil Sci Soc Am J 44:892–8.

Viney NR, Sivapalan M, Deeley D. 2000. A conceptual model of nutrient mobilisation and transport applicable at large catchment scales. J Hydrol 240:23–44.

Walker TW, Syers JK. 1976. Fate of phosphorus during pedogenesis. Geoderma 15:1–19.

WDATCP. 2015. DATCP Public FTP Site. Wisconsin Department of Agriculture, Trade, and Consumer Protection ftp://ftp.datcp.state.wi.us/GIS/DATCP_FOOD_SAFETY.gdb.zip

WDNR. 2011. Original Vegetation Cover of Wisconsin. Madison, WI: Wisconsin Department of Natural Resources ftp://dnrftp01.wi.gov/geodata/orig_veg_cover/orig_veg_cover.zip

WDNR. 2013. Wisconsin Land Cover database - Version 1. Wisconsin Department of Natural Resources ftp://gomapout.dnr.state.wi.us/landcover/

WDNR. 2015. Searchable database of CAFO WPDES permittees. Wisconsin Department of Natural Resources http://dnr.wi.gov/topic/AgBusiness/data/CAFO/

Williams J. 1995. The EPIC Model. In: Singh VP, editor. Computer Models of Watershed Hydrology. Littleton, CO: Water Resources Publications. pp 909–1000.

Williams JR. 1975. Sediment-yield prediction with Universal Equation using runoff energy factor. In: Present and Prospective Technology for Predicting Sediment Yield and Sources. Vol. ARS-S-40. U.S. Dept. Agric. pp 244–52.

Williams JR, Hann RW. 1978. Optimal operation of large agricultural watersheds with water quality restraints. Texas Water Resources Institute, Texas A&M Univ, Tech Rept No 96.

Wischmeier WH, Smith DD. 1978. Predicting rainfall erosion losses. A guide to conservation planning. U.S. Department of Agriculture, Washington, D.C.

Young RA, Onstad CA, Bosch DD, Anderson WP. 1989. AGNPS: A nonpoint-source pollution model for evaluating agricultural watersheds. J Soil Water Conserv 44:168–73.

Zipper SC, Soylu ME, Booth EG, Loheide, II, Steven P. 2015. Untangling the effects of shallow groundwater and soil texture as drivers of subfield-scale yield variability. Water Resour Res 51:6338–58.

Chapter 3

The synergistic effect of manure supply and extreme precipitation on surface water quality

Melissa Motew, Eric G. Booth, Stephen R. Carpenter, Xi Chen, Christopher J. Kucharik

In preparation for Environmental Research Letters

Abstract

Over-enrichment of phosphorus (P) in terrestrial agroecosystems contributes to eutrophication of surface waters. In the Midwest U.S. as well as elsewhere in the world, climate change is expected and has begun to bring an increase in the frequency of high-intensity precipitation events, which can serve as a primary conduit of P transport within watersheds. Using a process-based modeling framework developed for an agricultural watershed in southern Wisconsin, we conducted a 2x2 factorial experiment to test the effects of (high/low) terrestrial P supply (PSUP) and (high/low) precipitation intensity (PREC) on surface water quality. Sixty-year simulations were conducted for each of the four runs, with annual results obtained for watershed average P yield and concentration at the field scale (220m x 220m grid cells), P load and concentration at the stream scale, and summertime total P concentration (TP) in Lake Mendota. ANOVA results were generated for the 2x2 factorial design, with PSUP and PREC treated as categorical variables. The results showed a significant, positive interaction ($p<0.01$) between the two drivers for dissolved P concentration at the field and stream scales, as well as total P concentration at the field, stream, and lake scales. The synergy in dissolved P concentration was linked to nonlinear dependencies between P stored in manure and daily

rainfall and runoff amounts. The synergy between PSUP and PREC in affecting dissolved P losses from the landscape may have important ecological consequences because dissolved P is highly bioavailable. The synergy may be particularly problematic in areas having intense livestock production and surface applications of manure. Overall, the results suggest that high levels of terrestrial P supplied as manure can exacerbate water quality problems in the future as the frequency of high-intensity rainfall events increases with a changing climate. Conversely, lowering manure P supply may help improve the resilience of surface water quality to extreme events.

Keywords: extreme precipitation, phosphorus, interaction, synergy, watershed, runoff, manure

3.1. Introduction

Phosphorus (P) abundance and climate change both pose significant challenges to the ongoing management of freshwater resources. As the chief limiting nutrient in freshwater ecosystems, excess phosphorus (P) causes accelerated eutrophication of streams, rivers, and lakes. Because P is readily bound to, and slowly released from, soils and sediments, accumulated or "legacy P" can affect freshwater bodies for years or decades (Jarvie et al. 2013; Motew et al. 2017), causing ecological degradation and substantial economic loss (Kudela et al. 2015). The problem of legacy P affects many regions in the world, including North America, Europe, Asia, Latin America, and Oceania (MacDonald and Bennett 2009, Sattari et al. 2014, Garnier et al. 2015; Powers et al. 2016; Motew et al. 2017). As such, a growing body of literature has evolved in the past several years to address the challenges associated with legacy P and the threat it poses

to long-term freshwater quality (Sharpley et al. 2013, 2016; Haygarth et al. 2014, Rowe et al. 2015).

In addition to excessive terrestrial P, surface water quality is also vulnerable to extreme precipitation events, which in some watersheds cause a disproportionate amount of total annual loading of sediment and P to water bodies (Haygarth and Jarvis 1997; Royer et al. 2006; Gonzalez-Hidalgo 2013, Carpenter et al. 2014). Models and theories of climate change have predicted an increase in both mean annual precipitation as well as the frequency of extreme precipitation events in many regions (Trenberth 2011), and observational studies have begun to confirm this expectation (Fischer and Knutti 2016, Usinowicz et al. 2017). Globally, annual maximum daily precipitation has been increasing (Westra et al. 2013). Sub-daily events may also be increasing across regional and global scales, although data and methods to assess sub-daily trends are lacking (Westra et al. 2014). In the conterminous U.S. and Upper Mississippi River Basin, changes in precipitation have been dominant controls on runoff generation, overshadowing the effects of temperature, land cover change, and agricultural practices (Schilling et al. 2008; Frans et al. 2013; Gupta et al. 2015; McCabe and Wolock 2016).

An increase in extreme precipitation events and the subsequent increase in surface runoff generation suggests that managers will have to contend with even greater rates of P mobilization from landscapes in the future. This may be the case even within ecosystems assumed to have low runoff risk such as intensively farmed grasslands (Bilotta et al. 2007), cropland converted to grassland (Horrocks et al. 2014), and successional forests on abandoned cropland (Scott et al. 2001). Together, excessive terrestrial P supply and extreme precipitation may pose the risk of a "chemical time bomb", a phenomenon characterized by the long-term storage and sudden remobilization of a harmful environmental contaminant (Stigliani et al. 1991).

In addition to surface runoff generation, terrestrial P loss depends on available sources of P such as manure, fertilizer, or P-rich soils. Soil P changes through time according to the mass balance of inputs and outputs of P to the soil system. The largest fluxes to the soil system typically include agricultural inputs of P in the form of manure or fertilizer, and the largest fluxes out include harvested crops, biomass, or animal products. Fluxes of P lost to surface runoff, leached deep into the soil column, or deposited by the atmosphere, are generally much smaller than agricultural inputs and harvested outputs (Bennett et al. 1999). Relatively tiny losses of P to surface waters can however have a significant effect on aquatic ecosystems (Vadas et al. 2004; Carpenter and Lathrop 2008).

Phosphorus lost to surface runoff may be in particulate or dissolved form. Both forms contribute to eutrophication, and together as total P serve as the best indicator of water quality in receiving waters (Correll 1998). Transport of the two forms is controlled by different physical mechanisms. Particulate P is initially transported via erosion. In general, perennial land cover types, vegetative buffer strips, and reduced soil disturbance can be effective ways to reduce erosion and particulate P losses from agricultural lands (Baker and Laflen 1983; Sharpley and Smith 1994; McDowell and McGregor, 1984; Diebel et al. 2009; Uusi-Kämppä et al. 2000). Dissolved P loss on the other hand moves directly with surface runoff and therefore depends principally on the amount of dissolved P in manure or the surface soil layer. Surface applications of fertilizer and manure that are not physically incorporated can lead to stratification of P in soil and higher losses of dissolved P in runoff. Many studies have shown that compared to conventional tillage methods, reduced tillage approaches increase dissolved P losses to surface waters (Baker and Laflen 1983; McDowell and McGregor 1984; Langdale et al. 1985; Sharpley and Smith 1994; Bundy et al. 2001; Zhao et al. 2001; Tiessen et al. 2010).

While some studies have shown a link between extreme weather events and water quality responses, such as massive toxic algae blooms in Lake Erie, US (Michalak et al. 2013) and Lake Taihu, China (Zhu et al. 2014), evidence for causal connections may be lacking (Michalak 2016). In a modeling study of the effects of legacy P on lake water quality, Motew et al. (2017) found that average watershed soil P concentration compounded the effect of extreme precipitation on in-lake total P concentration, a finding with important implications for regions having both elevated terrestrial P supply and a rising trend in extreme precipitation. The mechanism responsible for that finding however was unclear. In this study, using the same models as Motew et al. (2017), we investigated how changes in extreme precipitation might interact with an abundant supply of terrestrial P to affect surface water quality. Using a 2x2 factorial design of high and low levels of each driver, we investigated how both precipitation intensity (PREC) and terrestrial P supply (PSUP) affected P transport at three spatial scales: the field scale, indicated by P loss in runoff from grid-cells of size 220 m by 220 m; the stream scale, indicated by in-stream P loading and concentration; and the lake scale, indicated by summertime total P (TP) concentration in the epilimnion. We focused on the watershed of Lake Mendota, a well-studied lake in southern Wisconsin, USA. This watershed has an excessive amount of P stored in soils and channel bed sediments (Bennett et al. 1999; Kara et al. 2011, Motew et al. 2017), and has also seen a significant increase in the occurrence of extreme precipitation events over the past two decades (Kucharik et al. 2010; Gillon et al. 2016) and elevated runoff (Usinowicz et al. 2017).

3.2. Materials and Methods

3.2.1. Study Region

Our study region was the 686 km^2 Lake Mendota watershed (LMW) (43.1097° N, -89.4206° W), a subwatershed of the Yahara Watershed of southern Wisconsin. The LMW is characterized by a predominance of agricultural land, relatively flat slopes (~4%), and silt loam soils. High P loads to Lake Mendota have been primarily attributed to the large number of dairy operations located upstream of the lake within its drainage area (Lathrop 2007). Empirical P loading patterns to Lake Mendota are strongly influenced by heavy precipitation events, with 29 days per year delivering approximately 74% of the annual load (Carpenter et al. 2014). Over the 1930-2010 time period, annual precipitation has increased in the Yahara Watershed at a rate of 2.1 mm y^{-1}, and the frequency of large storm events (>50mm) has increased from 9.5 events per decade (1931-1990) to 18 events per decade (1991-2010) (Gillon et al. 2016). These trends are consistent with other regional studies (Kunkel et al. 2007, Qian et al. 2007, Peterson et al. 2008, Pryor et al. 2009, Kucharik et al. 2010, Baker et al. 2012, Villarini et al. 2013).

3.2.2. Overview of Models

We used a deterministic, process-based watershed modeling framework to simulate field-to-lake flows of water, sediment, and phosphorus, as well as lake water quality in the Yahara Watershed. The framework consisted of Agro-IBIS, a terrestrial ecosystem model that simulates carbon, water, energy, momentum, nitrogen and phosphorus cycles in the soil-vegetation-atmosphere system (Kucharik et al. 2000; Kucharik et al. 2003); THMB (Terrestrial Hydrology and Biogeochemistry Model), a hydrologic and nutrient routing model (Coe 1998, 2000; Donner

et al. 2002); and the Yahara Water Quality Model (YWQM) that estimates lake water quality conditions in the Yahara lake chain (Carpenter and Lathrop 2014). Agro-IBIS and THMB have been used together previously to study hydrology and nitrogen transport in the Mississippi Basin (Donner and Kucharik 2003; Donner et al. 2004; Donner and Kucharik 2008). Recently, phosphorus cycling and transport were added to Agro-IBIS and THMB, and both models were linked with the YWQM for use in the Yahara Watershed (Motew et al. 2017). The three-model suite was previously calibrated and validated using local observations of streamflow, sediment and nutrient loading to the Yahara lakes, as well as local and regional observations of soil test P concentration and runoff P (Motew et al. 2017).

3.2.3. Representation of P yield and in-stream transport

Because we used models to investigate a possible interaction between terrestrial P supply and precipitation intensity, any interaction between these drivers would necessarily arise from the model equations governing P loss and transport. At the grid-cell level, which we equate with the field scale, P loss to runoff is calculated in Agro-IBIS for both dissolved and particulate forms. The equations governing the loss of these two forms are functionally different. For dissolved P, the equations in Agro-IBIS come from the SurPhos model which estimates daily dissolved P loss from manure, fertilizer, and soil (Vadas et al. 2007). The empirically-derived equations of SurPhos were originally based on observations of P desorption from soil (Sharpley et al. 1981), and later modified to capture P loss from dairy, swine, and poultry manures at a daily timescale (Kleinman et al. 2002a, Vadas et al. 2004). For manure, the main dependencies in the calculation of dissolved P loss (DP_{conc}, mg L^{-1}) include daily runoff (RUN, mm), daily rainfall ($RAIN$, mm), the amount of water extractable P leached from manure (LCH_m, kg), and

grid-cell area (*AREA*, ha) (Vadas et al. 2007, Appendix K.1). These quantities are related in the following manner:

$$DP_{man} = \left(\frac{10\ LCH_m}{(RAIN)(AREA)}\right)\left(\frac{RUN}{RAIN}\right)^{0.225} \quad [1]$$

For P leached from fertilizer to runoff, the same calculation is made except water extractable P leached from manure is replaced by water extractable P leached from fertilizer. The concentration of dissolved inorganic P in runoff from soil is calculated using a linear extraction coefficient (0.005) that is multiplied with the concentration of P in the surface soil labile pool (Appendix K.1.).

Daily losses of sediment P in runoff are calculated using the total amount of P in the surface 2 cm of soil, daily sediment yield, and a P enrichment value (McElroy et al. 1976; Williams and Hann 1978; Sharpley 1985) (Appendix K.2.). Sediment yield is calculated using the Modified Universal Soil Loss Equation (MUSLE) (Williams 1975). In-stream transport of P is modeled by THMB, which represents dissolved P as a conservative solute that follows water flow, and sediment P using an in-channel erosion process based on the model of Viney et al. (2000) (Appendix L). In addition to Appendices K and L, further details regarding the terrestrial P module in Agro-IBIS as well as the P transport module in THMB can be found in the appendix of Motew et al. (2017), available online at www.github.com/mmotew/Appendix-Motew-et-al-2017.git.

3.2.4. Experimental Design

The model suite consisting of Agro-IBIS, THMB, and the YWQM was used to conduct a 2x2 factorial experiment to test the interaction effect of (low/high) terrestrial P supply and (low/high) precipitation intensity on surface water quality. Sixty-year simulations were conducted for each of the four runs, named HIHI, HILO, LOHI, and LOLO, with the first two characters signifying the P supply regime and the last two characters the precipitation intensity regime. Annual results were obtained for P concentration and P yield at the field scale (220m by 220m grid-cell resolution), P concentration and loads in the inlet river to Lake Mendota, and TP in Lake Mendota. A sixty year simulation period was chosen in order to obtain an adequate sample size (n = 60), to minimize computational run-time, and to generate many random combinations of storm event timing in relation to nutrient applications. At the field and stream scales, total P concentration, yield, and load were each broken down into their dissolved and sediment forms since the processes controlling the release and transport of each form were different. For each simulation, a land cover rotation for the years 2006 to 2009 was repeated 15 times. Land cover and nutrient application rates were previously obtained for the 1986-2013 period (Motew et al. 2017). For the two factorial runs having high soil P levels (HIHI and HILO), Agro-IBIS was spun up for 25 y prior to the start of the simulation to bring the surface soil layer P concentration to equilibrium at ~200 ppm (Bray-1 soil test P equivalent). Soil P was maintained at this level, approximately in equilibrium, using a scale factor of 1.3 applied to historical nutrient application rates. For the low soil P runs (LOHI and LOLO), the same procedure was used except the spin up was conducted using a scale factor of 0.6, yielding an equilibrium soil P concentration of ~65 ppm. These values were chosen to represent highly contrasting yet plausible levels of terrestrial P supply (Supp Fig. 1).

Two 60-year climate scenarios consisting of daily precipitation, maximum and minimum air temperature, wind speed, relative humidity, and incoming shortwave radiation were generated with the following characteristics: (1) very similar mean annual values for each variable reflecting the current climate (1996-2015) of Madison, Wisconsin (Menne et al. 2012); and (2) contrasting daily precipitation distributions where one had more intense rainfall events (the HI regime) compared to the other (the LO regime). The contrasting precipitation regimes therefore varied in average length between storm events yet summed to similar annual values. Simulated runoff for the two regimes showed a greater event-based runoff to rainfall ratio in the HI case (Fig. 1). A greater runoff to rainfall ratio was a plausible biophysical response since rainfall rates would be more likely to exceed infiltration rates as storm intensity increased.

We began by creating the two daily precipitation time series using the rectangular Poisson pulse model (Eagleson, 1978). In this model, precipitation is simulated as a series of discrete random storms where storm intensity, storm duration, and inter-storm period are each expressed by an exponential distribution:

$$f_y(y) = \frac{1}{\bar{y}} exp\left(-\frac{y}{\bar{y}}\right) \qquad [2]$$

where y is an exponentially distributed random variable with mean $\bar{y}$. We used [2] to simulate storm intensity (i), storm duration (t_r), and inter-storm period (t_b). Storm depth (h) is the product of storm intensity and duration, $h=i*t_r$. The simulated rainfall time-series was created by randomly drawing from these three distributions to determine i, t_r and t_b for the subsequent precipitation event, under the assumption that i, t_r and t_b were independent. To represent seasonality of precipitation, the parameter of each function was varied on a monthly basis using

data from Hawk and Eagleson (1992) for La Crosse, Wisconsin. To generate two time-series with highly diverging precipitation regimes, we adjusted the initial monthly parameters for mean storm intensity and inter-storm period using a multiplier (Table 1). The HI regime had multipliers greater than 1 indicating a higher mean storm intensity and longer inter-storm period, while the LO regime had multipliers less than 1. The differing parameter multipliers were adjusted to meet the requirement that mean annual precipitation be nearly identical in both regimes (Table 1). The LO regime was similar to the current climate in terms of the 99th percentile daily precipitation rate and the HI regime was similar to the more extreme CMIP3 climate projections for the region (Notaro et al. 2014) (Table 1). Once generated, the hourly storm events were tabulated into daily precipitation amounts to be input to Agro-IBIS. To support model performance of streamflow in THMB, as determined through prior calibration (Motew et al. 2017), daily precipitation amounts in Agro-IBIS were distributed evenly over an 8 hr storm period. While this may have caused a loss of detail in daily storm characteristics, the imposed 8 hr storm length was applied to both precipitation regimes and was therefore not expected to interfere with the underlying differences between HI and LO.

Maximum and minimum daily air temperature (T_{max} and T_{min}) were generated using a method similar to that in the WGEN stochastic weather simulation model (Richardson and Wright 1984). Mean and standard deviation of T_{max} and T_{min} were determined for each day of the year for the Madison climate record (1996-2015), conditioned on wet and dry days. A Fourier model was then fit for all 8 statistics. A daily residual was calculated using a weakly stationary generating process that preserves serial- and cross-correlation and adds a stochastic component. Daily T_{max} and T_{min} were then calculated based on modeled mean and standard deviation and

whether the day is wet or dry. The remaining climate variables were estimated using regression models with a stochastic component presented in Booth et al. (2016).

3.2.5. Analysis

Annual average watershed values of P quantities, including yield, concentration, and load, were examined for an interaction between PSUP and PREC. We used 2-way ANOVA for a 2x2 factorial design to test the interaction effect of these drivers, treating both as categorical variables. An interaction between PREC and PSUP was statistically significant if the p-value for the interaction term was less than 0.05. We used QQ plots and the Shapiro-Wilk test to check for normality of each P quantity prior to performing ANOVA. If the distribution was not approximately normal, we used log10 transformation and re-tested for normality. Based on this procedure, all P quantities warranted a log10 transformation. However, we also performed ANOVA on the untransformed distributions since in some cases ANOVA revealed a significant interaction in the untransformed data and no significant interaction in the transformed data, indicating a removable interaction (de González and Cox 2007). Identifying removable interactions was helpful in the overall interpretation of results.

For those P quantities that had a significant interaction between PREC and PSUP, we generated interaction plots to visually examine the nature of the interaction, i.e. if they were synergistic, inversely related, etc. We also generated plots of probability density for those quantities using the ‘ksdensity’ kernel smoothing function in MATLAB. For analysis at the field scale, annual P yield was calculated as the sum of daily yield, averaged over all terrestrial grid-cells. For annual P concentration at the field scale, annual yield was divided by total annual runoff volume. A similar approach was taken at the stream scale, where daily direct drainage

loads to Lake Mendota were summed over the year to obtain annual load, and annual concentration was obtained by dividing annual load by total streamflow volume. All annual P variables were calculated for the water year, defined as Nov 1 – Oct 31. All statistical analyses were performed using the Statistics Toolbox in MATLAB (The MathWorks, Inc. 2015).

3.3. Results

At the field and stream scale, we detected significant interactions between precipitation intensity and terrestrial P supply for both dissolved and total P concentration. An interaction for total P concentration was also detected at the lake scale, signifying that the interaction was persistent in both dissolved and total P at all scales of the watershed (Table 2). The interaction applied only to dissolved P loss from manure, i.e. not from fertilizer or soil, and did not apply to sediment P loss or to mass-based quantities like P yield and load.

Aside from the significant interactions detected for dissolved P concentration at field and stream scales, as well as total P concentration at all three scales, no other P quantities had significant interactions (Table 2). Dissolved P yield (field scale) and load (stream scale) had removable interactions, as did total P yield and load. Sediment P yield at the field scale and sediment P load and concentration at the stream scale had no interaction whether transformed or not. Sediment P concentration at the field scale had a removable interaction. Interaction plots showed that for all quantities having a significant interaction between PREC and PSUP, the interaction was synergistic (Fig. 2). This means that when PSUP was HI, the effect of PREC was greater than it was when PSUP was LO.

The probability distributions of annual P quantities having a significant interaction showed that for the HIHI run, the distribution was spread wider and was centered at the highest

value of all four runs. This indicated a higher probability of very high annual P concentration at each scale, and relatively low probability of low P concentration at each scale (Fig. 3). Likewise for the LOLO run, there was a high probability of low P concentration events and conversely a low probability of very high P concentration events. The HILO and LOHI runs had similar mean values across the P concentration variables, but at the grid-cell level LOHI had a narrower distribution than HILO. This may suggest that PSUP played a bigger role in determining the spread of the distribution at the grid-cell level (since HILO was spread wider than LOHI). At the stream level, HILO and LOHI had very similar distributions, suggesting that the effects of PREC and PSUP effectively countered one another at that scale. At the lake scale, both of the HI PSUP runs caused very long, high-valued tails compared with the LO PSUP runs. HILO and LOHI had similar distribution shapes at the lake scale, but LOLO was highly skewed toward low values of TP.

The proportions of mean annual total P concentration in dissolved and sediment forms varied across model runs, but were similar between the field and stream scales (Fig. 4). Dissolved P constituted more than 60% of annual total P concentration at the grid cell scale for the runs with PREC=HI. For the runs with PREC=LO, the proportions were roughly half and half, with the LOLO run having the lowest proportion of dissolved P (49%) and the HILO run having the second-lowest (54%). At the stream scale, the HIHI run had the highest proportion of dissolved P (67%), while LOLO had the lowest (41%). Across the four runs, the proportion of sediment P increased. This may have been the result of lower amounts of fertilizer and manure being applied, which are both rich in dissolved P. It may also reflect the importance of heavy rainfall events in the process of dissolved P leaching, as evident in the synergistic interaction for dissolved P concentration at the field and stream scales. In the absence of both high terrestrial P

supply and heavy rainfall, mean annual dissolved P loss was at its lowest amount, which was reflected in its low proportion relative to sediment P.

The relative proportions of mean annual dissolved P from manure, fertilizer, and soil are shown in Fig. 5 for cropland cells. Soil contributions were the largest overall, followed by manure and then fertilizer. We found the interaction between PSUP and PREC to be significant for manure but not for fertilizer or soil. This suggested that dissolved P loss from manure to runoff was the causal mechanism driving the synergy within the watershed. This can be understood in terms of the nonlinear relationships governing dissolved P loss from manure to runoff (Eq. 1, Appendix K.1). Despite being governed by similar equations, losses from fertilizer did not display an interaction. This is likely due to the fact that liquid fertilizer leaches quickly into the soil, and is not present long enough at the ground surface to be exposed to hydrologic fluxes. Losses from soil are governed by a simple linear extraction coefficient, which likely explains why they displayed no interaction between PSUP and PREC.

Mathematically, the synergy could be explained in terms of the direct dependence of dissolved P on both the runoff to rainfall ratio and the supply of water-extractable P in manure (Eq. 1 and Appendix K.1). The runoff to rainfall ratio was significantly different between the HI and LO precipitation regimes (Fig. 1), and the 0.225 in the ratio's exponent (Eq. 1) implied high sensitivity to the ratio. Sediment P on the other hand did not display a high sensitivity to rainfall intensity, despite its high dependence on surface runoff (Appendix K.2.). Dissolved P yield and load did not display a significant interaction either, whereas concentration did. This may be explained by the fact that the calculation of P yield involves multiplying the dissolved P concentration by runoff volume. According to Eq. 1, dissolved P concentration is inversely related to daily rainfall, but proportional to runoff. Multiplying by runoff volume would be

expected to increase the variance which might potentially obscure an interaction. This idea was supported by the R^2 values and ANOVA results for dissolved P concentration and yield at the field scale, which showed that R^2 was greater for concentration (Table 2), and the MS error and SS error terms were lower (not shown). Specifically, R^2 was 0.49 and 0.33, MS error was 0.001 and 0.01, and SS error was 0.31 and 3.4, for concentration and yield, respectively.

3.4. Discussion

Our study highlights the potential danger posed by climate change whereby extreme precipitation will exacerbate losses of dissolved P from manure, an important source of P in many watersheds (Hansen et al. 2002). The synergy between P supply and precipitation intensity was linked to the nonlinear relationships governing dissolved P loss from manure (Eq. 1 and Appendix K.1). These equations involve both the supply of dissolved P in manure as well as daily values of runoff and rainfall. The presence of runoff in the numerator and rainfall in the denominator of Eq. 1 indicates an important role for storm event intensity, whereby a greater runoff to rainfall ratio will increase P loss from manure, despite the diluting effects of rainfall (Vadas et al. 2005).

The synergy in dissolved P loss from manure was strong enough to persist in total P loss at the field scale and beyond to the stream and lake scales. This finding underscores the importance of manure and its unique interactions with the hydrologic cycle in driving water quality outcomes throughout watersheds. These dynamics were first observed in laboratory and field studies (Sharpley and Moyer 2000; Kleinman et al. 2002a, 2002b, 2004; Kleinman and Sharpley 2003; Sharpley et al. 2004; Vadas et al. 2007b), and have since been coded into the SurPhos model, which predicts dissolved P loss from fields (Vadas et al. 2004, 2005, 2007a).

The SurPhos equations were designed to be modular and easily integrated into daily time-step models. As such, their implementation into other watershed models is ongoing. Collick et al. (2016) implemented the SurPhos equations into the Soil and Water Assessment Tool (SWAT2012 version 586) and found an improvement in the sensitivity of dissolved P loss to key factors in nutrient management, including timing, rate, method, and form of application. The important role of manure interactions with water fluxes in our study further emphasizes the benefit of including the SurPhos equations into watershed models to more accurately capture the effects of manure P applications in conjunction with changes in hydrology. In watersheds such as the Yahara, where dissolved P loss from manure constitutes a relatively large proportion of total P delivered to surface waters (e.g. Galeone 1999), proper characterization of manure dynamics, transformations, and interactions with the hydrologic cycle is essential.

In a study of Lake Erie, Michalak et al. (2013) found that long-term trends in agricultural management practices and extreme precipitation during spring contributed to an historic algae bloom in 2011. Using the SWAT model, they found that dissolved reactive P (DRP) yields, which were a likely factor driving the blooms, were sensitive to precipitation intensity, fertilizer application timing, and tillage practices, with precipitation having the strongest influence and fertilizer the weakest. To our knowledge, the version of SWAT used did not have the SurPhos equations, which include explicit representation of surface manure and fertilizer pools. As simulated by SurPhos, these pools have their own transformations and unique interactions with surface hydrologic fluxes that are independent of the soil pools. In the version of SWAT used by Michalak et al., fertilizer and manure P are added directly into the surface soil pools. Applications of manure or fertilizer P would still be expected to increase labile P and DRP loss, but the effect may be underestimated compared to estimates that use explicit representation of

surface fertilizer and manure pools. It is therefore possible that DRP losses during extreme rainfall events may have been underestimated in the Lake Erie study, especially if the events had occurred shortly after application. We suggest that the watershed modeling community adopt the methodology employed by SurPhos in order to best characterize the effects of surface applications of P and the timing of applications with respect to precipitation events.

For the 2011 Lake Erie bloom, increased DRP loads and meteorological conditions including extreme rainfall events in the spring followed by warm temperatures and low wind speed in the summer, created optimal conditions for cyanobacteria growth. The following several years also saw large blooms in Lake Erie, including a bloom in 2014 that necessitated the temporary suspension of the municipal water supply to Toledo, OH, and another record-setting bloom occurring in 2015 (Michalak 2016). Even though eutrophication can be sustained at low DRP concentrations (Hudson et al. 2000), high DRP concentrations are favorable to the growth of cyanobacteria such as *Microcystis* (Elser 1999). Thus, management practices that increase DRP loss may have important ecological consequences. Our results showed that extreme precipitation can exacerbate losses of dissolved P from manure. This in turn may help foster the conditions in lakes whereby toxic algae thrive.

The synergy observed in the present study was attributed to the supply of manure resting on the soil surface and its direct contact with overland flow. For all model simulations, all cropland grid cells receiving manure had 55% of the total annual application applied to the soil surface. These results suggest that reducing the amount of manure present on the soil surface may be an effective means for reducing the synergistic interaction with extreme precipitation. For regions such as the Lake Mendota Watershed in which livestock manure is a major source of P to receiving waters (Lathrop 2007), decreasing the amount of manure applied to fields, or

reducing surface application, is a challenging prospect. Incorporation of manure via tillage, for example, threatens to increase soil erosion and losses of sediment P (Baker and Laflen 1983). Manure injection, while still technologically evolving, can be costly in terms of both money and energy (Chen 2002).

High levels of soil test P observed in Wisconsin already reflect a dilemma of excess manure production and the problem of its disposal (Stuntebeck et al. 2011). And, despite significant water quality and other environmental hazards associated with high-intensity livestock production, economics in many regions still favor farms with higher animal numbers. In the Mendota Watershed, where agricultural lands are being encroached on by urban development, manure spreading must occur on a shrinking land area, which may result in very high rates of manure application per unit area (Gillon et al. 2016). With growing adoption of conservation tillage (Betz and Genskow 2012), applications are likely to be surface applied. Given the prospect of climate change, these trends may pose a substantial risk to water quality remediation efforts. On the other hand, if surface applications can be reduced or eliminated, there may be protective benefits to counter extreme precipitation. Further research should evaluate the effectiveness of manure injection or other incorporation technologies in mitigating the synergy.

Sediment P throughout the watershed failed to exhibit a synergy between PSUP and PREC. Despite the strong dependence of sediment yield and sediment P loss on daily runoff at the grid cell scale (Eq. 7), there was no interacting effect between P supply and hydrology on sediment P at any scale of analysis. This may be because P bound to soil particles comes mostly from non-labile (stable) pools, which make up the majority of total P in the soil system. The non-labile pools are only indirectly affected by water fluxes, whereas labile P in manure is directly affected by runoff and rainfall events. In most watersheds, sediment P losses are greater than

dissolved P losses (Sharpley et al. 1994; Owens and Walling 2002). Yet in regions where manure applications are prevalent, dissolved P loading can be comparable to sediment P loading (Lathrop 2007). High losses of dissolved P from our study region as well as other livestock-intensive regions may be attributable to the low N:P content of manure (4:1-5:1), and the fact that many farmers apply manure at rates to meet crop nitrogen demand (6:1-8:1) (Silveira et al. 2010). Applying manure to meet N demand can easily lead to high accumulation of P in soils, and when surface applied, stratification among soil layers (Andraski and Bundy 2003; Smith and Warnemuende-Pappas 2015). At both the grid cell and stream scales, we found that the proportion of total P in the sediment form increased between the HIHI and LOLO runs, indicating that under non-extreme conditions, dissolved P loss was not as prevalent. This effect may be attributable to the synergy itself, since dissolved P loss was enhanced when both PSUP and PREC were high. This implies that with climate change, and in the presence of manure, the proportion of total P loads in the dissolved form may increase.

For the runs where PSUP was HI, there were long, high-valued tails for log TP in Lake Mendota (Fig. 3). This suggests that PSUP played a stronger role at the lake scale than at the stream and grid-cell scales. Because of the roles of sedimentation and outflow from Lake Mendota, which are captured empirically in the Yahara WQ Model, the effect of PREC may have been dampened, leaving PSUP the dominant driver of log TP in the lake. The important role of PSUP in affecting lake water quality is consistent with the findings of Motew et al. (2017), who showed that terrestrial P supply has a strong influence on lake water quality.

3.5. Conclusion

In this simulation experiment, the supply of P in manure and extreme precipitation interacted synergistically to affect dissolved P concentration in the Lake Mendota Watershed. In the simplest terms, the results suggest that the effects of extreme precipitation on water quality are exacerbated in the presence of manure, and, vice-versa, that the effects of manure on water quality are exacerbated by extreme precipitation. Despite a mechanistic link with dissolved P losses only, the interaction was statistically significant for total P at the field scale, stream scale, and lake scale. This implies that manure management choices can have important implications for water quality at all spatial scales of a watershed. There is already a need for improved nutrient management in watersheds such as Lake Mendota to curtail the effects of excessive terrestrial P supply. Our results further underscore the importance of improved *manure* management to mitigate the compounding effects of extreme precipitation. Practices that reduce the amount of manure P applied, and/or avoid unincorporated surface application may prove the most useful.

Given the difficulty in measuring and attributing the effects of changing climate, land use, and land management on water quality (Gillon et all. 2016), process-based, spatially explicit models provide a way to isolate, identify, and scale-up important mechanisms affecting water quality (DeAngelis and Yurek 2017); in this case the specific interaction between terrestrial nutrient supply and the hydrologic cycle. Ecosystem models are however dependent on observations, such as the field studies on manure P leaching that formed the basis for the P cycling routines in SurPhos and Agro-IBIS. Together, field observations and modeling represent a powerful and complementary approach for furthering watershed and ecosystem science (Turner and Carpenter 2017, Seidl 2017). As such, improvements in both will be needed to better understand the challenges facing freshwater resources in the future.

Acknowledgements

We thank Peter Vadas for making helpful suggestions to guide our analysis and interpretation.

This study was supported by the National Science Foundation under Grant Nos. DEB-1440297 and DEB-1038759.

References

Andraski, T.W., L.G. Bundy, and K.C. Kilian. 2003. Manure history and long-term tillage effects on soil properties and phosphorus losses in runoff. Journal of Environmental Quality 32(5): 1782–1789.

Baker, J.M., T.J. Griffis, and T.E. Ochsner. 2012. Coupling landscape water storage and supplemental irrigation to increase productivity and improve environmental stewardship in the U.S. Midwest. Water Resour. Res. 48(5): W05301.

Baker, J.L., and J.M. Laflen. 1983. Water quality consequences of conservation tillage: New technology is needed to improve the water quality advantages of conservation tillage. J. Soil Water Conserv. 38(3): 186–193.

Bennett, E.M., T. Reed-Andersen, J.N. Houser, J.R. Gabriel, and S.R. Carpenter. 1999. A Phosphorus Budget for the Lake Mendota Watershed. Ecosystems 2(1): 69–75.

Betz, C., and K. Genskow. 2012. Farm Practices in the Lake Mendota Watershed: A Comparative Analysis of 1996 and 2011. Available at http://yaharaportal.com/sites/default/files/MendotaFPIReporFINALJune2012.pdf.

Bilotta, G.S., R.E. Brazier, and P.M. Haygarth. 2007. Processes affecting transfer of sediment and colloids, with associated phosphorus, from intensively farmed grasslands: erosion. Hydrol. Process. 21(1): 135–139.

Booth, E.G., J.X. Qiu, S.R. Carpenter, J. Schatz, X. Chen, C.J. Kucharik, S.P. Loheide, M.M. Motew, J.M. Seifert, and M.G. Turner. 2016. From qualitative to quantitative environmental scenarios: Translating storylines into biophysical modeling inputs at the watershed scale. Environmental Modelling & Software 85: 80–97.

Bundy, L.G., T.W. Andraski, and J.M. Powell. 2001. Management practice effects on phosphorus losses in runoff in corn production systems. J. Environ. Qual. 30(5): 1822–1828.

Carpenter, S.R., E.G. Booth, C.J. Kucharik, and R.C. Lathrop. 2014. Extreme daily loads: role in annual phosphorus input to a north temperate lake. Aquat. Sci. 77(1): 71–79.

Carpenter, S.R., and R.C. Lathrop. 2008. Probabilistic estimate of a threshold for eutrophication. Ecosystems 11(4): 601–613.

Carpenter, S.R., and R.C. Lathrop. 2014. Phosphorus loading, transport and concentrations in a lake chain: a probabilistic model to compare management options. Aquat. Sci. 76(1): 145–154.

Chen, Y. 2002. A liquid manure injection tool adapted to different soil conditions. Trans. ASAE 45(6): 1729–1736.

Coe, M.T. 1998. A linked global model of terrestrial hydrologic processes: Simulation of modern rivers, lakes, and wetlands. J. Geophys. Res. 103(D8): 8885–8899.

Coe, M.T. 2000. Modeling Terrestrial Hydrological Systems at the Continental Scale: Testing the Accuracy of an Atmospheric GCM. J. Clim. 13(4): 686–704.

Collick, A.S., T.L. Veith, D.R. Fuka, P.J.A. Kleinman, A.R. Buda, J.L. Weld, R.B. Bryant, P.A. Vadas, M.J. White, R.D. Harmel, and Z.M. Easton. 2016. Improved Simulation of Edaphic and Manure Phosphorus Loss in SWAT. J. Environ. Qual. 45(4): 1215–1225.

Correll, D.L. 1998. The role of phosphorus in the eutrophication of receiving waters: A review. J. Environ. Qual. 27(2): 261–266.

DeAngelis, D.L., and S. Yurek. 2017. Spatially Explicit Modeling in Ecology: A Review. Ecosystems 20(2): 284–300.

Diebel, M.W., J.T. Maxted, D.M. Robertson, S. Han, and M.J. Vander Zanden. 2009. Landscape planning for agricultural nonpoint source pollution reduction III: assessing phosphorus and sediment reduction potential. Environ. Manage. 43(1): 69–83.

Donner, S.D., M.T. Coe, J.D. Lenters, T.E. Twine, and J.A. Foley. 2002. Modeling the impact of hydrological changes on nitrate transport in the Mississippi River Basin from 1955 to 1994. Global Biogeochem Cycles 16(3): 1–19.

Donner, S.D., and C.J. Kucharik. 2003. Evaluating the impacts of land management and climate variability on crop production and nitrate export across the Upper Mississippi Basin. Global Biogeochem. Cycles 17(3): 1–16.

Donner, S.D., and C.J. Kucharik. 2008. Corn-based ethanol production compromises goal of reducing nitrogen export by the Mississippi River. Proc. Natl. Acad. Sci. U. S. A. 105(11): 4513–4518.

Donner, S.D., C.J. Kucharik, and J.A. Foley. 2004. Impact of changing land use practices on nitrate export by the Mississippi River. Global Biogeochem. Cycles 18(1): 1–21.

Eagleson, P.S. 1978. Climate, Soil, and Vegetation .2. Distribution of Annual Precipitation Derived from Observed Storm Sequences. Water Resour. Res. 14(5): 713–721.

Elser, J.J. 1999. The pathway to noxious cyanobacteria blooms in lakes: the food web as the final turn. Freshw. Biol. 42(3): 537–543.

Fischer, E.M., and R. Knutti. 2016. Observed heavy precipitation increase confirms theory and early models. Nat. Clim. Chang. 6(11): 986–991.

Frans, C., E. Istanbulluoglu, V. Mishra, F. Munoz-Arriola, and D.P. Lettenmaier. 2013. Are climatic or land cover changes the dominant cause of runoff trends in the Upper Mississippi River Basin? Geophys. Res. Lett. 40(6): 1104–1110.

Galeone, D.G. 1999. Calibration of paired basins prior to streambank fencing of pasture land. J. Environ. Qual. 28(6): 1853–1863.

Garnier, J., L. Lassaletta, G. Billen, E. Romero, B. Grizzetti, J. Némery, T.P.Q. Le, C. Pistocchi, N. Aissa-Grouz, T.N.M. Luu, L. Vilmin, and J.-M. Dorioz. 2015. Phosphorus budget in the water-agro-food system at nested scales in two contrasted regions of the world. Global Biogeochem. Cycles 29(9): 1348–1368.

Gillon, S., E.G. Booth, and A.R. Rissman. 2016. Shifting drivers and static baselines in environmental governance: challenges for improving and proving water quality outcomes. Regional Environ. Change 16(3): 759–775.

Gonzalez-Hidalgo, J.C., R.J. Batalla, and A. Cerda. 2013/3. Catchment size and contribution of the largest daily events to suspended sediment load on a continental scale. Catena 102: 40–45.

Gupta, S.C., A.C. Kessler, M.K. Brown, and F. Zvomuya. 2015. Climate and agricultural land use change impacts on streamflow in the upper midwestern United States. Water Resour. Res. 51(7): 5301–5317.

Hansen, N.C., T.C. Daniel, A.N. Sharpley, and J.L. Lemunyon. 2002. The fate and transport of phosphorus in agricultural systems. J. Soil Water Conserv. 57(6): 408–417.

Hawk, K.L., and P.S. Eagleson. 1992. Climatology of station storm rainfall in the continental United States: Parameters of the Bartlett-Lewis and Poisson rectangular pulse models, Department of Civil Engineering, R.M. Parsons Laboratory Report 336. Massachusetts Institute of Technology, Cambridge, MA.

Haygarth, P.M., H.P. Jarvie, S.M. Powers, A.N. Sharpley, J.J. Elser, J. Shen, H.M. Peterson, N.-I. Chan, N.J.K. Howden, T. Burt, F. Worrall, F. Zhang, and X. Liu. 2014. Sustainable phosphorus management and the need for a long-term perspective: the legacy hypothesis. Environ. Sci. Technol. 48(15): 8417–8419.

Haygarth, P.M., and S.C. Jarvis. 1997/1. Soil derived phosphorus in surface runoff from grazed grassland lysimeters. Water Res. 31(1): 140–148.

Horrocks, C.A., J.A.J. Dungait, L.M. Cardenas, and K.V. Heal. 2014. Does extensification lead to enhanced provision of ecosystems services from soils in UK agriculture? Land Use Policy 38: 123–128.

Hudson, J.J., W.D. Taylor, and D.W. Schindler. 2000. Phosphate concentrations in lakes. Nature 406(6791): 54–56.

Jarvie, H.P., A.N. Sharpley, B. Spears, A.R. Buda, L. May, and P.J.A. Kleinman. 2013. Water Quality Remediation Faces Unprecedented Challenges from "Legacy Phosphorus." Environ. Sci. Technol. 47(16): 8997–8998.

Kara, E.L., C. Heimerl, T. Killpack, M.C. Van de Bogert, H. Yoshida, and S.R. Carpenter. 2011. Assessing a decade of phosphorus management in the Lake Mendota, Wisconsin watershed and scenarios for enhanced phosphorus management. Aquat. Sci. 74(2): 241–253.

Kleinman, P.J.A., and A.N. Sharpley. 2003. Effect of broadcast manure on runoff phosphorus concentrations over successive rainfall events. J. Environ. Qual. 32(3): 1072–1081.

Kleinman, P.J.A., A.N. Sharpley, B.G. Moyer, and G.F. Elwinger. 2002a. Effect of mineral and manure phosphorus sources on runoff phosphorus. J. Environ. Qual. 31(6): 2026–2033.

Kleinman, P.J.A., A.N. Sharpley, T.L. Veith, R.O. Maguire, and P.A. Vadas. 2004. Evaluation of phosphorus transport in surface runoff from packed soil boxes. J. Environ. Qual. 33(4): 1413–1423.

Kleinman, P.J.A., A.N. Sharpley, A.M. Wolf, D.B. Beegle, and P.A. Moore. 2002b. Measuring Water-Extractable Phosphorus in Manure as an Indicator of Phosphorus in Runoff. Soil Sci. Soc. Am. J. 66: 2009–2015.

Kucharik, C.J. 2003. Evaluation of a Process-Based Agro-Ecosystem Model (Agro-IBIS) across the U.S. Corn Belt: Simulations of the Interannual Variability in Maize Yield. Earth Interact. 7: 1–33.

Kucharik, C.J., J.A. Foley, C. Delire, V.A. Fisher, M.T. Coe, J.D. Lenters, C. Young-Molling, N. Ramankutty, J.M. Norman, and S.T. Gower. 2000. Testing the performance of a dynamic global ecosystem model: Water balance, carbon balance, and vegetation structure. Global Biogeochem. Cycles 14(3): 795–825.

Kucharik, C.J., S.P. Serbin, S. Vavrus, E.J. Hopkins, and M.M. Motew. 2010. Patterns of Climate Change Across Wisconsin From 1950 to 2006. Phys. Geogr. 31(1): 1–28.

Kudela, R.M., E. Berdalet, S. Bernard, M. Burford, L. Fernand, S. Lu, S. Roy, P. Tester, G. Usup, R. Magnien, D.M. Anderson, A. Cembella, M. Chinain, G. Hallegraeff, B. Reguera, A. Zingone, H. Enevoldsen, and E. Urban. 2015. Harmful algal blooms: a scientific summary for policy makers. Available at http://ecite.utas.edu.au/101022 (verified 12 February 2017).

Kunkel, K.E., T.R. Karl, and D.R. Easterling. 2007. A Monte Carlo Assessment of Uncertainties in Heavy Precipitation Frequency Variations. J. Hydrometeorol. 8(5): 1152–1160.

Langdale, G.W., R.A. Leonard, and A.W. Thomas. 1985. Conservation practice effects on phosphorus losses from Southern Piedmont watersheds. J. Soil Water Conserv. 40(1): 157–161.

Lathrop, R.C. 2007. Perspectives on the eutrophication of the Yahara lakes. Lake Reserv. Manag. 23: 345–365.

MacDonald, G.K., and E.M. Bennett. 2009. Phosphorus Accumulation in Saint Lawrence River Watershed Soils: A Century-Long Perspective. Ecosystems 12(4): 621–635.

McCabe, G.J., and D.M. Wolock. 2016. Variability and Trends in Runoff Efficiency in the Conterminous United States. J. Am. Water Resour. Assoc. 52(5): 1046–1055.

McDowell, L.L., and K.C. McGregor. 1984. Plant nutrient losses in runoff from conservation tillage corn. Soil Tillage Res. 4: 79–91.

McElroy, A.D., S.Y. Chiu, J.W. Nebgen, A. Aleti, and F.W. Bennett. 1976. Loading functions for assessment of water pollution from nonpoint sources. EPA-600/2-76-151.

Menne, M.J., I. Durre, R.S. Vose, B.E. Gleason, and T.G. Houston. 2012. An Overview of the Global Historical Climatology Network-Daily Database. J. Atmos. Ocean. Technol. 29(7): 897–910.

Michalak, A.M. 2016. Study role of climate change in extreme threats to water quality. Nature 535(7612): 349–350.

Michalak, A.M., E.J. Anderson, D. Beletsky, S. Boland, N.S. Bosch, T.B. Bridgeman, J.D. Chaffin, K. Cho, R. Confesor, I. Daloglu, J.V. Depinto, M.A. Evans, G.L. Fahnenstiel, 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. Proc. Natl. Acad. Sci. U. S. A. 110(16): 6448–6452.

Motew, M., X. Chen, E.G. Booth, S.R. Carpenter, P. Pinkas, S.C. Zipper, S.P. Loheide, S.D. Donner, K. Tsuruta, P. Vadas, and C.J. Kucharik. 2017. The influence of legacy P on lake water quality in a Midwestern agricultural watershed. Ecosystems In press.

Notaro, M., D. Lorenz, C. Hoving, and M. Schummer. 2014. 21st century projections of snowfall and winter severity across central-eastern North America. J. Clim.Available at http://dx.doi.org/10.1175/JCLI-D-13-00520.1.

Owens, P.N., and D.E. Walling. 2002. The phosphorus content of fluvial sediment in rural and industrialized river basins. Water Res. 36(3): 685–701.

Peterson, T.C., X. Zhang, M. Brunet-India, and J.L. Vázquez-Aguirre. 2008. Changes in North American extremes derived from daily weather data. J. Geophys. Res. 113(D7): 1–9.

Powers, S.M., T.W. Bruulsema, T.P. Burt, N.I. Chan, J.J. Elser, P.M. Haygarth, N.J.K. Howden, H.P. Jarvie, Y. Lyu, H.M. Peterson, A.N. Sharpley, J. Shen, F. Worrall, and F. Zhang. 2016. Long-term accumulation and transport of anthropogenic phosphorus in three river basins. Nat. Geosci. 9(5): 353–356.

Pryor, S.C., J.A. Howe, and K.E. Kunkel. 2009. How spatially coherent and statistically robust are temporal changes in extreme precipitation in the contiguous USA? Int. J. Climatol. 29(1): 31–45.

Qian, T., A. Dai, and K.E. Trenberth. 2007. Hydroclimatic Trends in the Mississippi River Basin from 1948 to 2004. J. Clim. 20(18): 4599–4614.

Richardson, C.W., and D.A. Wright. 1984. WGEN: A model for generating daily weather variables. U.S. Department of Agriculture, Agricultural Research Service, ARS-8, Springfield, VA.

Rowe, H., P.J.A. Withers, P. Baas, N.I. Chan, D. Doody, J. Holiman, B. Jacobs, H. Li, G.K. MacDonald, R. McDowell, A.N. Sharpley, J. Shen, W. Taheri, M. Wallenstein, and M.N. Weintraub. 2015. Integrating legacy soil phosphorus into sustainable nutrient management strategies for future food, bioenergy and water security. Nutr. Cycling Agroecosyst.: 1–20.

Royer, T.V., M.B. David, and L.E. Gentry. 2006. Timing of riverine export of nitrate and phosphorus from agricultural watersheds in Illinois: implications for reducing nutrient loading to the Mississippi River. Environ. Sci. Technol. 40(13): 4126–4131.

Sattari, S.Z., M.K. van Ittersum, K.E. Giller, F. Zhang, and A.F. Bouwman. 2014. Key role of China and its agriculture in global sustainable phosphorus management. Environ. Res. Lett. 9(5): 054003.

Schilling, K.E., M.K. Jha, Y.-K. Zhang, P.W. Gassman, and C.F. Wolter. 2008. Impact of land use and land cover change on the water balance of a large agricultural watershed: Historical effects and future directions. Water Resour. Res. 44(7): 1–12.

Scott, C.A., M.F. Walter, G.N. Nagle, M. Todd Walter, N.V. Sierra, and E.S. Brooks. 2001. Residual phosphorus in runoff from successional forest on abandoned agricultural land: 1. Biogeochemical and hydrological processes. Biogeochemistry 55(3): 293–310.

Seidl, R. 2017. To Model or not to Model, That is no Longer the Question for Ecologists. Ecosystems 20(2): 222–228.

Sharpley AN, Ahuja LR, Menzel RG. 1981. The release of soil phosphorus to runoff in relation to the kinetics of desorption. J Environ Qual 10:386–91.

Sharpley, A.N. 1985. The selective erosion of plant nutrients in runoff. Soil Sci. Soc. Am. J. 49(6): 1527–1534.

Sharpley, A. 2016. Managing agricultural phosphorus to minimize water quality impacts. Sci. Agric. 73(1): 1–8.

Sharpley, A.N., S.C. Chapra, R. Wedepohl, J.T. Sims, T.C. Daniel, and K.R. Reddy. 1994. Managing agricultural phosphorus for protection of surface waters - issues and options. J. Environ. Qual. 23(3): 437–451.

Sharpley, A., H.P. Jarvie, A. Buda, L. May, B. Spears, and P. Kleinman. 2013. Phosphorus legacy: overcoming the effects of past management practices to mitigate future water quality impairment. J. Environ. Qual. 42(5): 1308–1326.

Sharpley, A.N., R.W. McDowell, and P.J.A. Kleinman. 2004. Amounts, Forms, and Solubility of Phosphorus in Soils Receiving Manure. Soil Sci. Soc. Am. J. 68(6): 2048.

Sharpley, A., and B. Moyer. 2000. Phosphorus Forms in Manure and Compost and Their Release during Simulated Rainfall. J. Environ. Qual. 29: 1462–1469.

Sharpley, A.N., and S.J. Smith. 1994. Wheat tillage and water quality in the Southern Plains. Soil Tillage Res. 30: 33–48.

Silveira, M.L., J.M.B. Vendramini, and L.E. Sollenberger. 2010. Phosphorus Management and Water Quality Problems in Grazingland Ecosystems. International Journal of Agronomy 2010: 1–8.

Smith, D.R., and E.A. Warnemuende-Pappas. 2015. Vertical tillage impacts on water quality derived from rainfall simulations. Soil Tillage Res. 153: 155–160.

Stigliani, W.M., P. Doelman, W. Salomons, R. Schulin, G.R.B. Smidt, and S. Van der Zee. 1991. Chemical Time Bombs: Predicting the Unpredictable. Environment 33(4): 4–30.

Stigliani, W.M., P. Doelman, W. Salomons, R. Schulin, G.R.B. Smidt, and S. Van der Zee. 1991. Chemical Time Bombs: Predicting the Unpredictable. Environment 33(4): 4–30.

Stuntebeck, T., M. Komiskey, M. Peppler, D. Owens, and D. Frame. 2011. Precipitation-runoff relations and water-quality characteristics at edge-of-field stations, Discovery Farms and Pioneer Farm, Wisconsin, 2003–08.

The MathWorks, Inc. 2015. MATLAB and Statistics Toolbox. The MathWorks, Inc., Natick Massachusetts, USA.

Tiessen, K.H.D., J.A. Elliott, J. Yarotski, D.A. Lobb, D.N. Flaten, and N.E. Glozier. 2010. Conventional and conservation tillage: influence on seasonal runoff, sediment, and nutrient losses in the Canadian Prairies. J. Environ. Qual. 39(3): 964–980.

Trenberth, K.E. 2011. Changes in precipitation with climate change. Clim. Res. 47(1): 123–138.

Turner, M.G., and S.R. Carpenter. 2017. Ecosystem Modeling for the 21st Century. Ecosystems 20(2): 211–214.

Usinowicz, J., J. Qiu, and A. Kamarainen. 2017. Flashiness and Flooding of Two Lakes in the Upper Midwest During a Century of Urbanization and Climate Change. Ecosystems 20(3): 601–615.

Uusi-Kämppä, J., B. Braskerud, H. Jansson, N. Syversen, and R. Uusitalo. 2000. Buffer Zones and Constructed Wetlands as Filters for Agricultural Phosphorus. J. Environ. Qual. 29: 151–158.

Vadas, P.A., W.J. Gburek, A.N. Sharpley, P.J.A. Kleinman, P.A. Moore Jr, M.L. Cabrera, and R.D. Harmel. 2007a. A model for phosphorus transformation and runoff loss for surface-applied manures. J. Environ. Qual. 36(1): 324–332.

Vadas, P.A., B.E. Haggard, and W.J. Gburek. 2005. Predicting dissolved phosphorus in runoff from manured field plots. J. Environ. Qual. 34(4): 1347–1353.

Vadas, P.A., R.D. Harmel, and P.J.A. Kleinman. 2007b. Transformations of soil and manure phosphorus after surface application of manure to field plots. Nutr. Cycling Agroecosyst. 77(1): 83–99.

Vadas, P.A., P.J.A. Kleinman, and A.N. Sharpley. 2004. A simple method to predict dissolved phosphorus in runoff from surface-applied manures. J. Environ. Qual. 33(2): 749–756.

Villarini, G., J.A. Smith, and G.A. Vecchi. 2013. Changing Frequency of Heavy Rainfall over the Central United States. J. Clim. 26(1): 351–357.

Viney, N.R., M. Sivapalan, and D. Deeley. 2000. A conceptual model of nutrient mobilisation and transport applicable at large catchment scales. J. Hydrol. 240(1–2): 23–44.

Westra, S., L.V. Alexander, and F.W. Zwiers. 2013. Global Increasing Trends in Annual Maximum Daily Precipitation. J. Clim. 26(11): 3904–3918.

Westra, S., H.J. Fowler, J.P. Evans, L.V. Alexander, P. Berg, F. Johnson, E.J. Kendon, G. Lenderink, and N.M. Roberts. 2014. Future changes to the intensity and frequency of short-duration extreme rainfall. Rev. Geophys. 52(3): 522–555.

Williams, J.R. 1975. Sediment routing for agricultural watersheds. JAWRA 11(5): 965–974.

Williams, J.R., and R.W. Hann. 1978. Optimal operation of large agricultural watersheds with water quality restraints. Texas Water Resources Institute, Texas A&M Univ., Tech. Rept. No. 96.

Zhao, S.L., S.C. Gupta, D.R. Huggins, and J.F. Moncrief. 2001. Tillage and nutrient source effects on surface and subsurface water quality at corn planting. J. Environ. Qual. 30(3): 998–1008.

Zhu, M., H.W. Paerl, G. Zhu, T. Wu, W. Li, K. Shi, L. Zhao, Y. Zhang, B. Qin, and A.M. Caruso. 2014. The role of tropical cyclones in stimulating cyanobacterial (Microcystis spp.) blooms in hypertrophic Lake Taihu, China. Harmful Algae 39: 310–321.

Regime	Mean storm intensity, $\bar{i}$ (mm hr^{-1}) MULTIPLIER	Mean storm duration, $\bar{t}_r$ (hr) MULTIPLIER	Mean interstorm period, $\bar{t}_b$ (hr) MULTIPLIER	Mean annual precipitation (mm)	99th percentile daily precipitation (mm)
HI	1.80	1	1.49	894.9	32.8
LO	0.90	1	0.71	892.5	25.9

Table 1. Parameter multipliers for the rectangular Poisson pulse model used to generate daily precipitation amounts for the two contrasting climate regimes along with summary statistics for each regime (mean annual precipitation and 99th percentile of daily precipitation).

	Field		Stream		Lake	
	R^2	F	R^2	F	R^2	F
Diss. P Conc. (mg L^{-1})	0.49	**7.75****	0.42	**9.07****		
Sed. P Conc. (mg L^{-1})	0.38	0.68^	0.06	1.46		
Tot. P Conc. (mg L^{-1})	0.62	**51.76*****	0.27	**8.37****	0.14	**7.42****
Diss. Yld or Load (kg ha^{-1} or kg)	0.33	3.42^	0.46	0.04^		
Sed. Yld or Load (kg ha^{-1} or kg)	0.24	1.24	0.12	0.34		
Tot. Yld or Load (kg ha^{-1} or kg)	0.31	2.02^	0.31	0.23^		

Table 2. ANOVA results for simulated annual phosphorus quantities including coefficient of determination and F score associated with an interaction between PSUP (P supply) and PREC (precipitation intensity). Significant F scores ($p<0.05$) are shown in bold.
*$p<0.05$
**$p<0.01$
***$p<0.001$
^significant interaction ($p<0.05$) removed after log 10 transformation

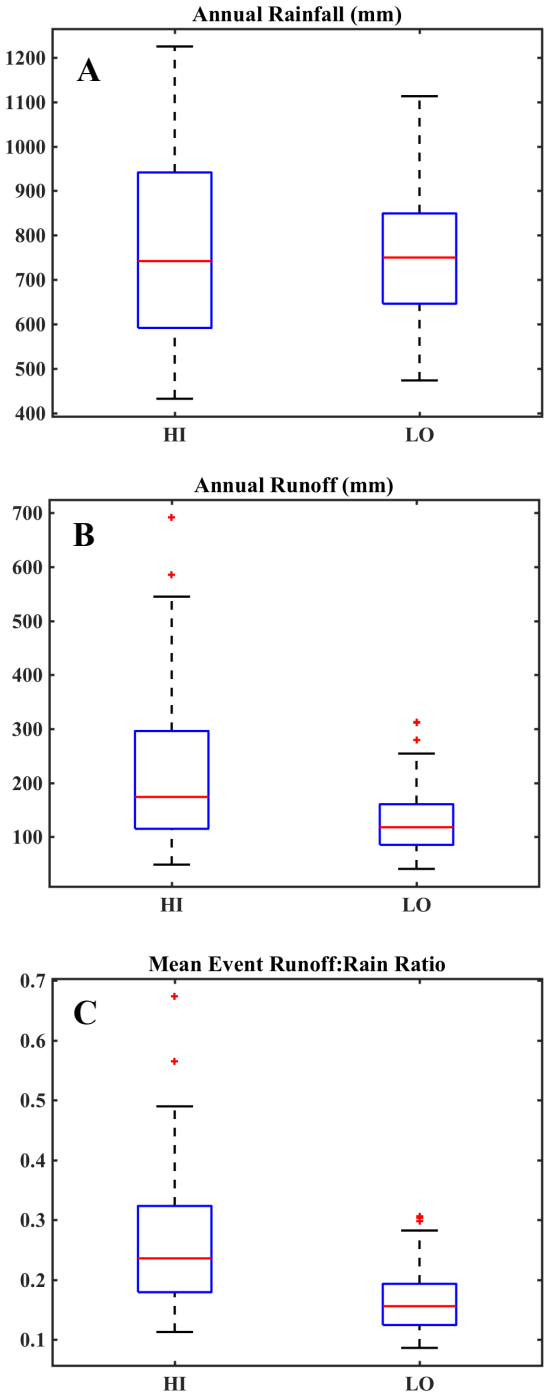


Figure 1. Distributions of annual rainfall (A), annual runoff (B), and mean annual event-based runoff to rain ratio for the HI and LO precipitation regimes.

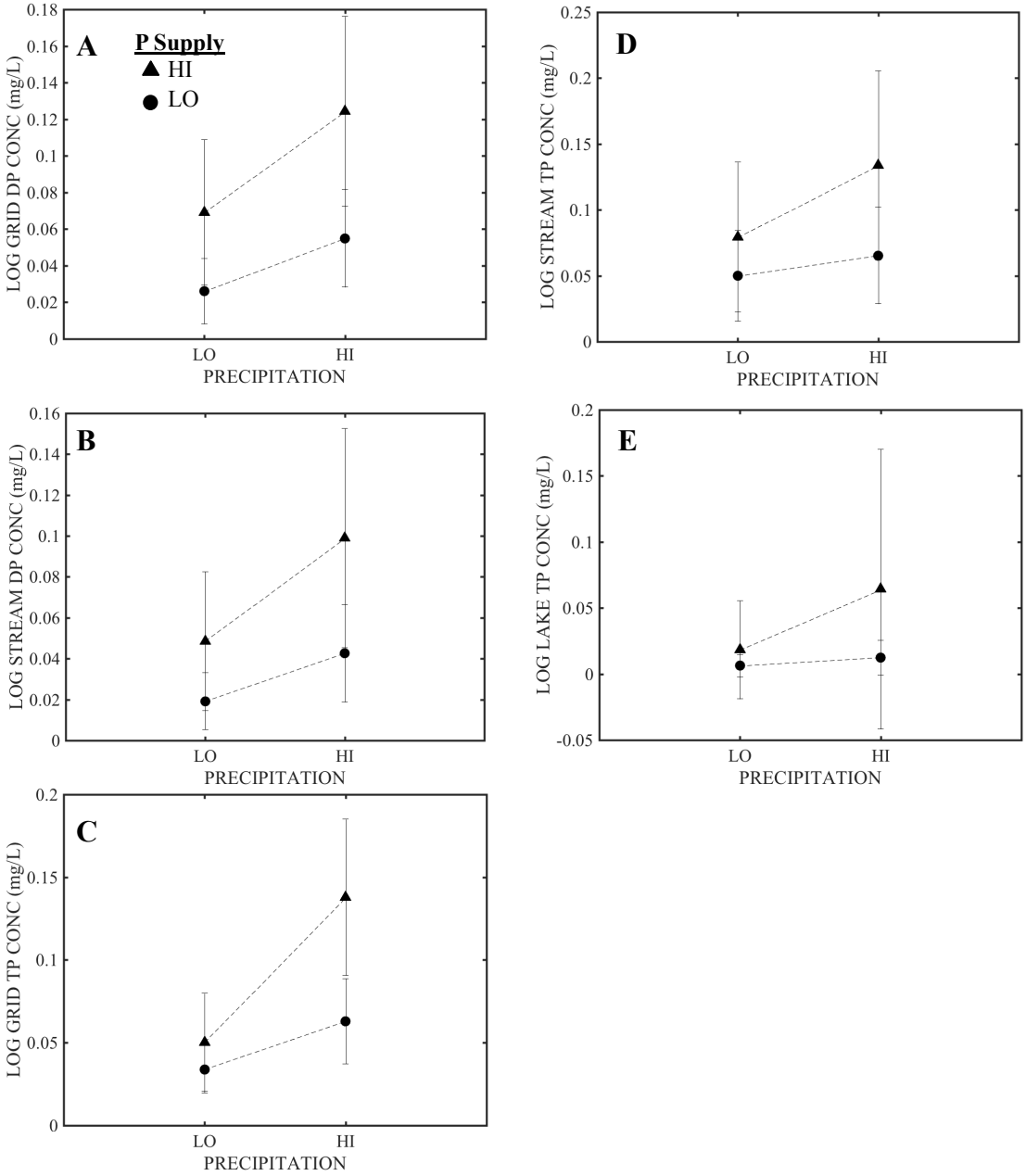


Figure 2. Interaction plots for mean annual values of P response variables for LO and HI levels of terrestrial P supply (PSUP) and precipitation intensity (PREC). Only those P indicator variables in Table 2 having a significant interaction are shown ($p<0.05$). P response variables include total annual dissolved P (DP) concentration in runoff at the field scale (A), total annual DP concentration in the inlet river to Lake Mendota (B), total annual TP concentration in runoff at the field scale (C), total annual TP concentration in the inlet river to Lake Mendota (D), and TP concentration in Lake Mendota (E). Error bars indicate ±1 standard deviation.

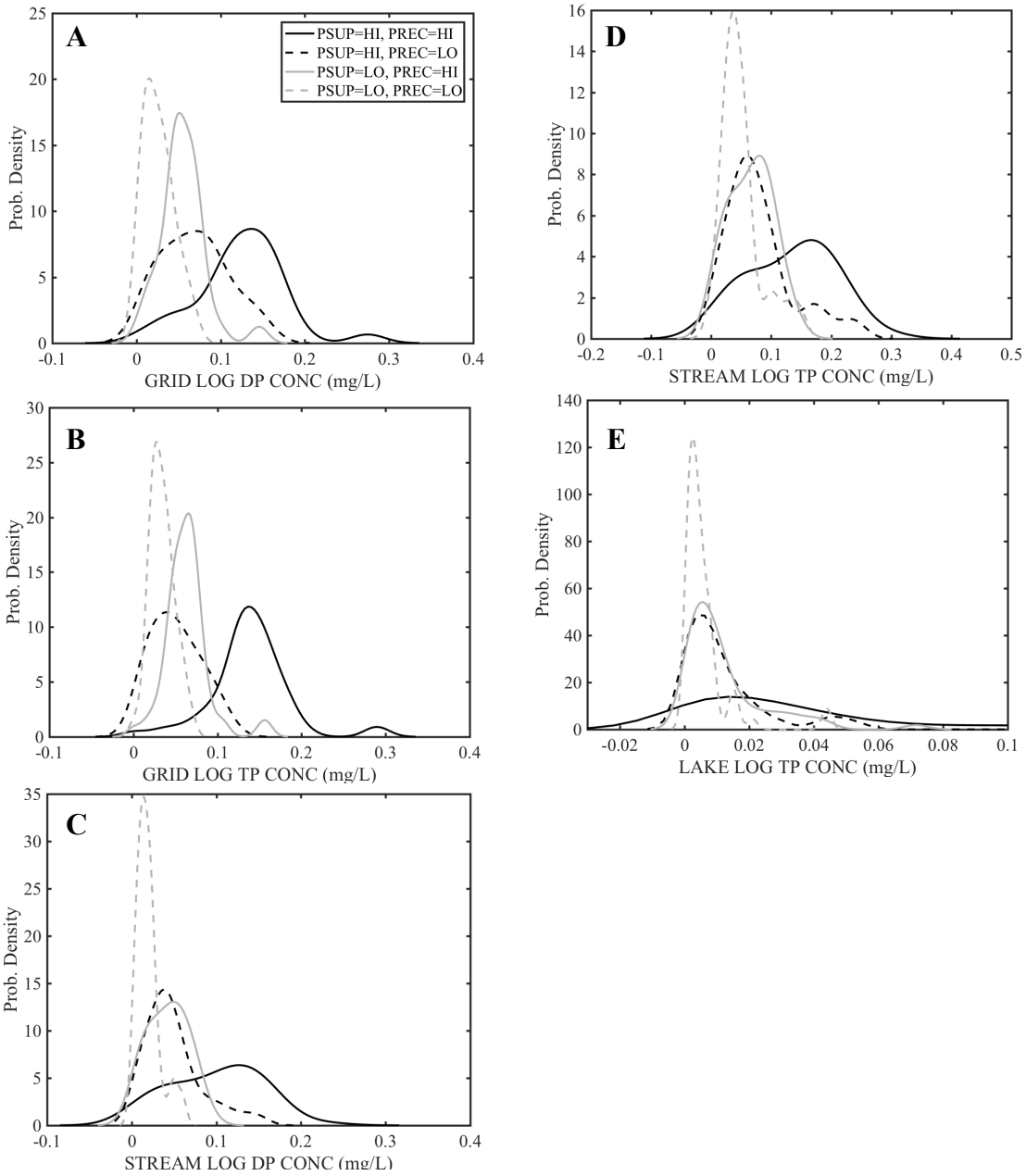


Figure 3. Probability density distributions of annual phosphorus variables. Distributions (n=60) are shown for each of the four factorial simulations combining HI and LO terrestrial P supply (PSUP) and precipitation intensity (PREC). Only the P indicator variables in Table 2 having a significant interaction ($p<0.05$) are shown. Variables include total annual dissolved P (DP) concentration in runoff at the field scale (A), total annual DP concentration in the inlet river to Lake Mendota (B), total annual TP concentration in runoff at the field scale (C), total annual TP concentration in the inlet river to Lake Mendota (D), and summer TP concentration in Lake Mendota (E).

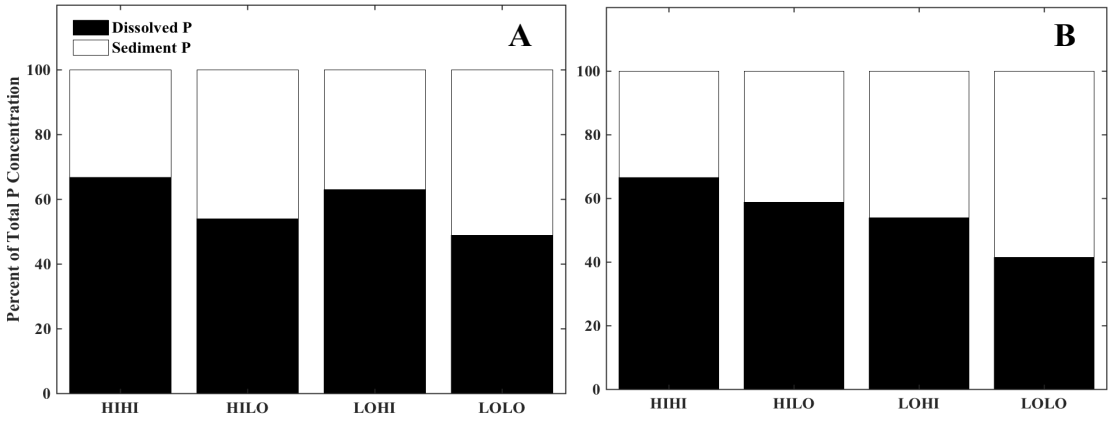


Figure 4. Percent of mean annual total P concentration that is in the dissolved or sediment form for the four simulations. The four simulations are denoted by HI/LO P supply followed by HI/LO precipitation intensity, e.g. HILO stands for HI P supply and LO precipitation intensity. Percentages are shown for P concentration in runoff at the field scale (A) and in the inlet river to Lake Mendota (B).

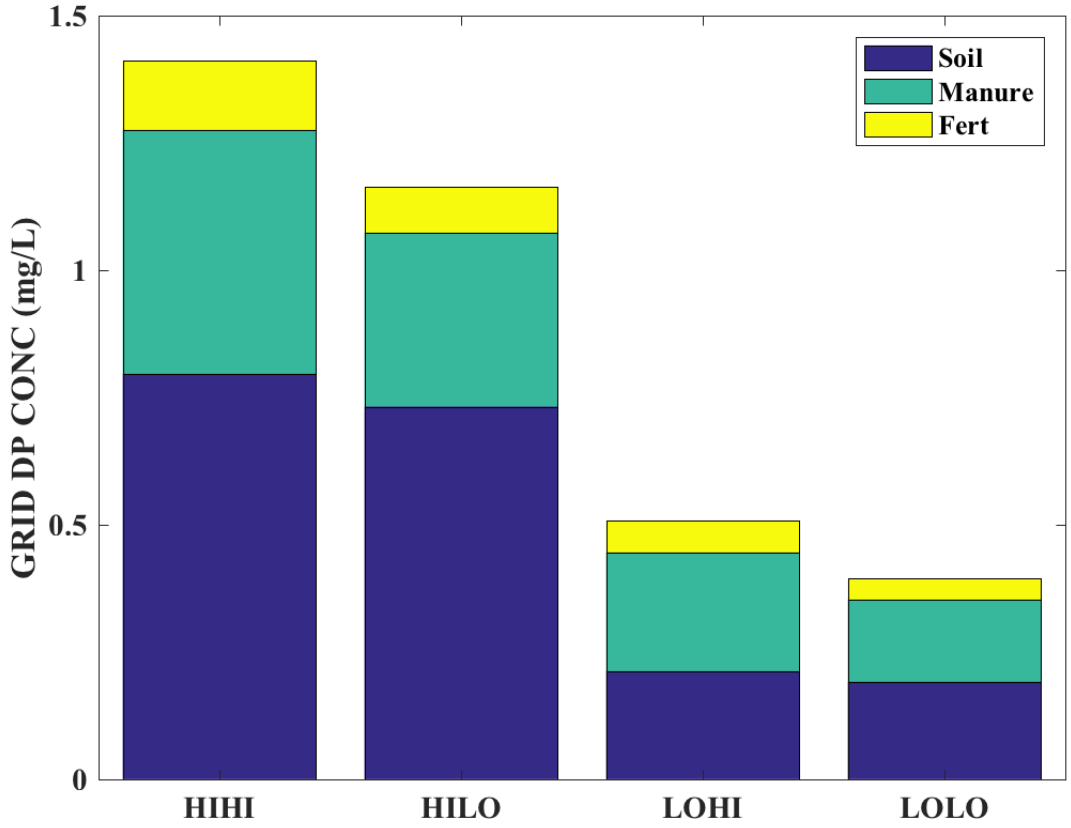


Figure 5. Mean annual dissolved P concentration in runoff for cropland grid cells in each simulation, divided into the three sources of terrestrial P: soil, manure, and fertilizer. The four simulations are denoted by HI/LO P supply followed by HI/LO precipitation intensity, e.g. HILO stands for HI P supply and LO precipitation intensity.

Appendix K. Equations Governing P Loss in Runoff

A.1. Dissolved P Loss in Runoff

For dissolved P loss, the equations in Agro-IBIS come from the SurPhos model which estimates daily amounts of dissolved P loss from manure, fertilizer, and soil (Vadas et al. 2007). The calculation for dissolved P loss from manure is comprised of the following equations.

$$DP_{man} = \frac{LCH_m}{RAIN \times AREA}(10)(PDFACTOR) \quad \text{[Eq. 1]}$$

where DP_{man} is the concentration of dissolved P in runoff from manure (mg L^{-1}), *RAIN* is daily rainfall (mm), *AREA* is the grid-cell area (ha), and LCH_m is water extractable P leached from dairy manure (kg), calculated as:

$$LCH_m = \left(1.2\ \frac{W}{W+73.1}\right) WEP \quad \text{[Eq. 2]}$$

WEP_m is the pool of water extractable P in manure (kg), and W is the water to soil ratio (mL g^{-1}), calculated as:

$$W = \frac{RAIN}{MASS}(COV)(100{,}000) \quad \text{[Eq. 3]}$$

where *MASS* is current manure dry matter mass (kg), and *COV* is the current area covered by manure (ha). The P distribution factor, *PDFACTOR*, is calculated as:

$$PDFACTOR = \left(\frac{RUNOFF}{RAIN}\right)^{0.225} \quad \text{[Eq. 4]}$$

where *RUNOFF* is daily runoff (mm).

Eqs. 1-3 were originally based on relationships governing P desorption from soil to water developed by Sharpley et al. (1981), which showed a dependence on the water-to-soil ratio as well as the time of interaction between water and soil. Kleinman et al. (2002a) showed that these concepts applied well to P losses from dairy, swine, and poultry manures in lab experiments. Subsequently, Vadas et al. (2004) used data from several sources (Sharpley and Moyer 2000, Kleinman 2002b, 2004, Kleinman and Sharpley 2003) as well as their own runoff box experiment to improve the equations as they apply to manure, and modify them to accommodate daily time-step models. The updates included a variable runoff-to-rainfall ratio factor (*PDFACTOR*, Eq. 4) and a deduction of immediate slurry drainage into the soil from the available WEP pool.

For P leached from fertilizer to runoff, Eq. 1 is also used, but leached water extractable P in manure (LCH_m) is replaced by fertilizer WEP leached. All applied fertilizer is water extractable and available to leaching, but a portion becomes incorporated into the soil and becomes part of the soil labile pool. The concentration of dissolved inorganic P in runoff from the soil labile pool (mg L^{-1}) is modeled using a simple linear extraction coefficient, as in:

$$DP_s = LAB \times 0.005 \quad \text{[Eq. 5]}$$

where *LAB* is labile inorganic P in the surface soil layer (mg kg^{-1}). The inorganic soil P pools used in Agro-IBIS resemble those used in SurPhos, with pools for labile, active, and stable forms (Vadas et al. 2007). The labile pool is assumed to be one half of Bray-1 soil test P (Vadas and White 2010). Agro-IBIS also contains two organic soil P pools, fresh and humic, that were included to allow cycling of plant litter and residue back into the soil system as well as simulation of particulate P loss in runoff (Motew et al. 2017, Neitsch et al. 2011).

A.2. Sediment P Loss in Runoff

Daily loss of sediment P is calculated using the total amount of P in the surface 2 cm of soil, daily sediment yield, and a P enrichment ratio (McElroy et al. 1976; Williams and Hann 1978; Sharpley 1985). Sediment yield is calculated using the Modified Universal Soil Loss Equation (MUSLE) (Williams 1975):

$$Y = \alpha\ (Q_{surf} \times Q_{peak})^{\beta} \times K \times C \times LS \times P \quad \text{[Eq. 6]}$$

where *Y* is daily sediment yield (tons), Q_{surf} is surface runoff (m^3), Q_{peak} is peak runoff rate ($m^3\ s^{-1}$), *K* is the soil erodibility factor, *C* is the cover and management factor, *LS* is the topographic factor, and *P* is the support practice factor. *K*, *C*, *LS*, and *P* are all unitless

factors ranging from 0-1. The coefficient and power constants for the MUSLE, given as 11.8 and 0.56 respectively for metric units in Williams (1975), are location-specific parameters that are often calibrated (Sadeghi et al. 2014). Motew et al. (2017) calibrated these parameters to be 5.9 and 0.35, respectively.

The calculation for sediment P (*SP*) is made in units of kg ha^{-1} using the following equation:

$$SP = \frac{0.001(P_{soil})(S)(\varepsilon_{sed})}{AREA} \quad \text{[Eq. 7]}$$

where P_{soil} is the concentration of sediment P in soil (g P ton^{-1} soil), S is sediment yield (tons), and ε_{sed} is the P enrichment value, defined as the ratio of the concentration of P transported with sediment to the concentration in the surface soil layer (Sharpley 1985). The concentration of sediment P in soil is calculated using the sum of P in all soil pools:

$$P_{soil} = \frac{P_{tot}}{(D_b)\,(Depth)} \quad \text{[Eq. 8]}$$

where P_{tot} is the sum of P in the surface soil layer for the three inorganic pools as well as the two organic pools (kg ha^{-1}), D_b is the soil bulk density (g cm^3), and *Depth* is the depth of the surface soil layer (mm).

Further details regarding the terrestrial P module, including losses of P in runoff and soil P transformations are given in the appendix of Motew et al. (2017), available online at www.github.com/mmotew/Appendix-Motew-et-al-2017.git.

Appendix L. Stream Level Dissolved and Sediment P Transport

The in-stream transport of dissolved and sediment P is modeled by THMB. In the modeling framework, dissolved P is considered to be chemically inactive. As a result, dissolved P is simulated as a conservative solute, which follows water flow in stream networks. For sediment P, the in-channel erosion process is modeled using the following equations (Viney et al., 2000):

$$X_p = \frac{k_{XP}}{Y_E^b} \quad \text{[Eq. 9]}$$

$$E = k_E X_p Y_E \quad \text{[Eq. 10]}$$

where X_p is in-channel enrichment ratio; k_{XP} is enrichment coefficient; Y_E is sediment erosion rate (kg s^{-1}); b is enrichment optimization parameter; E is sediment P erosion rate (kg s^{-1}); and k_E is sediment P erosion coefficient. In-channel sediment P deposition is assumed to be proportional to sediment deposition, which is quantified using the following equation:

$$D_P = D_S(\frac{C_P}{C_S}) \quad \text{[Eq. 11]}$$

where D_P is sediment P deposition rate (kg s^{-1}); D_S is sediment deposition rate (kg s^{-1}); C_P is sediment P concentration (kg m^{-3}); and C_S is sediment concentration (kg m^{-3}). The in-channel sediment P storage is tracked in THMB at each time step and the amount of sediment P erosion will not exceed the storage in each grid cell.

References

Kleinman PJA, Sharpley AN. 2003. Effect of broadcast manure on runoff phosphorus concentrations over successive rainfall events. J Environ Qual 32:1072–81.

Kleinman PJA, Sharpley AN, Moyer BG, Elwinger GF. 2002a. Effect of mineral and manure phosphorus sources on runoff phosphorus. J Environ Qual 31:2026–33.

Kleinman PJA, Sharpley AN, Veith TL, Maguire RO, Vadas PA. 2004. Evaluation of phosphorus transport in surface runoff from packed soil boxes. J Environ Qual 33:1413–23.

Kleinman PJA, Sharpley AN, Wolf AM, Beegle DB, Moore PA. 2002b. Measuring Water-Extractable Phosphorus in Manure as an Indicator of Phosphorus in Runoff. Soil Sci Soc Am J 66:2009–15.

McElroy AD, Chiu SY, Nebgen JW, Aleti A, Bennett FW. 1976. Loading functions for assessment of water pollution from nonpoint sources. US Environmental Protection Agency. Washington, D.C. EPA-600/2-76-151.

Motew M, Chen X, Booth EG, Carpenter SR, Pinkas P, Zipper SC, Loheide SP, Donner SD, Tsuruta K, Vadas PA, Kucharik CJ. 2017. The Influence of Legacy P on Lake Water Quality in a Midwestern Agricultural Watershed. Ecosystems In press:1–15.

Neitsch SL, Arnold JG, Kiniry JR, Williams JR. 2011. Soil and Water Assessment Tool Theoretical Documentation Version 2009.

Sadeghi SHR, Gholami L, Darvishan AK, Saeidi P. 2014. A review of the application of the MUSLE model worldwide. Hydrol Sci J 59:365–75.

Sharpley A, Moyer B. 2000. Phosphorus Forms in Manure and Compost and Their Release during Simulated Rainfall. J Environ Qual 29:1462–9.

Sharpley AN. 1985. The selective erosion of plant nutrients in runoff. Soil Sci Soc Am J 49:1527–34.

Sharpley AN, Ahuja LR, Menzel RG. 1981. The release of soil phosphorus to runoff in relation to the kinetics of desorption. J Environ Qual 10:386–91.

Vadas PA, Gburek WJ, Sharpley AN, Kleinman PJA, Moore PA Jr, Cabrera ML, Harmel RD. 2007. A model for phosphorus transformation and runoff loss for surface-applied manures. J Environ Qual 36:324–32.

Vadas PA, Kleinman PJA, Sharpley AN. 2004. A simple method to predict dissolved phosphorus in runoff from surface-applied manures. J Environ Qual 33:749–56.

Vadas PA, White MJ. 2010. Validating Soil Phosphorus Routines in the SWAT Model. Transactions of the ASABE 53:1469–76.

Viney NR, Sivapalan M, Deeley D. 2000. A conceptual model of nutrient mobilisation and transport applicable at large catchment scales. J Hydrol 240:23–44.

Williams JR. 1975. Sediment-yield prediction with Universal Equation using runoff energy factor. In: Present and Prospective Technology for Predicting Sediment Yield and Sources. Vol. ARS-S-40. U.S. Dept. Agric. pp 244–52.

Williams JR, Hann RW. 1978. Optimal operation of large agricultural watersheds with water quality restraints. Texas Water Resources Institute, Texas A&M Univ, Tech Rept No 96.

Chapter 4

How climate and land use/land management drive outcomes of water quality: an investigation into their relative effects and underlying mechanisms at field, stream, and lake scales

Melissa Motew, Xi Chen, Stephen R. Carpenter, Eric G. Booth, Jenny Seifert, Jiangxiao Qiu, Steven P. Loheide II., Jason Schatz, Monica G. Turner, Samuel C. Zipper, Christopher J. Kucharik

Abstract

Eutrophication is a major problem in watersheds where excessive non-point source pollution of phosphorus (P) occurs. Factors that affect P cycling and transport, including climate and land use/land management (LULM), are changing rapidly, making future freshwater resources uncertain. In this study, focusing on the agricultural Yahara Watershed of southern Wisconsin, we explored the relative influences of LULM and climate on indicators of water quality over a span of six decades and at three spatial scales: field, stream, and lake. We also investigated which aspects of LULM were the most important for influencing water quality indicators at each scale. Using a set of biophysical models, we simulated long-term scenarios developed for the YW (yahara2070.org) that explored the consequences of particular human actions including technological innovation, government reform, shifting social values, and environmental negligence. Results showed that climate exerted a stronger influence on water quality overall, yet LULM also played an important role in driving outcomes at field, stream, and lake scales. The most important influence of LULM on water quality came from the average mass balance of P inputs and outputs on agricultural fields, superseding erosion risk and land cover type. The effect of P balance was strongest at the field scale,

but attenuated at the stream and lake scales where the influence of weather variability became greater. This finding, which underscored the dominant role of climate in driving nutrient fluxes within the hydrologic network, suggests an inherent limitation for field scale LULM to influence water quality within streams and lakes. Nevertheless, reducing over-application of P throughout the watershed was an effective management strategy under the four climates investigated, even during decades with wetter conditions and more frequent extreme precipitation events.

4.1. Introduction

Eutrophication of surface waters due to phosphorus (P) enrichment is a worldwide concern, yet manifests as a result of multiple interacting factors. Some factors affecting water quality are relatively stable through time, such as geology, soils, and topography (Kyllmar et al. 2006), but other factors can change more quickly in response to human activities (Vanni et al. 2001). Factors that change temporally include land use type and intensity (Puckett 1995, Gergel et al. 2002, Johnson et al. 1997, Tong and Chen 2002, Smith et al. 2013), composition and configuration of land cover (O'Neill et al. 1997, Clément et al. 2017), management practices (Sharpley et al. 1994, Crossman et al. 2016), engineering structures that alter nutrient flows (Jarvie et al. 2006, Gentry et al. 2007, Arnscheidt et al. 2007), as well as changes in climate and hydrology (Michalak 2016). All of these factors may be shifting and changing simultaneously over a given time period, making it difficult to attribute changes in water quality to any one factor (Gillon et al. 2016). And, continued change in these factors promises to challenge our ability in the future to manage landscapes in ways that promote clean water.

Nonpoint source pollution from agriculture is a primary factor contributing to the impairment of freshwater ecosystems (Carpenter et al. 1998). As such, water quality remediation efforts in many agricultural watersheds have focused on best management practices (BMPs) that decrease nutrient runoff from agricultural lands (Sharpley 2016). However, the effectiveness of BMPs has varied widely within and among watersheds (Baker and Richards 2002, Jordan et al. 2005b, 2007b, Sharpley et al. 2009, Fiener and Auerswald 2009). At the field scale, agricultural and nutrient management programs have reduced P losses, but there has been less reported success at stream and watershed scales (Meals et al. 2010, Jarvie et al. 2013a, b). Thus there exists a research gap in understanding changes in water quality due to LULM across spatial scales, including questions related to the magnitude and timing of responses to management practices (Sharpley et al. 2009, Wood et al. 2005).

In addition to the effects of LULM, climate change may also substantially alter freshwater quality by increasing fluxes of water and P as well as bringing warmer temperatures that favor the growth of harmful algal blooms (Gkelis et al. 2014, Michalak 2016). Nutrient transport can be significantly altered as the timing and magnitude of runoff and soil moisture, lake levels, groundwater availability, and river discharge regimes change (Bates et al. 2008, Crossman et al. 2016). Regions where extreme rain events deliver a disproportionate amount of annual sediment and P loading to water bodies will be particularly vulnerable to increases in rainfall intensity (Haygarth and Jarvis 1997; Royer et al. 2006; Carpenter et al. 2014; Gonzalez-Hidalgo 2013, Motew et al. 2017 *in prep*).

While future projections of climate and LULM are fraught with uncertainty (Carpenter et al. 2006), changes in both drivers are likely to affect water quality outcomes across spatial and temporal scales in the future. Already there is evidence of this, for example in the Lake Erie Basin, where a

concerted effort to implement BMPs over the past two decades has been counteracted by changes in both agricultural practices and climate. These included an increased use of tile drainage, an efficient subsurface transport mechanism of P to surface water bodies, increased surface losses of dissolved P stemming from conservation tillage, and greater frequency of extreme rain events (Michalak et al. 2013). Lake Erie's water quality has remained poor, resulting in several years in the 2010s having massive, harmful cyanobacteria blooms (Michalak 2013). Another watershed exhibiting a diversity of effects from both LULM and climate is the Yahara Watershed (YW) of southern Wisconsin, where BMPs have been implemented with little success. Several counteracting factors have been blamed, including the intensification of dairy agriculture (Gillon et al. 2016, Smith et al. 2013), increased frequency of heavy rain events (Kucharik et al. 2010), as well as the gradual release of legacy P from soils and sediments (Motew et al. 2017).

It is important to understand the challenges and opportunities available at a local level to improve water quality in the face of future uncertainty. Because climate is driven by human activities occurring at a global scale, LULM represents an important avenue by which managers and decision makers can affect outcomes of water quality at a local scale. In this study, using the YW as an exemplar of urbanizing agricultural watersheds, we examined the following two questions: (1) What is the relative importance of variable climate and LULM on water quality indicators at three scales, including field scale P yield, stream P load, and lake total P concentration? (2) Which aspect of human land use and land management has the strongest influence on each indicator? Answers to these questions were intended to provide perspective on the overall potential of LULM to affect surface water quality, and to help set realistic goals for watershed management in the face of climate change. Additionally, knowing which aspects of LULM have the most effect within the watershed

system can help focus and prioritize management strategies to be as effective and efficient as possible.

Because of the long residence time and slow movement of P through watersheds (Hamilton 2012, Sharpley et al. 2013), understanding changes in water quality requires a long-term perspective. To conduct this research, we used 57-year scenarios (section 4.2.2.) in conjunction with biophysical models. Scenarios are plausible stories about how social-ecological systems unfold in response to current factors, new factors, and alternative human choices (Raskin 2005, Carpenter et al. 2006). Scenario development incorporates stakeholder perspectives as well as archetypes from the scenarios literature to describe long-term change in social-ecological systems (Wardropper et al. 2016). Biophysical models, when used to simulate scenarios, help bound the limits of what is physically possible within the narratives, and also enrich the scenarios with quantitative information, for example providing estimates of water quality indicators, or food production (Carpenter et al. 2015). The use of models in conjunction with scenarios also helps to identify important factors controlling the provisioning of ecosystem services, including interactions, feedbacks, and tradeoffs among different services (Qiu and Turner 2013, Qiu et al. 2017 *in review*).

4.2. Methods

4.2.1. Study Area

The 1345 km^2 Yahara Watershed (YW) of southern Wisconsin encompasses a chain of eutrophic lakes, Mendota, Monona, Waubesa, and Kegonsa, connected by the Yahara River. Roughly half the landscape in the YW is devoted to agriculture, with corn, soy, and dairy being

the principal products. The northern and southern parts of the watershed are predominantly agricultural, with dairy operations more common in the north and cash grain in the south. The Wisconsin state capital city of Madison (43°6'N, 89° 24'W) and surrounding urban area comprises roughly one quarter of the watershed and is centered on an isthmus between Lakes Mendota and Monona. The remaining quarter of the watershed is covered in natural vegetation, including forest, wetland, and prairie. The YW is characterized by relatively flat slopes (~4%), and silt loam soils. Nonpoint pollution reduction programs, including the use of best management practices, have been ongoing for several decades, however lake water quality has not improved (Lathrop and Carpenter 2013). Lack of improvement may in part be attributable to the intensification of dairy production, an increase in precipitation and frequency of extreme rain events, as well as the slow release of legacy P from soils and sediments (Gillon et al. 2016, Motew et al. 2017).

4.2.2. Scenarios

The Yahara 2070 scenarios were designed to explore plausible trajectories to the year 2070 in the Yahara Watershed's social-ecological system under different regimes: no action on environmental trends, accelerated technological development, strong intervention by government, and shifting values toward sustainability (Carpenter et al. 2015). The scenarios were developed through an iterative process of adapting archetypal drivers of global change to the perspectives, social processes, and environmental conditions of the YW, within the constraints created by coupling the storylines with biophysical models. A detailed description of the scenario development process and the scenarios themselves is given by Carpenter et al. (2015). Here we

briefly summarize the scenario narratives. Key climatic and land cover characteristics are also provided in Supp. Table 1.

> ***Abandonment and Renewal (AR)*** *explores what could happen if people of the YW are not prepared for environmental challenges that include climate change, deteriorating water quality, and the emergence of a new toxic species. Severe climate disasters elsewhere in the U.S. lead to a national food crisis, putting pressure on the Midwest to increase food production. This leads to an intensification of agriculture and deterioration of water quality which spurs the environmental health disaster. In the wake of the disaster, the watershed undergoes dramatic transformation as humans abandon the area, ending large-scale agriculture and leaving urban areas in ruins. The largely feral landscape allows prairies, forests, and wetlands to rejuvenate. The few remaining people live in dense small towns or on subsistence farms, and survival is the main concern.*
>
> ***Accelerated Innovation (AI)*** *explores what could happen if the U.S. prioritizes technological solutions to environmental challenges including climate change. The YW emerges as one of the nation's solution centers, and money from both public and private sectors is directed toward innovation, specifically in environment, health, and biotechnology. By 2070 the watershed's population has increased sharply as entrepreneurs and businesses establish themselves and infrastructure for innovation and technology expands. Advances such as cultured meats has decreased demand for agricultural land, and new technological capabilities have been focused on improving water quality and assuaging climate change impacts. The physical landscape has become highly engineered as a result.*
>
> ***Connected Communities (CC)*** *explores what could happen if there is a global shift in values toward community and sustainability, spurred by looming environmental and political collapses caused by climate-related disasters and inadequate political response. A worldwide youth movement to embrace sustainability catalyzes the Great Transition, a new paradigm in which*

connectivity, community and environmental sustainability pervade policy and culture. People in the YW live a community oriented and sustainable lifestyle, and the landscape reflects this in a mosaic of small farms, restored natural lands, and small communal areas.

***Nested Watersheds (NW)** explores what could happen if the U.S. reforms how it governs freshwater resources in response to increasing water insecurity. A national water crisis caused by climate change creates public outcry, and the federal government responds with a new water governance framework which draws jurisdiction for water governance around the boundaries of the country's major watersheds. The Yahara Watershed Unit becomes a subunit of the Upper Mississippi Watershed Unit, and is charged with supplying clean water to the country's water-scarce regions. By 2070, water conservation has become the norm, and the YW Subunit has created policies and practices to assist in meeting the mandated targets. Perennial biofuels become ubiquitous in the watershed, and other practices are promoted that reduce erosion and runoff and replenish groundwater.*

4.2.3. Description of Models

The biophysical modeling framework included Agro-IBIS, a terrestrial ecosystem model, THMB, a hydrologic and nutrient routing model, and the Yahara Water Quality (WQ) Model which estimates water quality indicators in the four mainstem Yahara lakes. The Agro-IBIS model simulates the movement of water, energy, momentum, carbon, nitrogen, and phosphorus in both natural and managed ecosystems. The structure of Agro-IBIS has been described in detail (Foley and others 1996; Kucharik and others 2000; Kucharik 2003), and many components of the model have been validated across a range of ecosystems at various spatial and temporal scales (El Maayar and others 2001; Kucharik and Brye 2003; Kucharik and others 2006; Kucharik and Twine 2007; Soylu and others 2014; Zipper and others 2015). Recently,

biogeochemical cycling of P was added to Agro-IBIS to enable field scale simulation of P loss to runoff, i.e. P yield, an indicator of surface water quality. The terrestrial P module in Agro-IBIS features P application, transformation, and loss of dissolved P to runoff; in-soil cycling of organic and inorganic forms of P; and loss of particulate-bound P with erosion. Details describing the recent addition of P cycling to Agro-IBIS as well as the addition of urban and other land cover types can be found in Motew et al. (2017).

Within the model framework, Agro-IBIS is coupled to THMB, a physically-based hydrologic routing model that simulates transport of water, sediment, N, and P at the watershed scale, including delivery of P loads to the YW lakes and the Rock River. THMB links topographic data and river morphological characteristics within the stream network to a set of linear reservoir functions, thereby simulating temporal variability of water flow and storage in the hydrologic system (Coe 1998, 2000; Coe et al. 2008). Nitrogen and phosphorus transport have both been added to THMB (Donner 2002, Motew et al. 2017). Direct drainage loads of P from THMB are passed to the Yahara WQ Model which predicts summer water quality variables in the four mainstem Yahara lakes. The model computes a mass balance for each lake using empirical relationships (Carpenter and Lathrop 2014). Total annual loads to Lake Mendota are calculated using direct drainage loads only since there are no P inputs from other lakes. For the downstream lakes, total annual loads are determined using direct drainage loads from THMB as well as estimates of P export from upstream lakes. Summer water quality is computed from terms of the P mass balance using regressions (Lathrop and Carpenter 2013; Carpenter and Lathrop 2014). All annual quantities are computed for the water year, defined as November 1 – October 31. Development and validation of the entire model framework including Agro-IBIS, THMB, and the Yahara WQ Model, can be found in Motew et al. (2017).

4.2.4. Model Simulations

The four core scenarios, AR, AI, CC, and NW, were each comprised of a set of climate-related drivers and a set of land use-related drivers, both of which were consistent with the core scenario narratives (Booth et al. 2016). The climate drivers were assumed to be driven by factors external to the YW, thus the climate and land use driver datasets could be analyzed independently. We simulated all combinations of LULM and climate, for a total of 16 simulations (Table 1). Our analyses focused on the effects of the four LULM driver data sets (herein referred to as LULM scenarios) under the four different climate driver data sets (climate scenarios).

Translating the four core scenarios into climate and LULM driver data consisted of (1) deriving daily weather inputs by combining climate model projections and a stochastic weather generator; and (2) spatially distributing annual, narrative-based watershed-scale land use/ land cover (LULC) using transition rules and annual manure and fertilizer (N and P) inputs based on current farm and livestock data (Booth et al. 2016). Input drivers related to LULM included LULC category per grid cell (corn, soy, wheat, hay, pasture, alfalfa, urban, wetland, forest, and grassland), the MUSLE C factor that empirically represented the effect of land cover type and management on sediment loss (Williams 1975), as well as nutrient application details (kg and % dry matter content of P and N in manure and fertilizer). Manure P applications were conducted in all receiving cropland grid cells by applying 10% of manure P in February (surface broadcast), 45% at planting (tilled to a 10 cm depth), and 45% applied in early October (surface broadcast). 100% of annual fertilizer P and N was applied at planting. No fertilizer was applied to urban lawns. Climate inputs included daily meteorological quantities (daily temperature, precipitation,

wind speed, solar radiation, and relative humidity) as well annual atmospheric CO_2 concentrations (Appendix M).

Each of the 16 factorial simulations (Table 1) was run over the 2014-2070 time period. Prior to each scenario simulation, a 200 year spin-up was conducted to bring carbon and nitrogen pools to equilibrium in soils and tree biomass, as well as a 25 year spin-up to bring soil P pools to approximate equilibrium. Details regarding historical driver data (pre-2014), as well as calibration and validation of Agro-IBIS and THMB can be found in Motew et al. (2017).

4.2.5. Analysis

Land cover was altered at the beginning of each decade in all four LULM scenarios, making land cover and land management constant during each decade. We therefore focused the analysis on decadal scale changes in the water quality indicators. The water quality indicators included annual total P yield from Agro-IBIS averaged over the Lake Mendota subwatershed (kg ha^{-1} y^{-1}), annual direct drainage load to Lake Mendota from THMB (DDL, kg y^{-1}), and mean annual lake total P concentration in Lake Mendota from the Yahara WQ Model (TP, mg L^{-1}). We first compared the water quality indicators at the decadal scale by grouping by LULM scenario and computing a mean. For example, for P yield in the AR LULM scenario (green curve in Fig. 1a), an average was computed using the four scenarios having AR driver data for LULM (i.e. ARAI, ARAR, ARCC, and ARNW in Table 1). Percent deviations from the mean were also calculated by subtracting the mean from each of the four individual scenarios. The same approach to obtain mean indicator response and percent deviation per decade was also applied to climate, where water quality indicators were grouped by climate scenario rather than LULM scenario.

We performed an attribution analysis to identify the most important LULM factors affecting the three water quality indicators. Multiple regression analysis was used, with eligible predictors chosen to include only those LULM factors that were directly controlled by human decision making. These included mean annual values of manure P application rate, fertilizer P application rate, P harvest rate, the Modified Universal Soil Loss Equation "cover and management" or "C" factor, and annual percent coverage of the watershed in cropland, corn, soy, alfalfa, urban, grassland, forest, and pasture. We simplified the list of predictors by combining manure P, fertilizer P, and harvest P into a composite mass balance index, PBAL, equal to manure P + fertilizer P – harvest P. The analysis began with a full regression model computed using all predictors. From there, correlations were examined between all statistically significant predictors. Any significant predictor was then removed from the regression analysis if there was a high correlation with another predictor found to be more significant, i.e. had a lower *p* value. For example, PBAL as well as percent cover in alfalfa and pasture were each significant predictors of P yield ($p < 0.05$). However, because both land cover types were highly correlated with PBAL (Supp. Table 2), but less significant predictors of P yield than was PBAL, they were eliminated from the regression analysis for P yield. Coefficients of determination as well as partial coefficients of determination, R^2 and partial R^2, respectively, were used to compare the importance of various predictors in determining each water quality indicator.

We compared the relative impacts of climate versus LULM on each output variable by plotting the annual range in each output variable as observed across the LULM and climate scenarios, respectively. Annual differences in the range caused by climate and LULM were calculated, and the number of years with a greater range due to either climate or LULM was tallied and graphically compared. Finally, to better understand temporal trends in the two forms

of P (dissolved and sediment), we visually compared outcomes in watershed averaged sediment P yield and dissolved P yield averaged across each LULM scenario. All analyses were conducted using MATLAB (The MathWorks, Inc. 2015).

4.3. Results

4.3.1. Outcomes of water quality indicators: decadal responses

Mean annual P yield, DDL, and TP followed similar trajectories when averaged over the six decades and grouped by LULM scenario (Fig. 1a-c). The general trajectory showed each P response variable peak in the third decade, and then decrease over the last four decades. The highest value of mean annual P yield occurred in decade two of the AR LULM scenario, at 1.29 kg ha^{-1}. The lowest value of P yield occurred in the last decade in the NW scenario, at 0.42 kg ha^{-1}. For DDL and TP, the highest values both occurred in decade three for the AR LULM scenario, at 49,880 kg y^{-1} and 0.076 mg L^{-1}, respectively. The lowest values of DDL and TP both occurred in decade six for the NW LULM scenario, at 21,138 kg y^{-1} and 0.028 mg L^{-1}, respectively. The second decade featured P yield and DDL at higher mean values than during the first decade, however TP declined in decade two before increasing in decade three.

The standard error in P yield, DDL, and TP, indicated by error bars in Fig. 1, reflected the range in response due to variations in climate in each decade. For P yield, the least variation due to climate occurred in CC during decade two (ste = 0.05 kg ha^{-1}), and the most variation occurred in NW during decade three (0.23 kg ha^{-1}). For DDL, the least variability due to climate occurred in NW in the second decade (2,008 kg y^{-1}), and the greatest variability during AR in the first

decade (13,366 kg y^{-1}). For TP, NW in decade five featured the least variability due to climate (0.002 mg L^{-1}), while the first decade in AR featured the greatest variability (0.04 mg L^{-1}).

4.3.2. Rankings of the four LULM scenarios

Decadal trends in the four LULM scenarios were assessed for P yield, DDL, and TP, in terms of percent annual deviation from the mean under a given climate (Fig. 2). The deviations showed relatively consistent trends across the climate scenarios and across the three water quality indicators. During the first two decades, P yield had the highest values across all four climates in the AR LULM scenario (Fig. 2a-d). CC meanwhile had the lowest P yield during the first three decades. In the third and fourth decades, NW had the highest P yield, followed by CC in decades five and six. The lowest P yields occurring in decades four and five were in AR, and by decade six the lowest P yield was in NW. Overall, the AI LULM scenario had the smallest deviations in P yield. For all decades and climates, percent deviation in P yield ranged from -23 to +32 ± 13% (standard deviation).

The patterns in DDL and TP were similar to P yield. With the exception of the CC climate (Fig. 2g and 2k), the highest deviations in DDL and TP during the first three decades occurred in AR. For the AR and NW climates, the highest deviations in DDL and TP during the fourth decade occurred during the NW LULM scenario. In the fifth and sixth decades, all climates had the highest deviations in DDL and TP occurring in CC. The lowest values of DDL and TP, signifying the best water quality at the stream and lake scales, occurred first in CC (roughly decades one through three), followed by AR in decade four (with the exception of the

NW climate), and then NW in decade six. Among all decades and climates, percent deviation in DDL ranged from -16 to +25 ± 9%. Percent deviation in TP ranged from -17 to +22 ± 8%.

For similar plots grouped by LULM, which showed the deviations caused by climate (Fig. 3), percent deviation in P yield, DDL, and TP ranged from -19 to 27 ± 13%, -30 to +45 ± 17%, and -46 to +106 ± 29%, respectively. The best (worst) decadal rankings of the climates generally corresponded to the driest (wettest) (Supp. Table 1), and were similar across the three water quality indicators. The worst climate (i.e. the climate with the highest P value) in decade one was NW, followed by AR in decades three and four, and ending with a combination of CC and AI in decades five and six. The best climate in decade one occurred in CC, whereas all climates in decade two had similar effects on the water quality indicators (i.e. the deviations from zero were relatively small). In decades three and four AI had the best climate, and in decades five and six AR had the best climate.

4.3.3. Relative influences of climate and LULM

A visual comparison of the ranges in climate and LULM showed that the influence of LULM on P yield, DDL, and TP (Fig. 4a-c), was small initially but then grew to be larger by the second and third decade. The influence of climate on DDL and TP (Fig. 4d-f) was relatively large in the first few years of the simulation time period, small in the second decade, and large again in the third decade. This was because the NW climate scenario featured heavy precipitation in the first seven years of simulation (i.e. the first "decade" of analysis), and AR featured heavy precipitation events in the third decade. The effect of heavy precipitation and flooding in both of those decades resulted in increased loads, TP, and ultimately a wide range in results among

LULM scenarios. A numeric comparison made between the ranges due to climate and LULM showed that for P yield, the range due to LULM was greater than the range due to climate in 36 of the 57 simulated years (the difference was greater than zero), whereas the range due to climate was greater in only 21 years (difference was less than zero) (Fig. 5). In contrast, the range in DDL and TP was most often dominated by climate. For DDL, only 3 years had a larger range caused by LULM and 54 years had a larger range caused by climate. For TP, 5 years had a larger range in LULM and 52 years had a larger range caused by climate. However, ranges in LULM and climate were similar in magnitude for DDL and especially for TP during the last several decades, suggesting that the dominance of climate was in fact rather small in most years. This weakened dominance of climate over LULM in DDL and TP in the later decades is likely due to the smaller range among the climate scenarios driven by relatively dry conditions in the four scenarios (Supp. Table 1).

4.3.4. Attribution

Linear regression showed that for TP, PBAL, CFAC, alfalfa and pasture were all significant predictors, however all were highly correlated, and PBAL was the most significant. Thus, for decadal P yield, DDL, and TP, PBAL was the primary predictor (Table 2). When broken down into annual fertilizer, manure, and harvest rates, manure was the most significant predictor of P yield and DDL (in terms of partial R^2), followed by fertilizer and then harvest rate. For TP, fertilizer rate was the most important predictor, followed by manure rate and harvest rate which were equally important.

PBAL explained more variability in P yield ($R^2 = 0.81$) than in DDL ($R^2 = 0.58$) and TP ($R^2 = 0.32$), indicating that LULM had its greatest influence at the field scale (Table 2 and Fig. 6). Decadal responses in P yield, DDL, and TP, including the "best" and "worst" scenarios in each decade, were generally consistent with the decadal trends in PBAL (Fig. 1d). The exception occurred between the second and third decades, where PBAL decreased and DDL and TP increased. This could have been due to a greater relative influence of climate on DDL and TP as well as the legacy effects of high PBAL in preceding years.

4.3.5. Sediment P and dissolved P yield

Decadal trends in sediment P and dissolved P yield varied among the LULM scenarios and highlighted how different factors control cycling and transport of each form (Fig. 7). For AI, which was the closest-to-average LULM scenario of the four (Fig. 2a-d), dissolved P yield was greater than sediment P yield in the first three decades and then less than sediment P yield in the last three decades. For AR, dissolved P yield was greater than sediment P yield throughout all decades. This was due in part to high rates of PBAL during the first several decades of AR and the legacy effects of elevated soil P in the decades following. AR featured low sediment P losses after the third decade because following the disaster and human abandonment of the watershed, natural ecosystems became dominant. Regrowth in perennial grasslands and forests did not substantially change soil P reserves, but did decrease the susceptibility to erosion and hence loss of sediment P. By the last decade of AR, in which some humans had resettled in the watershed and begun farming, water quality worsened, and resulted in the second highest levels of total P yield, DDL, and TP among the LULM scenarios. For CC, both forms of P changed little during

the first two decades, as opposed to the other scenarios which generally featured increases in dissolved P yield. By decade three, sediment P began to decrease at a faster rate than dissolved P. The last three decades of CC featured the highest rate of either form of P yield (dissolved in this case) among all four LULM scenarios, and explained why CC had the worst water quality in those decades. Sediment P yield in CC was lower than in AI by the fourth decade due to the widespread conversion to pasture throughout the watershed. However, PBAL remained high compared to the other scenarios since harvest rates were low and applications of manure persisted. This resulted in relatively high losses of dissolved P and hence total P from the landscape. For NW, the first three decades saw relatively high rates of dissolved P yield attributable to a combination of the legacy effects of high soil P at the onset of the simulation, and the slow adoption of fertilizer and manure reductions. By decade six however, NW had the lowest values of total P yield, DDL, and TP of all LULM scenarios (Fig. 1), spurred by the sharp reductions in both sediment P and dissolved P yield over the last three decades (Fig. 7d). Conversion to perennial land cover types having low inputs and high harvest rates (i.e. biofuel production), gave NW a negative PBAL in addition to a low risk of erosion. These two factors together gave NW the best decade of water quality observed among all simulations.

4.4. Discussion

Our study identified climate as the dominant influence on water quality indicators at field, stream, and lake scales. However, our study also highlighted how local management plays a key role in future outcomes. The impact of LULM, driven chiefly by the average field scale P budget on agricultural land, was (1) substantial at field, stream, and lake scales, evidenced by percent deviations from a climate-averaged mean ranging from -23 to +32%, -

16 to +25%, and -17 to +22%, for P yield, DDL, and TP, respectively (Fig. 2); (2) consistent across the four future climate scenarios, as indicated by similar temporal trends in percent deviations from a climate-averaged mean (Fig. 2); and (3) variable throughout the 57 year time period, as demonstrated by changes in the rankings of best/worst LULM scenario per decade. Overall, the results indicated that LULM had an important role in affecting water quality outcomes that was independent from the effects of climate. This suggests that local decision making and management, particularly as they relate to field scale P budgeting, can play a vital role in determining future outcomes of water quality in agricultural watersheds under climate change. As the simulations showed, LULM decisions have the power to improve or degrade water quality in both the near and long-term future.

The dominant influence of climate at stream and lake scales was likely due to a strong association between precipitation and surface hydrology downstream from grid cells. In other words, streamflow was more sensitive to precipitation than was overland flow generation, and streamflow was responsible for driving in-channel rates of erosion and deposition of sediment P (Viney and Sivapalan 1999). The greater importance of climate at the stream and lake scales is consistent with previous observations that precipitation is a dominant driver of inter-annual variability in P loading to the Yahara lake chain (Carpenter et al. 2014, Carpenter and Lathrop 2014). And, other studies have linked nutrient loading to riverine discharge rates (McElroy 1976, Williams and Hann 1978, Littlewood 1995, Correll et al. 1999, Royer et al. 2006). A diminished effect of LULM in streams and lakes implies that there may be an inherent limitation to the effects of land management interventions on downstream water quality. Under this assumption, if extreme precipitation and discharge increase with climate change, the effects of management practices may become less efficient

in improving stream and lake water quality. This, in conjunction with increases in field-scale P losses, will make controlling the source supply of P on the landscape even more important.

Previous studies have shown that the effectiveness of field scale management practices in improving water quality varies widely within and among watersheds (Baker and Richards 2002, Jordan et al. 2005b, 2007b, Sharpley et al. 2009, Fiener and Auerswald 2009). The greater influence of climate at stream and lake scales may help explain why traditional BMPs (e.g. buffers, no-till, etc.) are often successful at stemming P loss at the field scale, but not as successful in improving downstream water quality and ecology (Sharpley et al. 2009, Jarvie et al. 2013). A dominant influence of climate may also help explain why in places such as the YW where more frequent extreme events in recent decades have coincided with increased implementation of BMPs, there has been no improvement in lake P loads or water quality (Gillon et al. 2016).

In a study of thousands of lakes in Wisconsin, Rose et al. (2017) found that annual precipitation modulated the dominant landscape features affecting water quality. They found that during dry years, watershed features such as percent land use/land cover in agriculture became more important predictors of lake water quality than in wet years. This is consistent with our findings that show climate variability strongly drives P loads to Lake Mendota, and in dry years, the effects of LULM are more pronounced (Fig. 5 and Supp. Table 1).

The best/worst outcomes of water quality indicators attributable to LULM in each decade changed primarily in response to the mass balance of P inputs and outputs to the soil-vegetation system. The importance of balancing inputs and outputs, i.e. manure and fertilizer applications in conjunction with harvest/removal rates, superseded erosion risk as well as land cover category. The greater importance of PBAL over erosion risk was evident in the

CC scenario. That scenario featured a widespread conversion of row crops to perennial pasture throughout the watershed, and a subsequent decrease in erosion risk and sediment P yield. However, because P removed by grazing cattle was recycled back to the soil system each year as manure, losses of dissolved P remained high compared with the other three scenarios, ultimately giving CC the worst water quality during the last three decades. NW also featured a conversion to a perennial landscape. However, the high harvest rates of biofuels, as opposed to the un-harvested pastures of CC, allowed for drawdown of soil P reserves and a marked decrease in dissolved P loss (Fig. 7). Reductions in both sediment and dissolved P loss earned NW the best observed water quality outcome of all simulations in the final decade.

The greater importance of PBAL over erosion risk was also evident in the AR scenario, which featured a widespread conversion of cropland to grassland. Despite a significant reduction in erosion risk (a MUSLE C factor of 0.10 compared with 0.80), dissolved P losses remained high for decades due to the release of soil legacy P (Motew et al. 2017). The P-rich ecosystems of AR were also vulnerable to excessive P loss during times of extreme precipitation (e.g. decade three of the AR climate, Supp Table 1 and Fig. 3). BMPs focused on stemming erosion would likely not have reduced erosion any more than a conversion to grassland. It can thus be argued that erosion mitigation measures would not have counteracted the high levels of dissolved P loss that occurred in AR. This assertion is supported by other studies that show when erosion is reduced on agricultural lands, dissolved P loss can still be substantial (Bundy et al. 2001, Kleinman et al. 2008). Reducing the overall P budget at the field scale appeared to be a robust approach for improved water quality in both CC and AR, even during periods of extreme precipitation (Fig. 1). This finding is

supported by previous research suggesting that reducing terrestrial P supply has protective effects on water quality during periods of extreme precipitation (Motew et al. 2017, Motew et al. 2017 *in prep*).

Our finding that PBAL was the best predictor of water quality is supported by a broad literature linking agricultural sources of P, i.e. fertilizer and manure, to water quality indicators at field (Kleinman et al. 2002, Kurz et al. 2005), stream and river (Johnson et al. 1997, David and Gentry 2000), lake (Bennett et al. 1999, Michalak et al. 2013), watershed (Correll et al. 1999, Tong and Chen 2002, Yuan et al. 2013), and basin (Turner and Rabalais 1991) scales. Reducing or eliminating the over-application of P to the landscape is important for limiting the buildup of legacy P, which is associated with elevated levels of soil test P and high losses of P in surface runoff (Sharpley et al. 1994, Andraski and Bundy 2003, Vadas et al. 2005), as well as subsurface flows (Heathwaite and Dils 2000, Simard et al. 2000). Legacy P can affect downstream water quality outcomes over timescales of years to decades (Jarvie et al. 2013, Powers et al. 2016). The lingering effect of legacy P was evident in the AR scenario, where erosion risk was dramatically lowered due to the regeneration of natural (perennial) ecosystems, but losses of dissolved P remained relatively high (Fig. 5). The risk for slow release of legacy P, already a problem in the Yahara (Motew et al. 2017), underscores the need for the removal of P from the soil-vegetation system.

A mass-balance approach to nutrient management that aims to reduce legacy P stores promises direct improvements to water quality, and may help protect against the effects of increasing precipitation and frequency of extreme events. First and foremost, there should be emphasis placed on avoiding excess P applications on farms. Increased monitoring and crediting of soil test P could help prevent applications to areas of farms that are already

saturated with P. Advancements in farm technologies may also play a role, such as precision manure spreading that allows exact rates of P to be applied to specific areas of a single field (Cabot et al. 2006). Then, growing harvestable vegetation in areas of P overabundance may help draw down soil P to agronomic levels over the span of years to a decade (McCollum 1991, Schulte et al. 2010). Within aquatic ecosystems, methods of immobilizing waterborne P (Rydin et al. 2001, Kopáček et al. 2005) or dredging of sediment P (Van der Does et al. 1992) may help reduce legacy P stores that are subject to re-entrainment and/or internal recycling (Søndergaard et al. 2003).

Percent coverage of the watershed in various land cover types was not as important a predictor of water quality indicators as PBAL, even though in other studies land cover type can be a useful predictor of nutrient concentrations (Johnes et al. 1996, Johnson et al. 1997, Gergel et al. 2002, Strayer et al. 2003, Gergel 2005). Land cover type implicitly played a role in the water quality outcomes in this study because land cover type was associated with rates of P application and removal. For example, alfalfa and pasture were significant predictors of P yield, DDL, and TP, due to the fact that they had relatively high rates of applied P (as manure) and low rates of harvest. Their high correlation with PBAL (Supp. Table 2) yet lower significance indicated that the balance of P inputs and outputs was in actuality the chief biophysical mechanism underlying nutrient delivery to waterways. While not considered in this study, the spatial characteristics of land cover, such as shape, configuration, and distance to water bodies, can be important factors affecting surface water quality (Osborne and Wiley 1988, Gergel et al. 2002, Clément et al. 2017). Future research into the spatial arrangement of PBAL and its interactions with hydrologically sensitive areas,

could inform management efforts that seek to target critical source areas of the landscape (Sharpley et al. 2011, Giri et al. 2016).

Even though the LULM scenarios included many different field scale combinations of erosion risk, land cover types, nutrient application rates, and removal rates, the effects of other management practices were not explicitly examined, such as alternative tillage or incorporation methods (e.g. manure injection: Chen 2002), and the seasonal timing of nutrient applications (Vadas et al. 2011, Collick et al. 2016). Future research should evaluate the effects of such strategies at the watershed scale because these methods directly affect the amount of surface P available to runoff, as well as the likelihood for high surface P concentrations to coincide with seasonal patterns of precipitation and runoff (e.g. Correll et al. 1999, Royers et al. 2006).

4.5. Conclusion

Using long-term integrated scenarios and biophysical models we examined the effects of LULM and climate on surface water quality indicators over a span of six decades. Variation in climate was more important than variation in LULM overall, but LULM was a significant predictor of water quality indicators at all scales. The average annual net P balance in 220-m agricultural fields was the most important LULM driver at all scales and under the four climates examined. The results suggested that nutrient management planning that avoids and reduces an excess balance of P (inputs – outputs) at the field scale is the best LULM strategy for improving downstream water quality in the future under climate change. Strategies to block nutrient transport from land to waterways may be less effective, since

perennial land cover types with low erosion risk may still be susceptible to high losses of dissolved P when legacy P is high.

Our findings showed that the influence of LULM was greatest at fine scales, whereas climate was the dominant driver of P loading within the surface hydrologic network. An interesting implication of this finding is that the effects of improved nutrient budgeting or transport control practices that take place on the landscape may be limited in their ability to improve downstream water quality. Such a limitation could become increasingly important as managers cope with the challenges of increasing precipitation and frequency of extreme events which are likely to increase nutrient fluxes to waterways. The assumption however that LULM practices are limited in their ability to influence downstream water quality may hinge on the presence of large amounts of legacy P that effectively drive average water quality conditions (Motew et al. 2017) yet are subject to the dynamics of streamflow. In other words, the slow time scales of watershed P accumulation and depletion, and its overarching influence on average water quality conditions, are juxtaposed by the rapid timescales of interrannual weather variation which closely correlate with annual TP. Thus LULM may act more slowly than precipitation, but not necessarily be less important in driving water quality outcomes across the watershed.

We conclude that LULM efforts should focus on inducing a negative P balance on farms in order to induce a drawdown of legacy P and avoid further buildup in soils and downstream sediments. Until legacy P can be eliminated however, watershed managers may need to consider water quality remediation efforts that also focus within aquatic ecosystems themselves, for example by controlling for invasive species and food web dynamics (Walsh et al. 2016), removing P stored in stream or lakebed sediments (Van der Does et al. 1992),

and using treatments that immobilize P in freshwater bodies (Rydin et al. 2001, Kopáček et al. 2005).

Acknowledgements

This research was supported by the National Science Foundation under Grant Nos. DEB-1440297 and DEB-1038759.

References

Ainsworth, E.A., Long, S.P., 2005. What have we learned from 15 years of free-air CO2 enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising CO2. New Phytol. 165, 351–372.

Andraski, T.W., Bundy, L.G., 2003. Relationships between phosphorus levels in soil and in runoff from corn production systems. J. Environ. Qual. 32, 310–316.

Arnscheidt, J., Jordan, P., Li, S., McCormick, S., McFaul, R., McGrogan, H.J., Neal, M., Sims, J.T., 2007. Defining the sources of low-flow phosphorus transfers in complex catchments. Sci. Total Environ. 382, 1–13.

Baker, D.B., Richards, R.P., 2002. Phosphorus budgets and riverine phosphorus export in northwestern Ohio watersheds. J. Environ. Qual. 31, 96–108.

Bates, B., Kundzeqics, Z.W., Wu, S., Palutikof, J.P. (eds), 2008. Climate change and water, Tech. Pap. VI Intergovermental Panel Clim. Change, IPCC Secretariat, Geneva, Switzerland.

Bennett, E.M., Reed-Andersen, T., Houser, J.N., Gabriel, J.R., Carpenter, S.R., 1999. A Phosphorus Budget for the Lake Mendota Watershed. Ecosystems 2, 69–75.

Bishop, K.A., Leakey, A.D.B., Ainsworth, E.A., 2014. How seasonal temperature or water inputs affect the relative response of C3 crops to elevated [CO2]: a global analysis of open top chamber and free air CO2 enrichment studies. Food and Energy Security 3, 33–45.

Booth, E.G., Qiu, J.X., Carpenter, S.R., Schatz, J., Chen, X., Kucharik, C.J., Loheide, S.P., Motew, M.M., Seifert, J.M., Turner, M.G., 2016. From qualitative to quantitative environmental scenarios: Translating storylines into biophysical modeling inputs at the watershed scale. Environmental Modelling & Software 85, 80–97.

Bundy, L.G., Andraski, T.W., Powell, J.M., 2001. Management practice effects on phosphorus losses in runoff in corn production systems. J. Environ. Qual. 30, 1822–1828.

Cabot, P.E., Pierce, F.J., Nowak, P., Karthikeyan, K.G., 2006. Monitoring and predicting manure application rates using precision conservation technology. J. Soil Water Conserv. 61, 282–292.

Carpenter, S.R., Bennett, E.M., Peterson, G.D., 2006. Scenarios for ecosystem services: an overview. Ecol. Soc. 11, 29.

Carpenter, S.R., Booth, E.G., Gillon, S., Kucharik, C.J., Loheide, S., Mase, A.S., Motew, M., Qiu, J., Rissman, A.R., Seifert, J., Others, 2015. Plausible futures of a social-ecological system: Yahara watershed, Wisconsin, USA. Ecol. Soc. 20, 10.

Carpenter, S.R., Booth, E.G., Kucharik, C.J., Lathrop, R.C., 2014. Extreme daily loads: role in annual phosphorus input to a north temperate lake. Aquat. Sci. 77, 71–79.

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.

Carpenter, S.R., Lathrop, R.C., 2014. Phosphorus loading, transport and concentrations in a lake chain: a probabilistic model to compare management options. Aquat. Sci. 76, 145–154.

Chen, Y., 2002. A liquid manure injection tool adapted to different soil conditions. Trans. ASAE 45, 1729–1736.

Clément, F., Ruiz, J., Rodríguez, M.A., Blais, D., Campeau, S., 2017/1. Landscape diversity and forest edge density regulate stream water quality in agricultural catchments. Ecol. Indic. 72, 627–639.

Coe, M.T., 2000. Modeling Terrestrial Hydrological Systems at the Continental Scale: Testing the Accuracy of an Atmospheric GCM. J. Clim. 13, 686–704.

Coe, M.T., 1998. A linked global model of terrestrial hydrologic processes: Simulation of modern rivers, lakes, and wetlands. J. Geophys. Res. 103, 8885–8899.

Coe, M.T., Costa, M.H., Howard, E.A., 2008. Simulating the surface waters of the Amazon River basin: impacts of new river geomorphic and flow parameterizations. Hydrol. Process. 22, 2542–2553.

Collick, A.S., Veith, T.L., Fuka, D.R., Kleinman, P.J.A., Buda, A.R., Weld, J.L., Bryant, R.B., Vadas, P.A., White, M.J., Harmel, R.D., Easton, Z.M., 2016. Improved Simulation of Edaphic and Manure Phosphorus Loss in SWAT. J. Environ. Qual. 45, 1215–1225.

Correll, D.L., Jordan, T.E., Weller, D.E., 1999. Transport of nitrogen and phosphorus from Rhode River watersheds during storm events. Water Resour. Res. 35, 2513–2521.

Crossman, J., Futter, M.N., Palmer, M., Whitehead, P.G., Baulch, H., Woods, D., Jin, L., Oni, S., Dillon, P.J., 2016. The effectiveness and resilience of phosphorus management practices in the Lake Simcoe Watershed, Ontario, Canada. J. Geophys. Res. Biogeosci. 2015JG003253.

David, M.B., Gentry, L.E., 2000. Anthropogenic Inputs of Nitrogen and Phosphorus and Riverine Export for Illinois, USA. J. Environ. Qual. 29, 494–508.

Donner, S.D., Coe, M.T., Lenters, J.D., Twine, T.E., Foley, J.A., 2002. Modeling the impact of hydrological changes on nitrate transport in the Mississippi River Basin from 1955 to 1994. Global Biogeochem Cycles 16, 1–19.

El Maayar, M., Price, D.T., Delire, C., Foley, J.A., Black, T.A., Bessemoulin, P., 2001. Validation of the Integrated Biosphere Simulator over Canadian deciduous and coniferous boreal forest stands. J. Geophys. Res. 106, 14339.

Fiener, P., Auerswald, K., 2009. Effects of hydrodynamically rough grassed waterways on dissolved reactive phosphorus loads coming from agricultural watersheds. J. Environ. Qual. 38, 548–559.

Foley, J.A., Prentice, I.C., Ramankutty, N., Levis, S., Pollard, D., Sitch, S., Haxeltine, A., 1996. An integrated biosphere model of land surface processes, terrestrial carbon balance, and vegetation dynamics. Global Biogeochem. Cycles 10, 603–628.

Gentry, L.E., David, M.B., Royer, T.V., Mitchell, C.A., Starks, K.M., 2007. Phosphorus transport pathways to streams in tile-drained agricultural watersheds. J. Environ. Qual. 36, 408–415.

Gergel, S.E., 2005. Spatial and non-spatial factors: When do they affect landscape indicators of watershed loading? Landsc. Ecol. 20, 177–189.

Gergel, S.E., Turner, M.G., Miller, J.R., Melack J.M., Stanley, E.H., 2002. Landscape indicators of human impacts to riverine systems. Aquat. Sci. 64, 118-128.

Gillon, S., Booth, E.G., Rissman, A.R., 2016. Shifting drivers and static baselines in environmental governance: challenges for improving and proving water quality outcomes. Regional Environ. Change 16, 759–775.

Giri, S., Qiu, Z., Prato, T., Luo, B., 2016. An Integrated Approach for Targeting Critical Source Areas to Control Nonpoint Source Pollution in Watersheds. Water Resour. Manage. 1–14.

Gkelis, S., Papadimitriou, T., Zaoutsos, N., Leonardos, I., 2014. Anthropogenic and climate-induced change favors toxic cyanobacteria blooms: Evidence from monitoring a highly eutrophic, urban Mediterranean lake. Harmful Algae 39, 322–333.

Gonzalez-Hidalgo, J.C., Batalla, R.J., Cerda, A., 2013/3. Catchment size and contribution of the largest daily events to suspended sediment load on a continental scale. Catena 102, 40–45.

Hamilton, S.K., 2012. Biogeochemical time lags may delay responses of streams to ecological restoration. Freshw. Biol. 57, 43–57.

Haygarth, P.M., Jarvis, S.C., 1997/1. Soil derived phosphorus in surface runoff from grazed grassland lysimeters. Water Res. 31, 140–148.

Heathwaite, A.L., Dils, R.M., 2000. Characterising phosphorus loss in surface and subsurface hydrological pathways. Sci. Total Environ. 251-252, 523–538.

Jarvie, H.P., Neal, C., Withers, P.J.A., 2006. Sewage-effluent phosphorus: a greater risk to river eutrophication than agricultural phosphorus? Sci. Total Environ. 360, 246–253.

Jarvie, H.P., Sharpley, A.N., Spears, B., Buda, A.R., May, L., Kleinman, P.J.A., 2013a. Water Quality Remediation Faces Unprecedented Challenges from "Legacy Phosphorus." Environ. Sci. Technol. 47, 8997–8998.

Jarvie, H.P., Sharpley, A.N., Withers, P.J.A., Scott, J.T., Haggard, B.E., Neal, C., 2013b. Phosphorus mitigation to control river eutrophication: Murky waters, inconvenient truths, and "postnormal" science. J. Environ. Qual. 42, 295–304.

Johnes, P.J., 1996. Evaluation and management of the impact of land use change on the nitrogen and phosphorus load delivered to surface waters: the export coefficient modelling approach. J. Hydrol. 183, 323–349.

Johnson, L., Richards, C., Host, G., Arthur, J., 1997. Landscape influences on water chemistry in Midwestern stream ecosystems. Freshw. Biol. 37, 193–208.

Jordan, P., Arnscheidt, A., McGrogan, H., McCormick, S., 2007. Characterising phosphorus transfers in rural catchments using a continuous bank-side analyser. Hydrol. Earth Syst. Sci. Discuss. 11, 372–381.

Jordan, P., Menary, W., Daly, K., Kiely, G., Morgan, G., Byrne, P., Moles, R., 2005. Patterns and processes of phosphorus transfer from Irish grassland soils to rivers—integration of laboratory and catchment studies. J. Hydrol. 304, 20–34.

Kleinman, P.J.A., Sharpley, A.N., Moyer, B.G., Elwinger, G.F., 2002. Effect of mineral and manure phosphorus sources on runoff phosphorus. J. Environ. Qual. 31, 2026–2033.

Kleinman, P.J.A., Sharpley, A.N., Saporito, L.S., Buda, A.R., Bryant, R.B., 2008. Application of manure to no-till soils: phosphorus losses by sub-surface and surface pathways. Nutr. Cycling Agroecosyst. 84, 215–227.

Kopacek, J., Borovec, J., Hejzlar, J., Ulrich, K., Norton, S.A., Amirbahman, A., 2005. Aluminum control of phosphorus sorption by lake sediments. Environ. Sci. Technol. 39, 8784–8789.

Kucharik, C.J., 2003. Evaluation of a Process-Based Agro-Ecosystem Model (Agro-IBIS) across the U.S. Corn Belt: Simulations of the Interannual Variability in Maize Yield. Earth Interact. 7, 1–33.

Kucharik, C.J., Barford, C.C., Maayar, M.E., Wofsy, S.C., Monson, R.K., Baldocchi, D.D., 2006. A multiyear evaluation of a Dynamic Global Vegetation Model at three AmeriFlux forest sites: Vegetation structure, phenology, soil temperature, and CO2 and H2O vapor exchange. Ecol. Modell. 196, 1–31.

Kucharik, C.J., Brye, K.R., 2003. Integrated BIosphere Simulator (IBIS) yield and nitrate loss predictions for Wisconsin maize receiving varied amounts of nitrogen fertilizer. J. Environ. Qual. 32, 247–268.

Kucharik, C.J., Foley, J.A., Delire, C., Fisher, V.A., Coe, M.T., Lenters, J.D., Young-Molling, C., Ramankutty, N., Norman, J.M., Gower, S.T., 2000. Testing the performance of a dynamic global

ecosystem model: Water balance, carbon balance, and vegetation structure. Global Biogeochem. Cycles 14, 795–825.

Kucharik, C.J., Serbin, S.P., Vavrus, S., Hopkins, E.J., Motew, M.M., 2010. Patterns of Climate Change Across Wisconsin From 1950 to 2006. Phys. Geogr. 31, 1–28.

Kucharik, C.J., Twine, T.E., 2007. Residue, respiration, and residuals: Evaluation of a dynamic agroecosystem model using eddy flux measurements and biometric data. Agric. For. Meteorol. 146, 134–158.

Kurz, I., Coxon, C., Tunney, H., Ryan, D., 2005. Effects of grassland management practices and environmental conditions on nutrient concentrations in overland flow. J. Hydrol. 304, 35–50.

Kyllmar, K., Carlsson, C., Gustafson, A., Ulén, B., Johnsson, H., 2006. Nutrient discharge from small agricultural catchments in Sweden: Characterisation and trends. Agric. Ecosyst. Environ. 115, 15–26.

Lathrop, R.C., Carpenter, S.R., 2013. Water quality implications from three decades of phosphorus loads and trophic dynamics in the Yahara chain of lakes. Inland Waters 4, 1–14.

Leakey, A.D.B., Ainsworth, E.A., Bernacchi, C.J., Rogers, A., Long, S.P., Ort, D.R., 2009. Elevated CO2 effects on plant carbon, nitrogen, and water relations: six important lessons from FACE. J. Exp. Bot. 60, 2859–2876.

Littlewood, I.G., 1995. Hydrological regimes, sampling strategies, and assessment of errors in mass load estimates for United Kingdom rivers. Environ. Int. 21, 211–220.

McCollum, R.E., 1991. Buildup and decline in soil phosphorus: 30-year trends on a typical Umprabuult. Agron. J. 83, 77–85.

McElroy, A.D., Chiu, S.Y., Nebgen, J.W., Aleti, A., Bennett, F.W., 1976. Loading functions for assessment of water pollution from nonpoint sources. US Environmental Protection Agency. Washington, D.C. EPA-600/2-76-151.

Meals, D.W., Dressing, S.A., Davenport, T.E., 2010. Lag time in water quality response to best management practices: a review. J. Environ. Qual. 39, 85–96.

Michalak, A.M., 2016. Study role of climate change in extreme threats to water quality. Nature 535, 349–350.

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

Motew, M., Chen, X., Booth, E.G., Carpenter, S.R., Pinkas, P., Zipper, S.C., Loheide, S.P., Donner, S.D., Tsuruta, K., Vadas, P.A., Kucharik, C.J., 2017. The Influence of Legacy P on Lake Water Quality in a Midwestern Agricultural Watershed. Ecosystems In press, 1–15.

Norby, R.J., Delucia, E.H., Gielen, B., Calfapietra, C., Giardina, C.P., King, J.S., Ledford, J., McCarthy, H.R., Moore, D.J.P., Ceulemans, R., De Angelis, P., Finzi, A.C., Karnosky, D.F., Kubiske, M.E., Lukac, M., Pregitzer, K.S., Scarascia-Mugnozza, G.E., Schlesinger, W.H., Oren, R., 2005. Forest response to elevated CO2 is conserved across a broad range of productivity. Proc. Natl. Acad. Sci. U. S. A. 102, 18052–18056.

O'Neill, R.V., Hunsaker, C.T., Jones, K.B., Riitters, K.H., 1997. Monitoring environmental quality at the landscape scale. Bioscience 47, 513–519.

Osborne, L.L., Wiley, M.J., 1988. Empirical relationships between land use/cover and stream water quality in an agricultural watershed. J. Environ. Manage. 26, 9–27.

Powers, S.M., Bruulsema, 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.M., Sharpley, A.N., Shen, J., Worrall, F., Zhang, F., 2016. Long-term accumulation and transport of anthropogenic phosphorus in three river basins. Nat. Geosci. 9, 353–356.

Puckett, L.J., 1995. Identifying the major sources of nutrient water pollution. Environ. Sci. Technol. 29, 408A–414A.

Qiu, J., Turner, M.G., 2013. Spatial interactions among ecosystem services in an urbanizing agricultural watershed. Proc. Natl. Acad. Sci. U. S. A. 110, 12149–12154.

Qiu J., Carpenter SR., Booth EG., Motew M., Zipper S., Kucharik CJ., Chen X., Loheide II SP., Seifert J., Turner M. (*in review*) Scenarios reveal pathways to sustain future ecosystem services in an agricultural landscape.

Raskin, P.D., 2005. Global Scenarios: Background Review for the Millennium Ecosystem Assessment. Ecosystems 8, 133–142.

Rose, K.C., Greb, S.R., Diebel, M., Turner, M.G., 2017. Annual precipitation regulates spatial and temporal drivers of lake water clarity. Ecol. Appl. 27, 632–643.

Royer, T.V., David, M.B., Gentry, L.E., 2006. Timing of riverine export of nitrate and phosphorus from agricultural watersheds in Illinois: implications for reducing nutrient loading to the Mississippi River. Environ. Sci. Technol. 40, 4126–4131.

Rydin, E., Huser, B., Welch, E.B., 2000. Amount of phosphorus inactivated by alum treatments in Washington lakes. Limnol. Oceanogr. 45, 226–230.

Schulte, R.P.O., Melland, A.R., Fenton, O., Herlihy, M., Richards, K., Jordan, P., 2010. Modelling soil phosphorus decline: Expectations of Water Framework Directive policies. Environ. Sci. Policy 13, 472–484.

Sharpley, A., 2016. Managing agricultural phosphorus to minimize water quality impacts. Sci. Agric. 73, 1–8.

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.

Sharpley, A.N., Chapra, S.C., Wedepohl, R., Sims, J.T., Daniel, T.C., Reddy, K.R., 1994. Managing Agricultural Phosphorus for Protection of Surface Waters: Issues and Options. J. Environ. Qual. 23, 437–451.

Sharpley, A.N., Kleinman, P.J.A., Flaten, D.N., Buda, A.R., 2011. Critical source area management of agricultural phosphorus: experiences, challenges and opportunities. Water Sci. Technol. 64, 945–952.

Sharpley, A.N., Kleinman, P.J.A., Jordan, P., Bergström, L., Allen, A.L., 2009. Evaluating the success of phosphorus management from field to watershed. J. Environ. Qual. 38, 1981–1988.

Simard, R.R., Beauchemin, S., Haygarth, P.M., 2000. Potential for preferential pathways of phosphorus transport. J. Environ. Qual. 29, 97–105.

Smith, A.P., Western, A.W., Hannah, M.C., 2013. Linking water quality trends with land use intensification in dairy farming catchments. J. Hydrol. 476, 1–12.

Soylu, M.E., Kucharik, C.J., Loheide, II, Steven, P., 2014. Influence of groundwater on plant water use and productivity: Development of an integrated ecosystem – Variably saturated soil water flow model. Agric. For. Meteorol. 189-190, 198–210.

Strayer, D.L., Edward Beighley, R., Thompson, L.C., Brooks, S., Nilsson, C., Pinay, G., Naiman, R.J., 2003. Effects of Land Cover on Stream Ecosystems: Roles of Empirical Models and Scaling Issues. Ecosystems 6, 407–423.

The MathWorks, Inc., 2015. MATLAB and Statistics Toolbox. The MathWorks, Inc., Natick Massachusetts, USA.

Tong, S.T.Y., Chen, W., 2002. Modeling the relationship between land use and surface water quality. J. Environ. Manage. 66, 377–393.

Turner, R.E., Rabalais, N.N., 1991. Changes in Mississippi River Water Quality This Century. Bioscience 41, 140–147.

Twine, T.E., Bryant, J.J., T Richter, K., Bernacchi, C.J., McConnaughay, K.D., Morris, S.J., Leakey, A.D.B., 2013. Impacts of elevated CO2 concentration on the productivity and surface energy budget of the soybean and maize agroecosystem in the Midwest USA. Glob. Chang. Biol. 19, 2838–2852.

Vadas, P.A., Jokela, W.E., Franklin, D.H., Endale, D.M., 2011. The effect of rain and runoff when assessing timing of manure application and dissolved phosphorus loss in runoff. JAWRA 47, 877–886.

Vadas, P.A., Kleinman, P.J.A., Sharpley, A.N., Turner, B.L., 2005. Relating soil phosphorus to dissolved phosphorus in runoff: a single extraction coefficient for water quality modeling. J. Environ. Qual. 34, 572–580.

Van der Does, J., Verstraelen, P., Boers, P., Van Roestel, J., Roijackers, R., Moser, G., 1992. Lake restoration with and without dredging of phosphorus-enriched upper sediment layers. Hydrobiologia 233, 197–210.

Viney, N.R., Sivapalan, M., 1999. A conceptual model of sediment transport: application to the Avon River Basin in Western Australia. Hydrol. Process. 13, 727–743.

Walsh, J.R., Carpenter, S.R., Vander Zanden, M.J., 2016. Invasive species triggers a massive loss of ecosystem services through a trophic cascade. Proc. Natl. Acad. Sci. U. S. A. 113, 4081–4085.

Wardropper, C.B., Gillon, S., Mase, A.S., McKinney, E.A., Carpenter, S.R., Rissman, A.R., 2016. Local perspectives and global archetypes in scenario development. E&S 21, 1–12.

Williams, J.R., 1975. Sediment routing for agricultural watersheds. JAWRA 11, 965–974.

Williams, J.R., Hann, R.W., 1978. Optimal operation of large agricultural watersheds with water quality restraints. Texas Water Resources Institute, Texas A&M Univ., Tech. Rept. No. 96.

Wood, F.L., Heathwaite, A.L., Haygarth, P.M., 2005. Evaluating diffuse and point phosphorus contributions to river transfers at different scales in the Taw catchment, Devon, UK. J. Hydrol. 304, 118–138.

Yuan, Y., Locke, M.A., Bingner, R.L., Rebich, R.A., 2013. Phosphorus losses from agricultural watersheds in the Mississippi Delta. J. Environ. Manage. 115, 14–20.

Zipper, S.C., Soylu, M.E., Booth, E.G., Loheide, II, Steven, P., 2015. Untangling the effects of shallow groundwater and soil texture as drivers of subfield-scale yield variability. Water Resour. Res. 51, 6338–6358.

		Decade 1 2014-2020	Decade 2 2021-2030	Decade 3 2031-2040	Decade 4 2041-2050	Decade 5 2051-2060	Decade 6 2061-2070
Mean Annual Prec (mm)	AI	1104	1002	961	922	1008	993
	AR	945	1068	1229	1156	877	891
	CC	900	1075	985	933	1089	963
	NW	1030	1056	1005	1012	795	846
Mean Annual Tmax (°C)	AI	15	16	16	16	16	16
	AR	16	17	18	17	20	20
	CC	16	16	17	17	17	18
	NW	16	17	17	18	18	18
Mean Annual Tmin (°C)	AI	3	4	4	4	5	5
	AR	5	6	7	7	9	8
	CC	5	5	6	6	7	7
	NW	5	6	6	7	7	7
Weekly 5" Events Per year	AI	4	7	6	3	4	10
	AR	4	8	11	10	3	0
	CC	7	6	5	5	6	5
	NW	13	5	6	7	5	3
Atmospheric CO_2 (ppm)	AI	407	431	461	494	524	547
	AR	419	456	506	561	621	683
	CC	412	439	477	517	562	605
	NW	414	443	484	528	577	625
Majority Land Cover Type in YW (% Cover)	AI	Corn(30%)	Corn(29%)	Corn(26%)	Corn(27%)	Corn(27%)	Corn(27%)
	AR	Corn(30%)	Corn(30%)	Grass(24%)	Grass(44%)	Grass(49%)	Grass(49%)
	CC	Corn(30%)	Corn(28%)	Corn(22%)	Corn(12%)	Pasture(13%)	Pasture(14%)
	NW	Corn(30%)	Corn(29%)	Corn(26%)	Corn(22%)	Hay(20%)	Hay(21%)

Supplementary Table 1. Climate and land cover characteristics of the four core scenarios.

	Man. P	Fert. P	Harv. P	PBAL	MUSLE C Factor	Crop	Corn	Soy	Alf	Urb	Past	Grs	For
Man. P	1.00	0.75	0.70	0.97	0.74	0.58	0.62	0.18	0.83	0.08	-0.57	-0.36	-0.60
Fert. P	0.75	1.00	0.93	0.86	0.98	0.83	0.93	0.35	0.91	0.56	-0.84	-0.70	-0.24
Harv. P	0.70	0.93	1.00	0.76	0.92	0.95	0.87	0.50	0.86	0.66	-0.91	-0.85	-0.01
PBAL	0.97	0.86	0.76	1.00	0.85	0.64	0.74	0.19	0.89	0.19	-0.64	-0.42	-0.58
MUSLE C Factor	0.74	0.98	0.92	0.85	1.00	0.81	0.97	0.37	0.87	0.61	-0.87	-0.70	-0.20
Crop	0.58	0.83	0.95	0.64	0.81	1.00	0.75	0.56	0.81	0.66	-0.88	-0.93	0.14
Corn	0.62	0.93	0.87	0.74	0.97	0.75	1.00	0.40	0.76	0.65	-0.85	-0.65	-0.13
Soy	0.18	0.35	0.50	0.19	0.37	0.56	0.40	1.00	0.19	0.27	-0.64	-0.41	0.36
Alf	0.83	0.91	0.86	0.89	0.87	0.81	0.76	0.19	1.00	0.35	-0.71	-0.64	-0.42
Urb	0.08	0.56	0.66	0.19	0.61	0.66	0.65	0.27	0.35	1.00	-0.69	-0.85	0.55
Past	-0.57	-0.84	-0.91	-0.64	-0.87	-0.88	-0.85	-0.64	-0.71	-0.69	1.00	0.78	-0.09
Grs	-0.36	-0.70	-0.85	-0.42	-0.70	-0.93	-0.65	-0.41	-0.64	-0.85	0.78	1.00	-0.39
For.	-0.60	-0.24	-0.01	-0.58	-0.20	0.14	-0.13	0.36	-0.42	0.55	-0.09	-0.39	1.00

Supplementary Table 2. Correlation matrix for LULM drivers including mean annual manure, fertilizer, and harvest P rates, PBAL (equal to manure + fertilizer – harvest), MUSLE “Crop and management” C factor, and percent cover of the Lake Mendota subwatershed in crop, corn, soy, alfalfa, urban, pasture, grassland, and forest cover types.

		Climate			
		AI	**AR**	**CC**	**NW**
LULM	**AI**	AIAI	AIAR	AICC	AINW
	AR	ARAI	ARAR	ARCC	ARNW
	CC	CCAI	CCAR	CCCC	CCNW
	NW	NWAI	NWAR	NWCC	NWNW

Table 1. The 16 factorial simulations conducted for this study that included all combinations of LULM and climate driver data sets. Name indicates LULM followed by climate, e.g. the ARCC simulation had LULM drivers from the AR core scenario, and climate drivers from the CC core scenario. LULM drivers included annual gridded LULC category as well as annual gridded nutrient application rates. Climate drivers include daily gridded meteorological quantities as well as annual atmospheric CO_2 concentration.

Predictor	Yield	DDL	TP
kg ha^{-1} y^{-1}		**R^2**	
PBAL	0.81	0.58	0.32
		Partial R^2	
Annual Fertilizer Rate	0.36	0.21	0.12
Annual Manure Rate	0.44	0.24	0.10
Annual Harvest Rate	0.17	0.16	0.10

Table 2. LULM predictors of decadal-scale P yield (Mendota subwatershed average), direct drainage load to Lake Mendota (DDL), and total P concentration (TP) in Lake Mendota. PBAL is equal to annual fertilizer rate + manure rate – harvest rate.

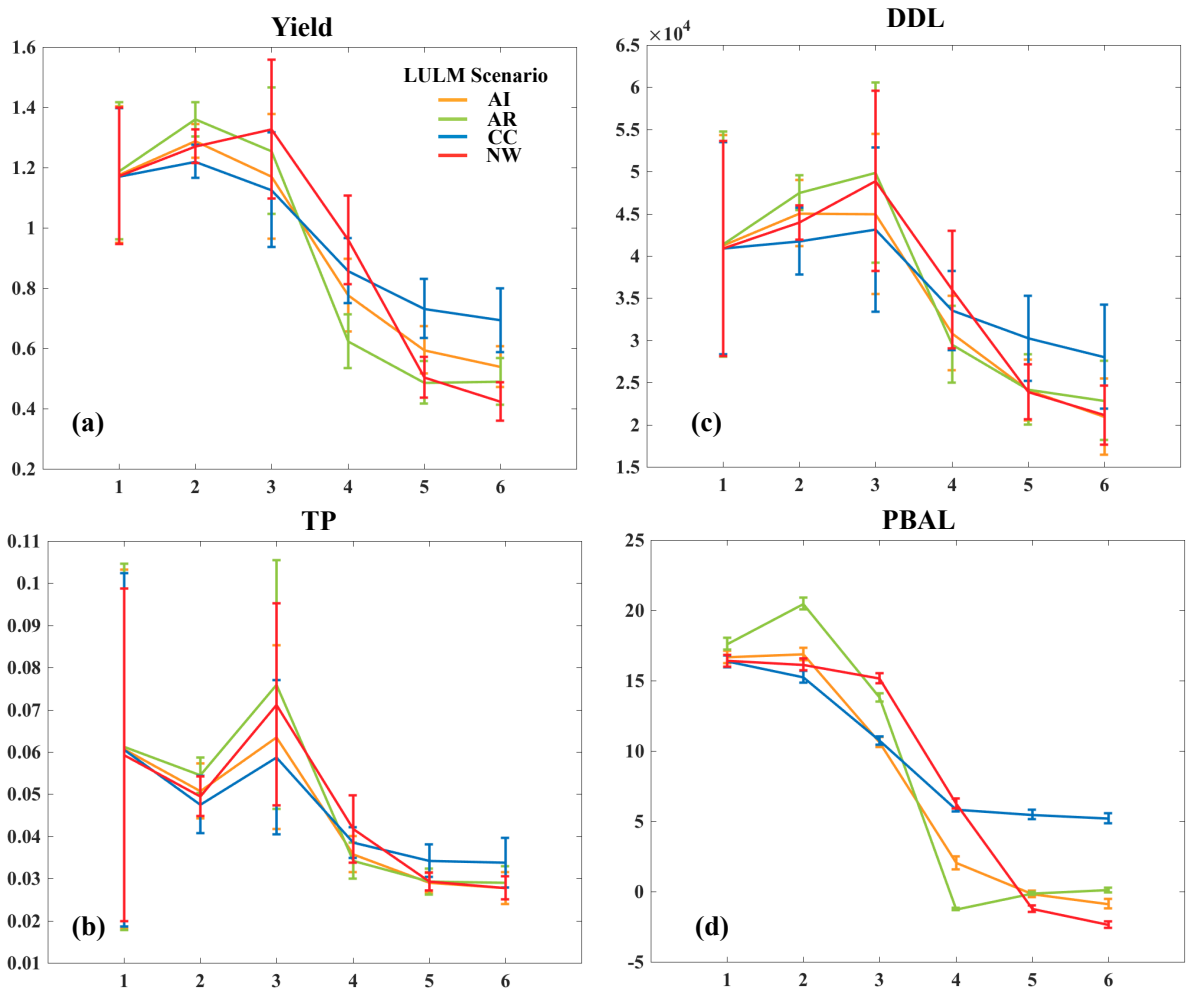


Figure 1. Mean decadal responses in water quality variables under the four LULM scenarios: watershed-averaged total P yield (kg ha^{-1} y^{-1}) (a), mean annual direct drainage load (DDL, kg y^{-1}) (b), mean summertime total P concentration in Lake Mendota (TP, mg L^{-1}) (c), and mean annual PBAL, equal to mean annual manure rate + fertilizer rate – harvest rate (kg ha^{-1} y^{-1}) (d). Error bars represent one standard deviation in the range across the four climate scenarios.

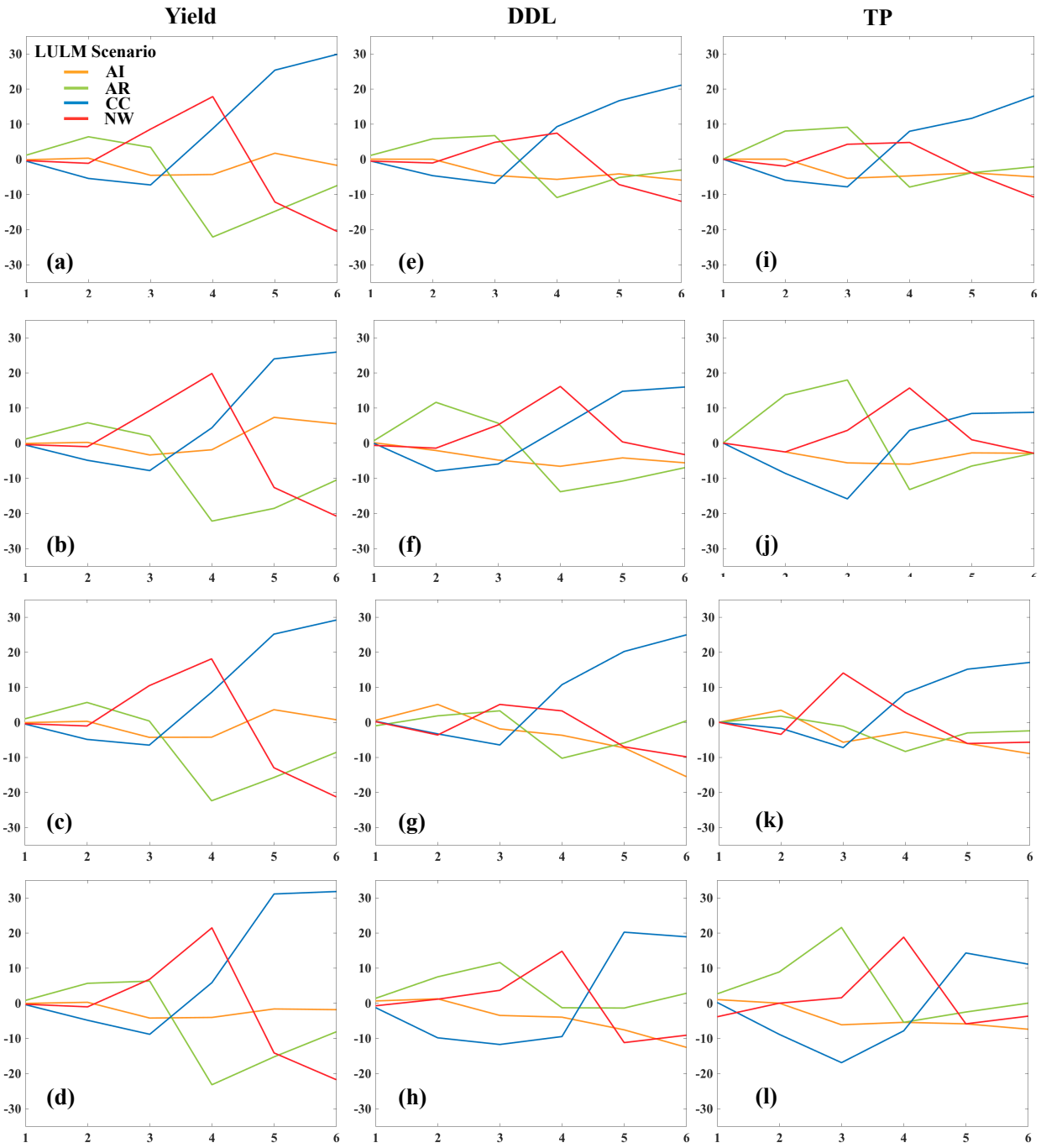


Figure 2. Percent deviation from the climate-averaged mean for each LULM scenario in decades 1-6. First column is watershed-averaged total P yield (kg ha^{-1} y^{-1}), second column is direct drainage load to Lake Mendota (DDL, kg y^{-1}), and third column is summertime lake TP concentration in Lake Mendota (TP, mg L^{-1}). Rows correspond to the climate scenarios AI, AR, CC, and NW, in order of top to bottom, respectively.

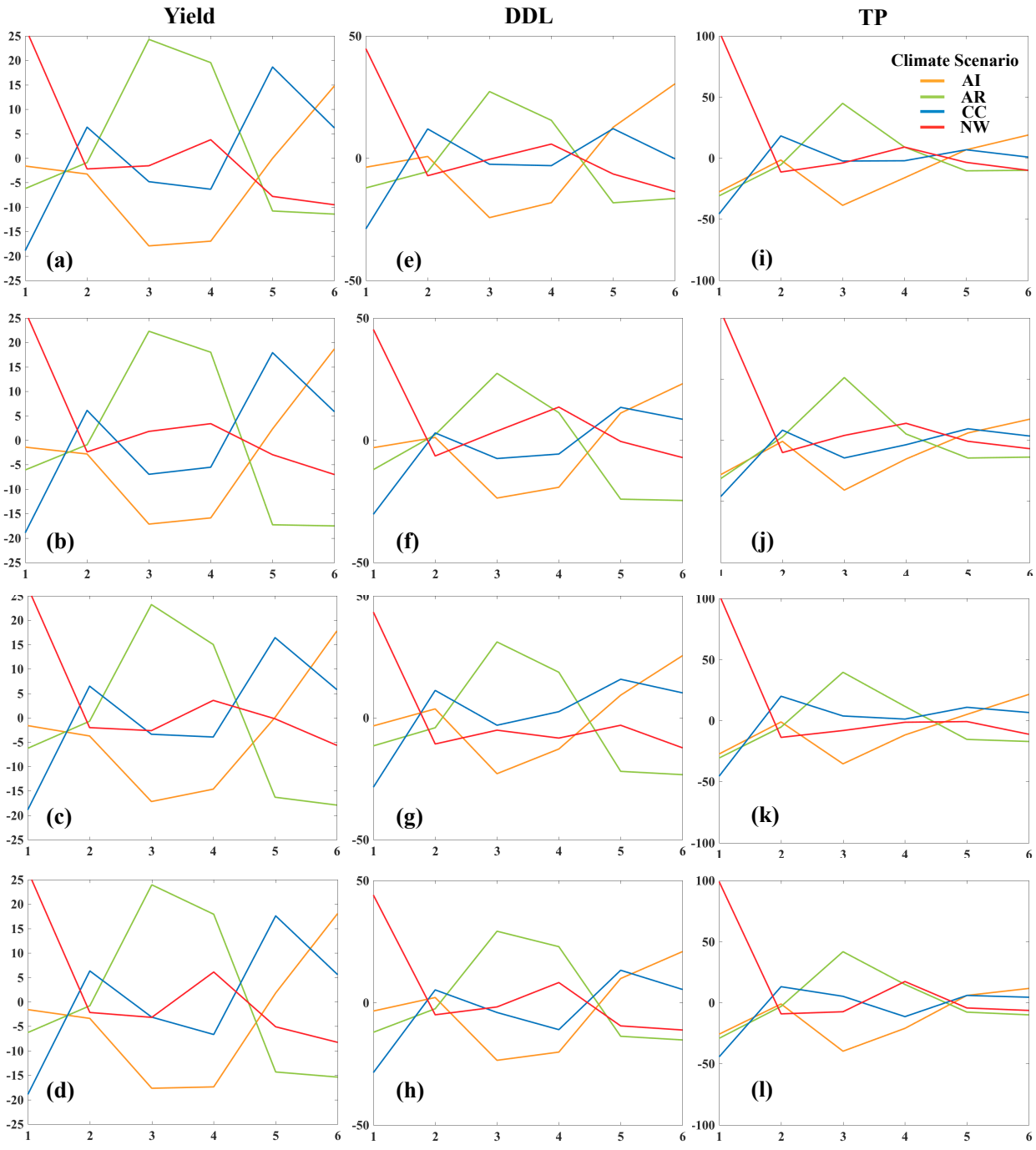


Figure 3. Percent deviation from the LULM-averaged mean for each climate scenario. First column is watershed-averaged total P yield (kg ha^{-1} y^{-1}), second column is direct drainage load to Lake Mendota (kg y^{-1}), and third column is summertime lake TP concentration in Lake Mendota (mg L^{-1}). Rows correspond to the LULM scenarios AI, AR, CC, and NW, in order of top to bottom, respectively.

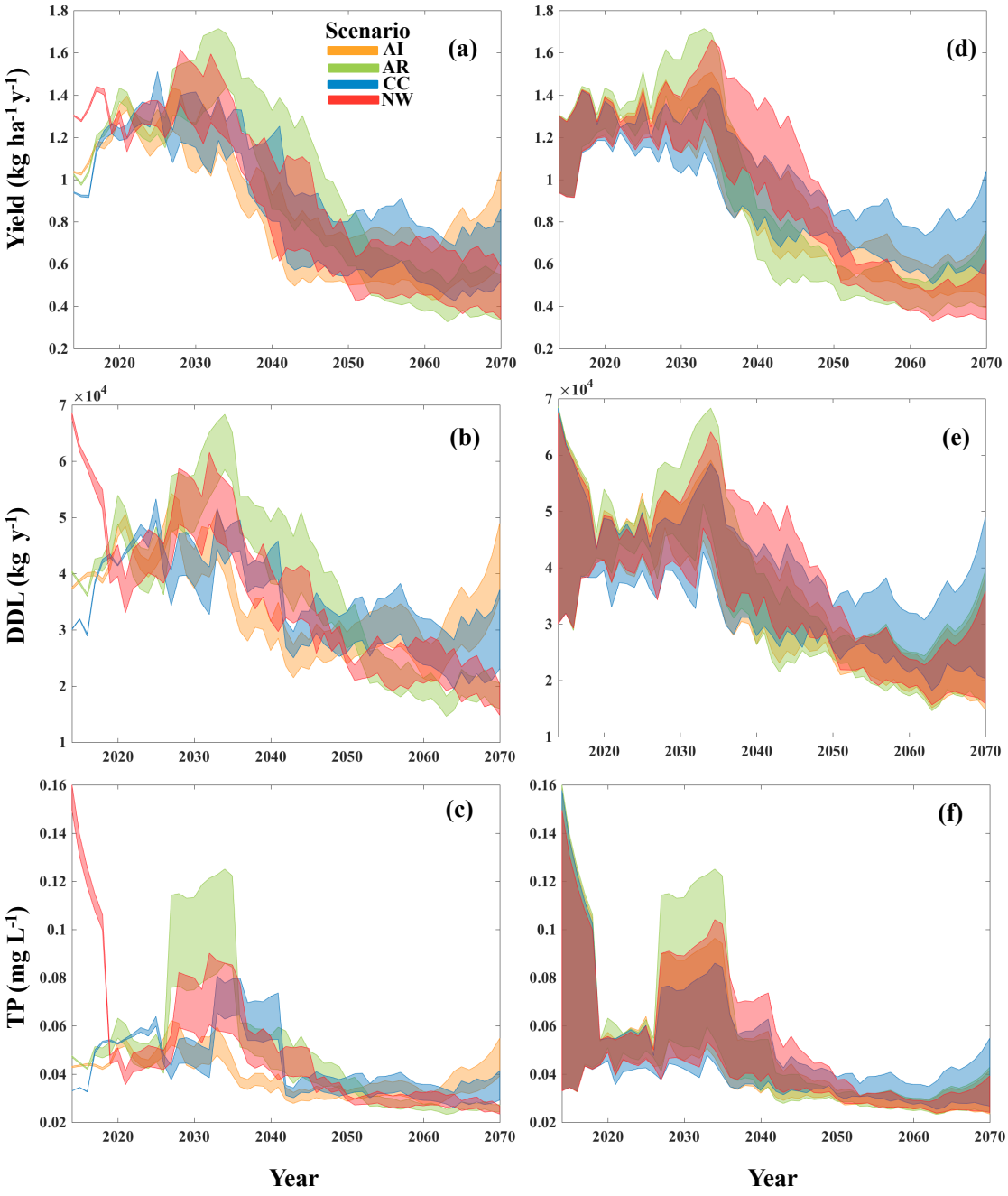


Figure 4. Annual range across the climate (a-c) and LULM scenarios (d-f) for yield, DDL, and TP.

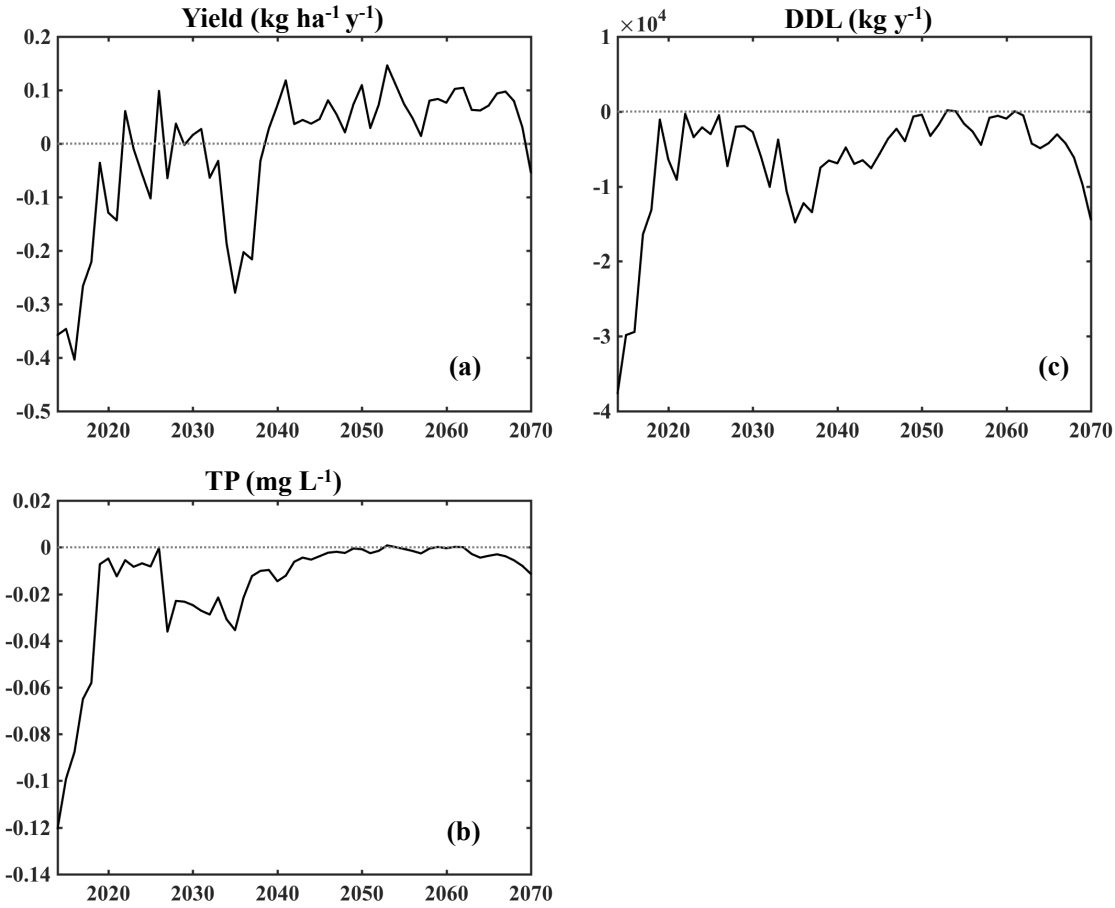


Figure 5. Annual range across LULM minus the annual range across climate. Years with positive (negative) values indicate the range in LULM (climate) was greater.

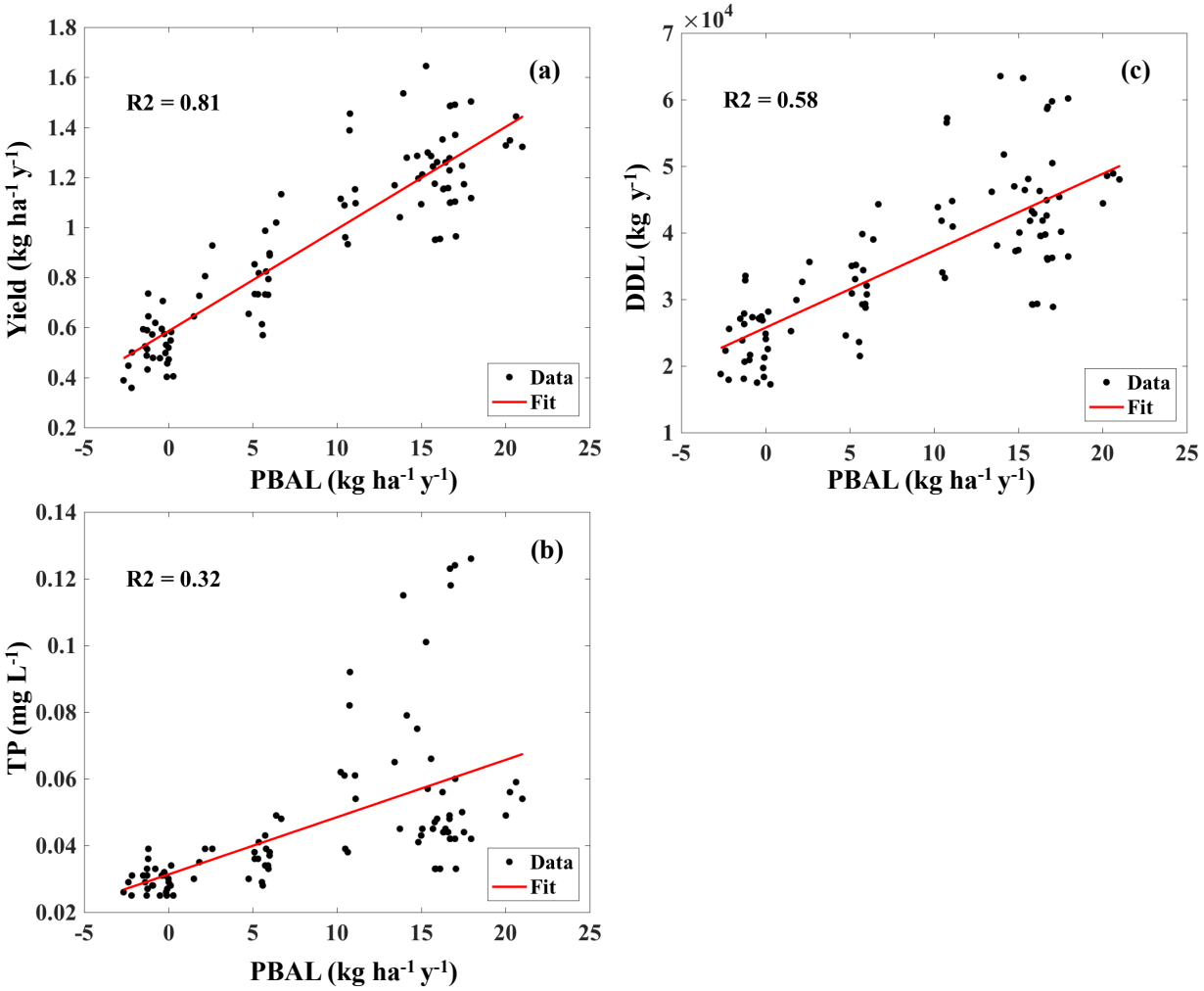


Figure 6. Scatter plots of decadal scale water quality variables versus PBAL (manure P rate + fertilizer P rate – harvest P rate).

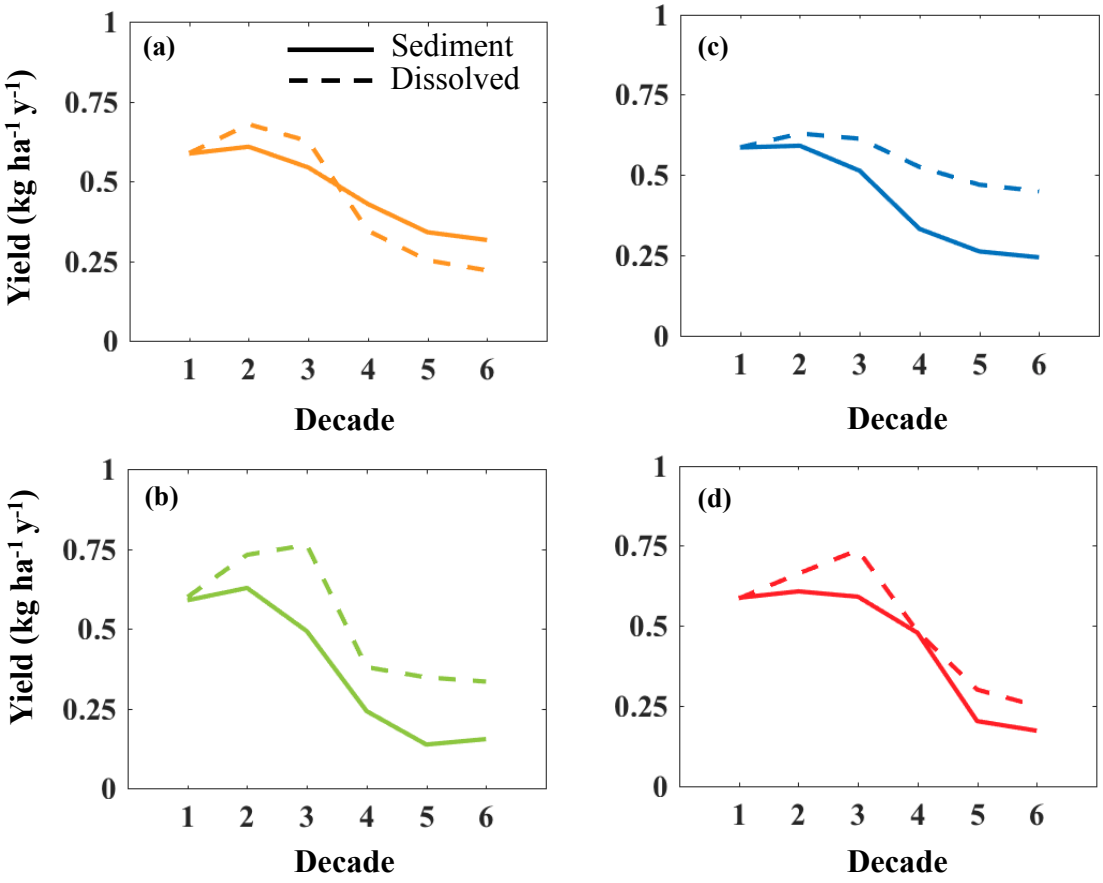


Figure 7. Mean annual P yield per decade for each LULM scenario, AI (a), AR (b), CC (c), and NW (d).

Appendix M. Atmospheric CO_2 estimation for the Yahara 2070 scenarios

For each of the four scenarios, atmospheric CO_2 was assigned the value 391.7 ppm in the year 2010, corresponding with the IPCC Special Report on Emissions Scenarios (SRES). The remainder of each scenario's CO_2 trajectory was based on specific emissions scenarios from the SRES. AI was largely based on the milder B1 emissions scenario, and AR, CC, and NW were all largely based on the A1B scenario, which assumed more rapid economic growth. The CO_2 trajectories were also adjusted according to the severity of climate change in each scenario, each of which was primarily dictated by temperature change. Decadal trends in both temperature and CO_2 are shown in Supp. Table 1.

A recent study showed that previous model parameterizations of plant physiological response to rising CO_2 may be overestimated for major crop types (Twine et al. 2013). To account for this, we performed a simple calibration procedure that generated a more realistic response in vegetation to rising levels of CO_2. For C3 plants specifically, we set targets of a 15% increase in net primary productivity at 550 ppm, and a 22.5% increase in productivity at 700 ppm (Ainsworth and Long, 2005, Bishop et al. 2014, Leakey et al. 2009, Norby et al. 2005). We determined the model parameters for CO_2 inputs and plant physiological response that were required to achieve these responses, and used those parameterizations as part of the final model simulations. Our corresponding productivity increases for C4 crops and grasses were minimal, around 2% for CO_2 concentrations of 550 and 700 ppm.

Chapter 5

Conclusions

5.1 Summary

My dissertation investigated how P cycling and transport within an agricultural watershed is affected by anthropogenic drivers of change including climate, land use, and land management. To conduct this research, I helped develop a comprehensive watershed modeling framework that simulates an array of ecosystem services including surface water quality at multiple spatial and temporal scales. I introduced biogeochemical cycling of P and loss of P in runoff to Agro-IBIS, which enabled simulation of terrestrial P dynamics and landscape contributions of P to aquatic ecosystems. The collaborative development and evaluation of the modeling framework was documented in the Appendices to Chapter 2, and each research chapter of my dissertation has used the modeling framework to investigate a particular aspect of surface water quality.

In Chapter 2, I quantified the impact of legacy P stored in soils and stream sediments on lake water quality. The results indicated that legacy P has a significant and long-lasting effect on water quality in the lakes as well as in runoff and streams. The lower lakes (Monona, Waubesa, and Kegonsa) were less sensitive to changes in legacy P than was Lake Mendota, due to the dominant influence of riverine loads from upstream lakes. In calibrating the historical simulation that was used to spin up Agro-IBIS, I discovered that there is a substantial overabundance of soil P in the watershed, roughly four times more than agronomic recommendations, on average. I also discovered evidence that terrestrial P supply may interact synergistically with extreme

rainfall in affecting lake TP concentration. Overall, this chapter emphasized the important role that past land management has on building up P reserves in soils and sediments, and how that can affect water quality outcomes well into the future.

Based on the clue from Chapter 2 that there might be a possible synergy between rainfall intensity and terrestrial P supply, I investigated the possibility of this more thoroughly in Chapter 3. Using a 2x2 factorial model design of high and low treatments of both rainfall intensity and terrestrial P supply, I found that manure supply on the landscape interacts synergistically with rainfall intensity to affect losses and loading of bioavailable dissolved P. The synergy persisted in total P yield, load, and lake TP concentration in Lake Mendota, indicating that the synergy was relevant at all spatial scales of the watershed. This chapter underscored the importance of manure management in affecting water quality outcomes throughout the watershed, and suggested that surface waters in regions of livestock production may be more vulnerable to increases in rainfall intensity expected with climate change.

Using the Yahara 2070 scenarios, Chapter 4 investigated the role of LULM in affecting long-term water quality outcomes under climate change. Dividing the four core scenarios into their LULM and climate driver counterparts, I examined the relative influence of the four LULM scenarios on water quality under the four contrasting future climates. Overall, LULM was observed to have a substantial effect on water quality at field, stream, and lake scales, with the most important LULM-based driver being the balance of P inputs and outputs at the field scale. The results also indicated that the effect of LULM was most important at the field scale and less important at stream and lake scales, where the effect of climate became increasingly dominant. This chapter showed that while LULM can play an important role in driving water quality outcomes, namely by inducing a negative balance of P inputs and outputs to the soil, the effect of

field scale management practices may be muted in downstream aquatic ecosystems. In addition to reducing the supply of P on the landscape, other management strategies that focus on the aquatic ecosystem themselves may be beneficial, for example in controlling invasive species.

5.2 Synthesis

An overarching message has emerged from my research. It indicates that the overabundance of P within the YW is a dominant biophysical control of surface water quality outcomes across spatial and temporal scales. The second and fourth chapters demonstrated that the effect of this overabundance is significant and long-lasting in the Madison lakes, persistent under changes in climate and across spatial scales, more important than the risk of erosion, and more important than changes in land use/land cover. Without radical intervention, the continued slow release of legacy P will persist in the coming decades, degrading water quality in the lakes and counteracting costly conservation practices and programs.

Increases in extreme precipitation that are expected with climate change already pose a significant threat to water quality. However, in combination with an abundant supply of terrestrial P, the threat of climate change becomes greater. As Chapter 3 demonstrated, the presence of surface-applied manure, ubiquitous throughout fields of the northern Yahara, will be especially vulnerable to increases in intense rainfall events, having the potential to lose even greater amounts of bioavailable P as that intensity increases. Reducing land applications of manure would help mitigate this synergy and may also provide protection from extreme precipitation.

Traditional approaches to soil and water conservation have long focused on retaining P on land and preventing its loss into waterways. Practices like conservation tillage, vegetative buffers, and cover crops provide important benefits within agroecosystems, such as protecting soils from erosion, helping build organic matter, promoting biodiversity, and reducing losses of sediment and sediment P to waterways. However, there is mounting evidence from the literature, the field of conservation, as well as the findings of my research, that these practices may be limited in their ability to holistically address the problem of nonpoint source P pollution. For example, some of these practices do not limit losses of dissolved P but rather may promote them, as in the case of conservation tillage (Bundy et al. 2001). My research underscored the importance of managing for both dissolved and particulate forms of P. Losses of both contribute to eutrophication, yet each are governed by different transport mechanisms that may respond differently to climate and management.

As my research highlighted, in addition to previous studies, there is very little room for error when it comes to containing P on the landscape, since compared with the current amounts of terrestrial P supplied in soils, manures, and fertilizers, only a tiny amount of P lost to waterways is needed to cause eutrophication. As demonstrated by the many combinations of land cover, management, and climate explored in the Yahara 2070 scenarios, constructing a landscape resilient to extreme precipitation will be extremely difficult. Thus, rather than focusing solely on practices that contain P on the landscape and prevent its loss into waterways, my research suggests that a better strategy would be one that focuses on reducing legacy P buildup while also balancing agricultural inputs and outputs of P at the field scale. If average field scale P balance can be substantially reduced, or preferably be reversed (inducing a drawdown of P reserves) as was demonstrated in the Nested Watersheds scenario of Chapter 4, significant improvements in

lake water quality could be attained within 1-2 decades. Traditional conservation practices may play a role in effective watershed management, but new management strategies should harness the strengths of these approaches while also addressing the weaknesses. For example, since conservation tillage is often accompanied by surface P applications that promote increased losses of dissolved P, injection-based or other technologies that help deliver P to deeper soil layers may help curb this unintended consequence (McConnell et al. 2013). Manure injection may also help mitigate the synergistic interaction with extreme rainfall, since the surface supply of manure would be displaced belowground.

An improved strategy for land management in the YW should draw from prior proven strategies, such as those focused on targeting critical source areas (Sharpley et al. 2011, Giri et al. 2016). The targeting approach, which was successfully used in a paired watershed study of the Pecatonica River Watershed of southwestern Wisconsin, is related to the concept of disproportionality. Disproportionality, when applied within a water quality management context, suggests that a small number of sources (farms) may act as outliers that contribute the majority of pollution to waterways (Nowak et al. 2006). Targeting only those sources may be an efficient means of achieving water quality improvements. In the Pecatonica River Watershed, this concept was tested by a collaboration of scientists, conservationists, farmers and other partners. They targeted a small number of farms within the treatment watershed for management interventions that included changes in crop rotations, adoption of conservation or no-till, installation of livestock stream crossings, and changes in manure handling methods. Two years after completing the interventions, they measured a 55% reduction in in-stream P loads during storm events compared with the control watershed. The results represented an important proof-of-concept for the targeting approach.

In light of the overabundance problem, a targeted approach that focuses on drawing down legacy P by inducing a negative P balance on farms may be an effective strategy for the YW. To understand P cycling and water quality impacts at the watershed scale, modeling could provide important information regarding where to target management interventions, and how much of an intervention is required from specific locations to meet water quality goals. On farms, there should be great emphasis placed on avoiding excess P application. Farmers, consultants, and conservationists alike will need to abandon the erroneous assumption of biophysical homogeneity on farms and embrace the need for precision approaches to nutrient management (Nowak 1998). Technological advances will be needed, for example precision manure spreading that takes into account within-field variations in soils, nutrient holding capacity, and current nutrient levels, and distributes manure only where it is needed (Cabot et al. 2006). Manure and fertilizer applications should focus on P in addition to N, since manure is often applied to meet crop N needs which leads to over application of P (Silveira et al. 2010). While precision approaches may improve farmers ability to not over apply fertilizer and manure, it may also exacerbate the problem of excess manure and its disposal in the YW. Thus, multi-farm or sub-watershed scale strategies to handle large amounts of manure will be needed as well, such as manure digesters that are becoming increasingly popular in Wisconsin (Güngör et al. 2008). Other technologies under development that recover P from manure, agricultural residues, and soils should be considered as well (Elser and Bennett 2011, Kahiluoto et al. 2015). And finally, continuous, long-term monitoring of stream and lake water quality will be essential in order to evaluate the effectiveness of installed strategies and any other changes taking place in the watershed (Jordan et al. 2007).

Anticipated challenges to changes in management must be kept in mind, particularly relating to the social and economic realities facing farmers. This is why, as was done in the Pecatonica River Watershed, the enactment of any new management strategies may best be facilitated through collaborations among scientists (social and biophysical), crop consultants, managers, farmers, and people capable of bridging relationships among these participants. In Chapter 5.5 *A Fifth Scenario*, I suggest a possible avenue for engaging scientists, conservationists, farmers, and other stakeholders in an interdisciplinary search for actionable watershed management strategies.

5.3 Broader Contributions

The inability of management interventions to improve water quality in the Yahara and other watersheds has been the topic of recent studies (e.g. Gillon et al. 2016). My research provides additional insight into possible barriers facing management within the YW and elsewhere, since the YW is an exemplar of agricultural and mixed land use watersheds. For example, Chapter 4 showed that P overabundance is the main biophysical factor under local human control that drives water quality throughout all scales of a watershed. Chapter 2 showed that long-term buildup of P within soils and sediments exerts a strong, continuous, and long-lasting effect on lake water quality. These findings will hopefully encourage management within other agricultural watersheds to focus more directly on the problem of P overabundance in soils and sediments, for example through measures that prevent over-application through precision nutrient management (Cabot et al. 2006), encouragement of P crediting based on soil P testing,

and exploration of other means and technologies for recovering or immobilizing P, such as P dredging (Van der Does et al. 1992) or flocculation (Noyma et al. 2016).

As climate change continues to alter hydrological cycles, water quality will be subject to its effects. My research emphasized the important role that climate plays in driving water quality at all scales of a watershed, which suggests to the broader community of scientists and policy makers that mitigating and adapting to global climate change will be imperative in order to protect freshwater resources in the future. As Chapter 4 showed, the effects of local actions will be critical in shaping future trajectories of water quality and other ecosystem services, but the effects of climate may ultimately dominate P transport within waterways. It is therefore essential that the most important aspect of LULM, the terrestrial supply of P, be minimized.

My dissertation made important advancements in watershed modeling. I incorporated the SurPhos model into Agro-IBIS, enabling state-of-the-art simulation of dissolved P loss from soils, fertilizers, and manures. SurPhos is able to capture the important stratification effects of surface applications and the rapid dynamics of P transformations following application. The discovery in Chapter 3 that a synergy exists between extreme precipitation and manure supply was only possible because SurPhos simulates an explicit manure layer on top of the soil, including its interactions with rain and runoff. SurPhos reflects the latest understanding of how dissolved P is lost from agroecosystems, thus the entire watershed modeling community should be encouraged to make use of these advancements.

By incorporating SurPhos into Agro-IBIS, as well as sediment yield and sediment P loss estimation capabilities, Agro-IBIS is the only land surface model to my knowledge containing physically-based representation of water balance (evapotranspiration, infiltration, surface runoff, and drainage), dynamic vegetation growth, and the best dissolved P loss estimation methods

available. My dissertation as a whole demonstrated the power of process-based modeling in examining novel conditions of change.

My participation in developing and testing the modeling framework for WSC has helped enable simulation of the Yahara 2070 scenarios as well as a number of other studies conducted by team members. These include investigations into the tradeoffs and synergies among ecosystem services (Qiu et al. 2017 *in review*, Qiu et al. 2017a *in prep*), the effects of shallow groundwater on ecosystem services (Qiu et al. 2017b *in prep*), the effects of climate and land use on water balance (Zipper et al. 2017 *in prep*), flood risk in the YW (Chen et al. 2017 *in prep*), and methods for estimating final ecosystem service outcomes (Booth et al. 2017 *in prep*).

5.4 Limitations

A key limitation to this research involved a lack of understanding of how diversity in management practices across the watershed (tillage, timing, application method, etc.) affected model results. For example, as discussed in Chapter 2, soil P can vary greatly across agricultural regions, and even across a single field. In order to model the watershed, assumptions were made regarding typical farming practices, including when applications were made[1], how much was incorporated or surface applied[2], and how much P was in each 220-m grid cell at the onset of simulations[3]. It would be valuable for future research efforts using these models to explore how sensitive water quality outcomes are to such assumptions, and whether or not additional

[1] Applications were made three times per year, once in mid-February, once at planting, and again on October 1.
[2] February and October applications were surface applied and the application at planting was tilled to a 10 cm depth.
[3] Initial soil P was determined using typical levels for land cover types in the watershed (Bennett et al. 2004) and were further adjusted during calibration and validation (Motew et al. 2017, Appendix D).

observational datasets (e.g. soil P, nutrient application rates, tillage practices, etc.) would improve model performance.

This research also did not address how the proximity to waterways or spatial arrangement of terrestrial P supply affects water quality. Management strategies that target critical source areas have been successful in some cases, and such an approach might be effective in the YW given the conspicuous presence of large dairy operations upstream of Lake Mendota. In the section *5.5 A Fifth Scenario*, I propose an exercise that integrates a targeted management approach with the biophysical lessons learned from my research as a way to identify effective yet plausible watershed management strategies.

Finally, representation of P cycling in Agro-IBIS had two notable deficiencies that should be addressed in future research. The first involved a lack of P limitation in plant growth. This was deemed a minor deficiency since soils in the YW are in general very high in P, suggesting that plant growth would only rarely be affected by P limitation. Future work that investigates water quality tradeoffs with crop production would however require proper representation of P limitation in plants. Implementing routines for this would be feasible since models of plant P limitation already exist (e.g. Jones et al. 1984).

The other model deficiency in P cycling related to the soil layer from which plants could uptake P. In other P models, such as those included in SurPhos and SWAT, plants are allowed to take up P across multiple soil layers. I attempted to employ a similar strategy when introducing P cycling to Agro-IBIS, however difficulties arose when simulating natural (non-crop) ecosystems. This was due to a problem in capturing organic P transformations correctly within the P module (Chapter 2, Fig. 2). Organic pools play a minor role in agroecosystems yet an important role in natural systems where litter and detritus are important components of P cycling. The organic

pool transformations may be poorly represented since the P models I referenced were designed for agroecosystems (Jones et al. 1984). In order to simulate reasonable soil P dynamics in both crop and non-crop ecosystems as required by the Yahara 2070 simulations, I implemented a solution where plant uptake occurred only from the top soil layer (2.5 cm). I ran independent tests using the APLE model (Vadas et al. 2012) to verify that this solution did not substantially alter simulated estimates of surface soil P or losses of P in runoff. To improve the representation of plant-P dynamics, such as the interplay between root depth and soil P stratification, this part of the model should be improved. Alterations to the equations governing organic P cycling may be required.

5.5 A Fifth Scenario

My research has identified new biophysical aspects of P cycling, transport, and surface water quality that deserve consideration in watershed management. In this section, I propose a collaborative activity that searches for effective yet plausible avenues of land management that combine our best biophysical understanding of the watershed with real world knowledge of farmers, managers, and stakeholders. A Fifth Scenario should seek to address the P overabundance problem, be operationally feasible and cost effective, and help to increase the resilience of both farms and freshwater resources to future changes in climate and other pressures.

Similar to the Yahara 2070 scenario building exercise, the proposed activity would entail an iterative process of modeling and scenario visioning. However, instead of using a fixed number of scenarios that span a wide range of possible outcomes as was done for Yahara 2070,

the proposed activity would focus on finding any number of practical LULM scenarios that improve water quality, with the goal of eventually narrowing down to a single, actionable scenario.

In addition to modeling, the scenario building process would require input from social scientists, policy makers, conservationists, managers, farmers, and other stakeholders that can help navigate the context and constraints that have historically challenged water quality programs in the YW (Wardropper et al. 2016). Participation from those having direct knowledge of farmer realities, such as volatility in milk prices, crop yield variability in response to droughts and other weather events, feed prices, herd health, and other demands, would help guide the scenarios toward solutions that are feasible.

The role of the biophysical scientists in this exercise would be to employ the models (used in my dissertation) to determine how a targeted approach toward land management might be best applied to achieve water quality improvements. The models would be used to identify the prime locations for management, the amount of yield reduction required per location, as well as specific interventions that could be used to achieve those reductions. The other collaborators, having knowledge or experience in farming, conservation, policy, etc., should then work together to envision different pathways of action that could deliver the required biophysical outcomes dictated by the models. For example, if the models indicate that a 50% reduction in P yield is required from all farms upstream of Lake Mendota losing >2 kg ha^{-1} on average, the scenario team should formulate alternative pathways that could achieve those reductions. Solutions should be holistic, spanning issues like cost-sharing and available funding to farmers, to specific farm technologies that can be used in interventions. The scenario team may in turn pose new pathways for the modelers to try, based on preferred avenues of action or economic constraints. Using

modeling and scenario visioning iteratively in this way, an effective yet actionable Fifth Scenario may emerge.

Improving water quality will require the perspectives and skills of many actors within the watershed, not least of all the farmers who initiate the P cycling and transport process through nutrient application and land use practices. Final implementation of a Fifth Scenario should therefore focus on engaging with farmers in a way that recognizes and emphasizes their problem solving capacity and direct knowledge of their land. As a scientist and modeler whose understanding of the watershed is centered on the biophysical aspects of P cycling, I understand that my ability to help improve water quality in the Yahara will depend on meaningful collaborations that produce action. Thus, I look forward to opportunities that bring together all perspectives of agriculture and water quality in an attempt to increase the resilience of both our freshwater resources and agricultural systems, both of which are vital to society.

References

Bennett EM, Carpenter SR, Clayton MK. 2004. Soil phosphorus variability: scale-dependence in an urbanizing agricultural landscape. Landsc Ecol 20:389–400.

Booth EG, Motew M, Chen X, Carpenter SR, Zipper SC, Kucharik CJ, Loheide II SP, Schatz J, Qiu J, Turner M. 2017. Final ecosystem service outcomes from four narrative scenarios: Yahara Watershed, Wisconsin, USA. *In prep.*

Bundy LG, Andraski TW, Powell JM. 2001. Management practice effects on phosphorus losses in runoff in corn production systems. J Environ Qual 30:1822–8.

Cabot PE, Pierce FJ, Nowak P, Karthikeyan KG. 2006. Monitoring and predicting manure application rates using precision conservation technology. J Soil Water Conserv 61:282–92.

Chen X, Motew M, Booth EG, Zipper SC, Loheide II SP, Kucharik CJ. 2017. Impacts of varied lake level management on flood risk to land, property, and people in the Yahara Watershed, Wisconsin, USA. *in prep*

Elser J, Bennett E. 2011. Phosphorus cycle: A broken biogeochemical cycle. Nature 478:29–31.

Gillon S, Booth EG, Rissman AR. 2016. Shifting drivers and static baselines in environmental governance: challenges for improving and proving water quality outcomes. Regional Environ Change 16:759–75.

Giri S, Qiu Z, Prato T, Luo B. 2016. An Integrated Approach for Targeting Critical Source Areas to Control Nonpoint Source Pollution in Watersheds. Water Resour Manage:1–14.

Güngör K, Karthikeyan KG. 2008. Phosphorus forms and extractability in dairy manure: a case study for Wisconsin on-farm anaerobic digesters. Bioresour Technol 99:425–36.

Jones CA, Cole CV, Sharpley AN, R WJ. 1984. A simplified soil and plant phosphorus model: I. Documentation. Soil Sci Soc Am J 48:1–3.

Jordan P, Arnscheidt A, McGrogan H, McCormick S. 2007. Characterising phosphorus transfers in rural catchments using a continuous bank-side analyser. Hydrol Earth Syst Sci Discuss 11:372–81.

Kahiluoto H, Kuisma M, Ketoja E, Salo T, Heikkinen J. 2015. Phosphorus in manure and sewage sludge more recyclable than in soluble inorganic fertilizer. Environ Sci Technol 49:2115–22.

McConnell DA, Ferris CP, Doody DG, Elliott CT, Matthews DI. 2013. Phosphorus losses from low-emission slurry spreading techniques. J Environ Qual 42:446–54.

Motew M, Chen X, Booth EG, Carpenter SR, Pinkas P, Zipper SC, Loheide SP, Donner SD, Tsuruta K, Vadas PA, Kucharik CJ. 2017. The Influence of Legacy P on Lake Water Quality in a Midwestern Agricultural Watershed. Ecosystems In press:1–15.

Nowak P. 1998. Agriculture and change: the promises and pitfalls of precision. Commun Soil Sci Plant Anal 29:1537–41.

Nowak P, Bowen S, Cabot PE. 2006. Disproportionality as a Framework for Linking Social and Biophysical Systems. Soc Nat Resour 19:153–73.

Noyma NP, de Magalhães L, Furtado LL, Mucci M, van Oosterhout F, Huszar VLM, Marinho MM, Lürling M. 2016. Controlling cyanobacterial blooms through effective flocculation and sedimentation with combined use of flocculants and phosphorus adsorbing natural soil and modified clay. Water Res 97:26–38.

Qiu J, Carpenter SR, Booth EG, Motew M, Zipper SC, Kucharik CJ, Chen X, Loheide II SP, Seifert J, Turner M 2017. Scenarios reveal pathways to sustain future ecosystem services in an agricultural landscape. Environ Res Lett. *In review.*

Qiu J, Carpenter SR, Booth EG, Motew M, Zipper SC, Kucharik CJ, Loheide II SP, Turner M 2017a. Understanding relationships among ecosystem services across scales. *In prep*.

Qiu J, Zipper SC, Motew M, Booth EG, Loheide II SP, Kucharik CJ 2017b. Importance of shallow groundwater on sustaining multiple ecosystem services under climate extreme scenarios. *In prep.*

Sharpley AN, Kleinman PJA, Flaten DN, Buda AR. 2011. Critical source area management of agricultural phosphorus: experiences, challenges and opportunities. Water Sci Technol 64:945–52.

Silveira ML, Vendramini JMB, Sollenberger LE. 2010. Phosphorus Management and Water Quality Problems in Grazingland Ecosystems. International Journal of Agronomy 2010:1–8.

Vadas PA, Joern BC, Moore PA. 2012. Simulating soil phosphorus dynamics for a phosphorus loss quantification tool. J Environ Qual 41:1750–7.

Van der Does J, Verstraelen P, Boers P, Van Roestel J, Roijackers R, Moser G. 1992. Lake restoration with and without dredging of phosphorus-enriched upper sediment layers. Hydrobiologia 233:197–210.

Wardropper CB, Chang C, Rissman AR. 2015. Fragmented water quality governance: Constraints to spatial targeting for nutrient reduction in a Midwestern USA watershed. Landsc Urban Plan 137:64–75.

Zipper SC, Motew M, Booth EG, Chen X, Qiu J, Loheide II SP, Kucharik CJ. 2017. Separating climate- and land use-driven changes in the watershed-scale water budget using principal components regression. *In prep*.