Appendix 2

This appendix contains a memorandum from Tetra Tech to Ohio EPA reviewing Soil and Water Assessment Tool (SWAT) models and publications.

TETRA TECH

Memo

To:	Josh Griffin and Paul Gledhill (Ohio EPA)
From:	Bill Carlson and Jon Butcher (Tetra Tech), authors Kevin Kratt (Tetra Tech), reviewer[1]
Date:	June 29, 2022
Subject:	Review of SWAT modeling

This memorandum presents a review of previous Soil and Water Assessment Tool (SWAT) modeling efforts that may be pertinent to the development of the Maumee Watershed Nutrient Total Maximum Daily Load (TMDL).

The Ohio Environmental Protection Agency (EPA) tasked Tetra Tech with identifying findings from previous SWAT modeling efforts that may help guide the agency with development of the implementation strategy for the TMDL. The TMDL and implementation will focus on the Annex 4 load reductions for total phosphorus (TP) and dissolved reactive phosphorus (DRP). As such, Tetra Tech's review focused on key findings for those two forms of phosphorus.

Ohio EPA also directed Tetra Tech to target the review of previous SWAT modeling efforts on the ensemble[2] modeling efforts for the Maumee River watershed. Other SWAT model publications were also reviewed, as were three of Tetra Tech's SWAT models developed to support other Ohio TMDLs in the Lake Erie basin.

To support the Maumee Watershed Nutrient TMDL, Ohio EPA identified two key study questions and several key concepts to explore with the first study question

1. What level of implementation effort is needed to meet the Annex 4 nutrient reduction goals?
    a. What is the role of commercial/inorganic fertilizer versus manure fertilizer?
    b. What is the impact of legacy soil phosphorus?
    c. What are the models' conclusions about meeting TP versus DRP targets?
2. What are the strengths and weaknesses of SWAT modeling for tracking DRP?

This memorandum summarizes Tetra Tech's findings on each of these topics.

[1] Ohio EPA and Tetra Tech would also like to gratefully acknowledge the review comments provided by Margaret Kalcic, Jay Martin, Don Scavia, Anna Apostel, Jeffrey Kast, Rebecca Muenich, and Grey Evenson.

[2] "Ensemble modeling" refers to the groups of four to six SWAT models of the Maumee River watershed that were developed independently by four to six organizations and whose model results were evaluated together. Publications of the results of the ensemble modeling also refer to "multiple models" and "multi-modeling". For convenience, in this memorandum, only the term "ensemble modeling" is used.

1.0 OVERVIEW

To support Ohio EPA with the development of the Maumee Watershed Nutrient TMDL, Tetra Tech reviewed published SWAT models with the objective of identifying findings that may help guide Ohio EPA with development of the implementation strategy for the TMDL. Tetra Tech reviewed four sets of SWAT models or associated publications (Table 1):

- **Evaluation of Ensemble Modeling**: Five publications associated with the ensemble modeling effort were reviewed: Evenson et al. 2021; Martin et al. 2019, 2021; Scavia et al. 2016, 2017. These studies combine and compare results from multiple independently developed models as a way to understand the uncertainty that is introduced by model algorithms, model setup choices, and model calibration. The ensemble modeling efforts are generally categorized as evaluations of critical source areas or as evaluations of best management practices (BMPs). Summaries of key findings for the ensemble modeling efforts are presented in Section 3.0.
- **Individual SWAT Models**: Six publications of additional SWAT models developed for watersheds in the Great Lakes (Christopher et al. 2017; Hanief and Laursen 2019; Her et al. 2016; Merriman et al. 2019; Wang et al. 2017) or Gulf of Mexico (Cho et al. 2010) basins were also reviewed; all of these SWAT models were developed to evaluate BMP scenarios. Summaries of the individual SWAT model publications are presented in Section4.1.
- **Maumee River SWAT Models**: Several teams of researchers developed SWAT models for the Maumee River watershed. Some of these models evolved from models developed for ensemble modeling. Six publications were reviewed: four publications focused upon BMPs and two publications focused on critical source area identification. Summaries of the Maumee River SWAT model publications are presented in Section 4.2.
- **Tetra Tech SWAT Models**: Tetra Tech developed SWAT models for the Black River, Sandusky River, and St. Joseph River watersheds. These models were developed to support TMDL development for impaired waters in the Lake Erie Basin. Simulated flow data were used to develop load duration curves in these TMDL projects; simulated phosphorus data were used to support the source assessment in each TMDL study. The Sandusky River SWAT model was later updated and used to evaluate several BMP scenarios. Summaries of the Tetra Tech SWAT models are presented in Section4.3.

Summaries of the findings of this review are presented in the following subsections that are organized by key study question and key concept. Additional details about SWAT model tracking of DRP, evaluations of the ensemble modeling efforts, and individual SWAT models are presented in Sections 2.0, 3.0, and 4.0, respectively.

Table 1. Summary of SWAT models

Reference	Watershed	Type of modeling
Apostel et al. (2021)	Maumee River (Western Lake Erie Basin)	Comparison of Maumee River watershed SWAT models
Cho et al. (2010)	Little River (Gulf of Mexico Basin)	BMP scenario evaluation
Christopher et al. (2017)	River Raisin (Western Lake Erie Basin)	BMP scenario evaluation
Evenson et al. (2021)	Maumee River (Western Lake Erie Basin)	Ensemble modeling: Critical source areas evaluation
Hanief and Laursen (2019)	Grand River (Eastern Lake Erie Basin)	BMP scenario evaluation
Her et al. (2016)	St. Joseph River (Western Lake Erie Basin)	BMP scenario evaluation
IDEM (2017)	St. Joseph River (Western Lake Erie Basin)	TMDL support
Kalcic et al. (2016)	Maumee River (Western Lake Erie Basin)	BMP scenario evaluation and sensitivity analyses
Kast et al. (2021a)	Maumee River (Western Lake Erie Basin)	BMP scenario evaluation and sensitivity analyses
Kast et al. (2021b)	Maumee River (Western Lake Erie Basin)	BMP scenario evaluation
Martin et al. (2019, 2021)	Maumee River (Western Lake Erie Basin)	Ensemble modeling: BMP scenario evaluation and sensitivity analyses
Merriman et al. (2019)	Upper East River (Lake Michigan Basin)	BMP scenario evaluation
Muenich et al. (2016)	Maumee River (Western Lake Erie Basin)	BMP scenario evaluation
Ohio EPA (2014)	Sandusky River (Western Lake Erie Basin)	TMDL support
Ohio EPA (2022)	Black River (Central Lake Erie Basin)	TMDL support
Scavia et al. (2016, 2017)	Maumee River (Western Lake Erie Basin)	Ensemble modeling: BMP scenario evaluation and sensitivity analyses
Wang et al. (2017)	Gully Creek (Lake Huron Basin)	BMP scenario evaluation
U.S. EPA (2017)	Sandusky River (Western Lake Erie Basin)	BMP scenario evaluation

1.1 ANNEX 4 PHOSPHORUS TARGETS

The Maumee Watershed Nutrient TMDL will address phosphorus-loading from the Maumee River watershed that causes harmful algal blooms in the Western Basin of Lake Erie and contributes to hypoxia in the Central Basin of Lake Erie. Ohio EPA is developing the TMDL consistent with the Annex 4 phosphorus targets.

Work under the U.S.-Canada Great Lakes Water Quality Agreement (GLWQA) – Annex 4 led to the establishment of binational phosphorus load reduction goals for the western and central basins of Lake Erie. The Annex 4 phosphorus reduction targets for Lake Erie are as follows (GLWQA 2015):

- **To minimize the extent of hypoxic zones in the waters of the central basin of Lake Erie:** A 40 percent reduction in annual TP entering the western and central basins of Lake Erie—from the United States and from Canada—to achieve an annual load of 6,000 metric tons to the central basin. This amounts to a reduction from the United States and Canada of 3,316 metric tons and 212 metric tons respectively.
- **To maintain algal species consistent with healthy aquatic ecosystems in the nearshore waters of the western and central basins of Lake Erie:** A 40 percent reduction in spring (March 1 through July 31) TP and DRP loads from the following watersheds where algae is a localized problem: in Canada, Thames River and Leamington tributaries; and in the United States, Maumee River, River Raisin, Portage River, Toussaint Creek, Sandusky River and Huron River (Ohio).
- **To maintain cyanobacteria biomass at levels that do not produce concentrations of toxins that pose a threat to human or ecosystem health in the waters of the western basin of Lake Erie:** A 40 percent reduction in spring (March 1 through July 31) TP and DRP loads from the Maumee River in the United States.

These targets were formally adopted by the United States and Canada in February 2016. Each of the affected States (Indiana, Ohio, and Michigan) developed domestic action plans that describe how the 40 percent reduction goals will be met. To account for years with very high spring precipitation or streamflow, the states, provinces, and countries seek to meet the Annex 4 targets in nine of ten years.

1.2 LEVEL OF IMPLEMENTATION EFFORT

The Maumee Watershed Nutrient TMDL will require phosphorus reductions from nonpoint sources in the Maumee River watershed because the vast majority of phosphorus is derived from nonpoint sources. As the land use in the watershed is predominantly agriculture, a key focus of nonpoint source phosphorus reduction will target agricultural runoff. In recent years, multiple organizations have developed SWAT models to evaluate watershed-scale phosphorus loading. Several SWAT model studies included development of scenarios to predict phosphorus reductions from various levels of implementation of agricultural BMPs. Ohio EPA seeks to estimate the level of implementation of agricultural BMPs that may be necessary to achieve Annex 4 phosphorus targets.

Recent ensemble modeling sought to determine the levels of implementation effort necessary across the Maumee River watershed to meet the Annex 4 targets. With the ensemble modeling efforts, generally, model scenario results did not meet Annex 4 targets for TP and DRP loads and flow-weighted mean concentrations

(FWMC), when averaging results across the 5 models for 2005-2014. To achieve Annex 4 targets, suites of BMPs will need to be adopted at higher implementation rates than were simulated in the models.

Tetra Tech categorized pertinent sensitivity analyses and BMP scenarios from the ensemble modeling effort into four groups (listed below). With regards to the sensitivity analyses, Ohio EPA is not considering management actions to eliminate all permitted point sources, to eliminate all manure land application, to eliminate all combined sewer overflows (CSOs)[3], or to require conversion of cropland to grassland.

- **Sensitivity Analyses**: The objective of these model simulations is to determine how sensitive model results are to various sources of TP and DRP. These analyses test increasing or decreasing various levels of certain sources. However, the sensitivity analyses are not meant to represent feasible "real-world" management actions.
    - ***Elimination of Sources***: Elimination of all point sources was simulated as a sensitivity analysis in two studies, elimination of all manure land application was simulated as a sensitivity analysis in one study, and one study simulated CSOs while another did not (thus allowing for evaluation of the elimination of CSOs). The point sources elimination analysis results in only small decreases in TP and DRP loads reaching Lake Erie. The manure land application analysis also resulted in small decreases in TP and DRP loads. The evaluation of models that did and did not simulate CSOs indicate no significant difference. With regards to CSOs, the two models used different sets of point sources inputs and the models were separately calibrated; thus, the CSO evaluation was an indirect comparison. Therefore, the elimination of all point sources, the elimination of all manure land application, and the elimination of all CSOs are each estimated to be insufficient in and of themselves to meet the Annex 4 targets (Martin et al. 2019, 2021).
    - ***Conversion of Cropland to Grassland***: Conversion of 10%, 25%, or 50% of cropland to grassland (targeting cropland with lowest crop yields and highest TP losses)[4] considerably reduced TP loads but did not reduce DRP loads as much (Scavia et al. 2016). These results indicate that without implementation of other BMPs, a quarter to one half of cropland would need to be taken out of production and converted to grassland to achieve Annex 4 TP and DRP targets (Scavia et al. 2017).
- **BMP Implementation Scenarios**: The objective of these model simulations is to evaluate certain suites of BMPs that could be implemented in an effort to develop a strategy to meet the Annex 4

---

[3] CSO communities must comply with their National Pollutant Discharge Elimination System permits, long-term control plans, and consent decrees or orders.

[4] Grassland was simulated as Shawnee switchgrass (*Panicum virgatum*) that was harvested in October each year. No phosphorus fertilizer was applied, but a once per year nitrogen fertilizer (56 kilograms nitrogen per hectare) was applied. The grassland was managed for wildlife habitat, with no livestock grazing. The baseline discussion says that tiling is included for the agricultural HRUs but there is no discussion of whether tiling was included in the scenario grassland-harvest HRU.

targets. Many of these BMP scenarios are intended to represent feasible “real-world” management actions.

- ***Implementation of In-Field Agricultural BMPs***: Several scenarios across the ensemble modeling efforts (Scavia et al. 2016, 2017; Martin et al. 2019, 2021) suggested that subsurface fertilizer placement and cover crops (aside from winter wheat) had the largest phosphorus reductions. Altering the timing of fertilizer application and changing all application from broadcast to incorporation had inconsistent effects on phosphorus losses. None of the scenarios of in-field agricultural BMPs met Annex 4 targets for both TP and DRP loads averaged over 2005-2014 (Scavia et al. 2016, 2017; Martin et al. 2019, 2021), nor were the FWMC targets met in 9 of 10 years (Martin et al. 2019, 2021).
- ***Implementation of In-Field and Edge-of-Field Agricultural BMPs***: Several scenarios across the two sets of ensemble modeling efforts (Scavia et al. 2016, 2017; Martin et al. 2019, 2021) indicated that buffer strips and wetlands targeted to areas with the greatest TP losses resulted in considerable reductions in TP loads but much less relative reduction with DRP loads. All but one scenario failed to achieve the Annex 4 targets. Note that Scavia et al. (2016, 2017) evaluated the 10-year (2005-2014) average loads with the Annex 4 target loads, while Martin et al. (2019, 2021) compared each of the 10 individual annual loads to the Annex 4 target loads and evaluated the 9-of-10 years goal.

  The *Targeted Series of Practices* (#8) scenario in Scavia et al. (2016, 2017) met both load targets, when averaged over 2005-2014, but it only barely met the DRP load target. This scenario was composed of subsurface application of phosphorus fertilizer, cereal rye cover crop in years without wheat, and medium-quality buffers; BMPs were targeted to the 50% of row cropland with the highest TP losses. *Random Series of Practices* (#9), with distribution to random cropland, did not meet the DRP target.

Evaluation of ensemble modeling of BMP implementation scenario results generally indicated that achievement of the Annex 4 targets in 9-of-10 years will be difficult even with widespread adoption of a suite of BMPs. While BMP scenarios for certain individual models resulted in the Annex 4 targets being met in 9-of-10 years, the average of model results for each BMP scenario never resulted in the Annex 4 TP load target being met in 9-of-10 years (Martin et al. 2019, 2021). None of the model scenarios averages of results met the Annex 4 target for DRP load in 9-of-10 years (Martin et al. 2019, 2021).

In the ensemble modeling evaluations, several BMP scenarios simulated different implementation rates of various in-field and edge-of-field suites of BMPs. A few scenarios were estimated to lead to Annex 4 target achievement; “[h]owever, all the successful pathways require broad implementation of both common and less common practices” (Scavia et al. 2017, p. 129). The authors further describe this effort as being a “daunting task and will require extensive changes in management and much greater investment of resources to achieve the required levels of implementation” (Scavia et al. 2017, p. 130).

Individual SWAT model publications also indicated that high levels of implementation are needed to significantly reduce phosphorus loads. In the Upper East River (Lake Michigan Basin), thousands of individual

BMPs would need to be implemented to appreciably reduce TP and DRP loading to the watershed (Merriman et al. 2019). In the Gully Creek watershed (Lake Huron Basin), watershed-scale reductions predicted for existing and planned BMPs were small because the total area of the existing and planned BMPs was small relative to the area of the watershed (Wang et al. 2017). In the River Raisin watershed (Western Lake Erie Basin), scenarios of 25%, 50%, and 100% implementation of two-stage ditches on headwaters streams reduced annual TP loads for the River Raisin watershed by 12%, 20%, and 31% (Christopher et al. 2017); 100% implementation of two-stage ditches on headwaters streams is not alone sufficient to meet the Annex 4 target of a 40% annual TP reduction. In the St. Joseph River watershed (Western Lake Erie Basin), "tangible improvement in water quality at a watershed scale would require wide implementation and appropriate placement of conservation practices" (Her et al. 2016, p. 261. Finally, in the Maumee River watershed, researchers concluded that the rates of implementation considered feasible for stakeholders (25%-33%) would not achieve Annex 4 phosphorus load goals and that significantly higher rates of implementation and targeting of critical DRP source areas are needed (Kalcic et al. 2016).

1.3 BMP-TARGETING

Ohio EPA did not identify BMP-targeting as one of the topics to be addressed in this review. However, ensemble and individual SWAT modeling efforts indicate that targeting BMPs to source areas with high phosphorus load delivery to streams yields greater reductions than randomly distributing BMPs. However, solely targeting BMP implementation to areas of high phosphorus loading may be unrealistic because multiple factors affect if and where farmers may install BMPs (Kast et al. 2021b); instead, BMP placement should target areas of high phosphorus loading with willing landowners.

In the ensemble modeling studies, targeting implementation of cover crops, riparian buffers, and subsurface fertilizer placement to existing areas with the highest phosphorus loads (versus random distribution) resulted in Annex 4 targets being met for more years (Martin et al. 2020; Scavia et al. 2017)[5].

The authors for the individual SWAT models came to similar conclusions. In the Little River experimental watershed, "the study showed that using subwatershed nonpoint source pollutant load as a prioritization criterion resulted in the most rapid water quality improvement" (Cho et al. 2010, p. 463). In the Grand River watershed (Eastern Lake Erie Basin), the authors concluded that multiple BMPs need to be implemented in high yield subwatersheds to achieve the Annex 4 target reductions (Hanief and Laursen 2019, p. 173). In the St. Joseph River watershed (Western Lake Erie Basin), the authors found that the highest performing BMPs were not widely implemented and BMPs were not implemented in the areas where they would do the most benefit (Her et al. 2016).

[5] Scavia et al (2016, 2017) and Martin et al. (2019, 2021) evaluated model results with Annex 5 targets differently.

1.4 ROLE OF COMMERCIAL/INORGANIC FERTILIZER VERSUS MANURE FERTILIZER

SWAT models of the Maumee River watershed, which had been developed by teams of researchers that also participated in ensemble SWAT modeling effort, were used to evaluate sources of phosphorus, including inorganic phosphorus fertilizer and manure application. During the March through July period, the baseline model indicated that inorganic fertilizer was the largest source of DRP (58%) and second largest source of TP (42%), while manure application was the third largest source of DRP (12%) and TP (8%)(Kast et al. 2021a). These results generally indicate that manure application contributes about as much phosphorus load as point sources, and both such sources are significantly smaller than contributions from current inorganic phosphorus fertilizer application and legacy soil phosphorus. These results are also consistent with results from the ensemble modeling efforts (Martin et al. 2019, 2021).

SWAT model scenarios developed before (Muenich et al. 2016), during (Martin et al. 2019, 2021), and after (Kast et al. 2021a) the ensemble modeling SWAT studies included sensitivity analyses that reduced or eliminated manure application and/or inorganic phosphorus fertilizer application. All of the SWAT modeling efforts suggest that load contributions from commercial fertilizers or manure increase under wetter conditions (e.g., higher spring precipitation, higher spring streamflow) due to both increased soil erosion and increased dissolved phase transport. Results also showed that modifying or eliminating inorganic fertilizer and manure application are not additive; other environmental factors affect phosphorus dynamics in surface runoff, tile drainage, and in-stream flow.

These SWAT modeling efforts showed that elimination of all manure land application will not achieve Annex 4 targets (Kast et al. 2021a; Martin et al. 2019, 2021). Loads contributed by manure land application are considerably smaller than loads contributed by inorganic phosphorus fertilizers and legacy soil phosphorus. The ensemble modeling and individual modeling efforts for the Maumee River watershed used sensitivity analyses and the authors acknowledged that the wholesale elimination of inorganic phosphorus fertilizers or manure application is not economically feasible. Small, un-regulated to large, permitted livestock operations that produce considerable manure are located throughout the Maumee River watershed.

Other SWAT models (Cho et al. 2010; Hanief and Laursen 2019) included phosphorus fertilizer reduction scenarios without specifying what type of fertilizer was being reduced. As such, they shed no light upon the roles of inorganic phosphorus fertilizer and manure application.

1.5 IMPACT OF LEGACY SOIL PHOSPHORUS

Legacy soil phosphorus is a key factor that must be considered in developing BMPs to meet Annex 4 targets. SWAT modeling indicates that the Maumee River watershed soils contain significant legacy soil phosphorus (Muenich et al. 2016). Model results showed that legacy soil phosphorus (45%) contributed considerably to watershed TP loading, as compared with inorganic fertilizers (42%) and other sources (Kast et al. 2021a). The model scenario results suggested that corn and soybean yields would not be affected by the cessation of all phosphorus fertilizer application for about 25 years (Muenich et al. 2016), although the result is dependent on the SWAT 2012 code representation of soil P availability to crops. TP and DRP loads do decrease over time with cessation of phosphorus fertilizer application, but the response time is relatively slow due to the legacy soil phosphorus (see Section 4.2.2).

Various climactic conditions were simulated in SWAT for 80-year periods using different fertilizer scenarios and sensitivity analyses: *business as usual*, *elimination of all fertilizers*, *elimination of phosphorus fertilizers*, and *elimination of inorganic phosphorus fertilizers*. Under dry conditions, Annex 4 targets can always be met under any fertilizer scenario. Under average conditions, Annex 4 targets can be met after only a few years; however, corn yields are immediately affected by the elimination of nitrogen fertilizer in the *elimination of all fertilizer* scenario. After about 25 years, corn and soybean crop yields are significantly affected by the lack of phosphorus fertilizer.

Under high spring precipitation and high spring streamflow conditions, Annex 4 DRP load targets can be met in about 25 years without manure application (*elimination of all fertilizer* and *elimination of phosphorus fertilizer*) and in about 45-50 years with manure application (*elimination of inorganic phosphorus fertilizer*). Under higher spring precipitation and higher spring streamflow conditions, Annex 4 TP load targets are never met. Without manure application (*elimination of all fertilizer* and *elimination of phosphorus fertilizer*), under higher spring precipitation conditions, Annex 4 TP load targets are almost met after about 45 years. Thus, in higher spring precipitation and streamflow conditions, without phosphorus fertilizer application, Annex 4 targets will not be met due in part to phosphorus losses from legacy soil phosphorus. Other phosphorus sources (e.g., in-stream erosion) would also contribute.

The large amount of legacy soil phosphorus will need to be addressed during BMP implementation. At present, the spatial distribution of higher and lower concentrations of legacy soil phosphorus are unknown (Kast et al. 2021a). The effects of commercial and manure fertilizer application in areas of higher versus lower legacy soil phosphorus are not known and could be important. Current phosphorus fertilizer application timing is also an important consideration, especially with DRP (Muenich et al. 2016). Implementation strategies will need to consider climate, legacy soil phosphorus, and current fertilizer practices and will need to implement an array of BMPs to address the different species and sources of phosphorus.

1.6 CONCLUSIONS ABOUT MEETING TP VERSUS DRP TARGETS

Achieving the DRP target will likely be more difficult than achieving the TP target. There are, however, concerns regarding the SWAT model's ability to represent DRP loading through tile drains that may limit the accuracy of model applications (see Section 2.0).

The two sets of ensemble modeling evaluated results differently. In Scavia et al. (2016, 2017), the ensemble model results (averaged across five models) were averaged over 2005-2015, and these 10-year averages were compared with the Annex 4 load targets. In Martin et al. (2019, 2021), the ensemble model results (averaged across five models) for each individual year were compared with the Annex 4 load targets and evaluated with the 9-of-10 years goal.

With the ensemble modeling effort, Scavia et al. (2016) found that the TP target (as an average across 2005-2014) was achieved in 5 of 13 scenarios, whereas the DRP target was achieved in only 3 scenarios and just barely at that. Similarly, Martin et al. (2020, p. 7) concluded that "implementation of most scenarios would make less progress in meeting the DRP loading targets compared to TP". The authors further concluded that "the agreement among model results provide confidence in the finding that the bundled-management scenarios are more likely to reach the TP target than the DRP target" (Martin et al. 2020, p. 6).

Researchers at the University of Michigan concluded that there "may be trade-offs in meeting multiple targets," for example, time of phosphorus fertilizer application influences the timing of DRP loading but did not influence TP loading to meet the annual target for hypoxia (Kalcic et al. 2016, p. 8138). Additionally, some BMPs are predicted to slightly increase DRP losses, while reducing TP losses. For example, "[w]inter cover crops held back nutrient runoff during the winter months, and reduced TP loading considerably throughout most of the year. … However, nutrients stored in the cover crop were released after the crop was killed in the spring, providing higher P at the soil surface available for export in the late spring and summer" (Kalcic et al. 2016, p. 8142).

2.0 SWAT MODEL TRACKING OF DRP

A key issue regarding the use of SWAT to simulate conditions in the Maumee River is the representation of DRP transport from the soil surface and soil matrix into tile drains. SWAT simulates phosphorus in organic and mineral forms. Mineral phosphorus may be present in stable (non-bioavailable) sorbed form, active sorbed form, or in solution. The latter form is DRP. Sorbed phosphorus may be transported with sediment, and DRP with flow processes. Transformations between DRP and mineral forms are represented by equilibrium partitioning. DRP can leach from the surface layer into the subsurface layer. However, earlier versions of the SWAT code (e.g., SWAT 2012) do not route DRP from the soil into tile drains, as was noted (and remedied) by Kalcic et al. (2016). Additionally, DRP from the soil surface can be routed into tile drains if the crack flow routine is turned on, and this option was used in several of the studies referenced in this memorandum. Further, DRP can be associated with groundwater discharge in the SWAT model, but the concentration is a user input and is not estimated on a mass-balance basis from soil P. The following sections summarize the evolution of how the SWAT model has simulated DRP transport through tile drains.

Transport from the soil surface can occur both via leaching through the soil matrix and by drainage of ponded water on the surface especially via soil macropores that connect to tile drains. Williams et al. (2016) showed that no-till agriculture tended to increase DRP loss through tile drains, while incorporating surface-applied phosphorus fertilizers reduces loss, suggesting that communication via macropores may be the dominant process. Radcliffe et al. (2015) provided a review of models to predict phosphorus losses in drained fields, including SWAT 2012 and the related APEX model, and found the performance of all models to be lacking, in particular because they did not fully account for the transport of dissolved phosphorus in pooled water on the land surface to tile drains via macropores. Lu et al. (2015) present an extension to SWAT 2012 called DrainP, developed in Denmark. This version was modified to predict DRP in multiple soil layers with improved, but still not impressive representation of DRP in tile flow. They also note that lack of a proper macropore routine appears to be a significant problem.

Older versions of SWAT (e.g., SWAT 2005 and earlier) approximated tile drainage as lateral flow in the soil. Versions of SWAT beginning in 2006, and including the SWAT 2009 and 2012 full releases, incorporate the Hooghoudt equations for tile flow used in the DRAINMOD model, in which flow into tiles is simulated as occurring when the perched water table rises to the invert elevation of the tiles. These equations were originally designed to estimate the time to drain to field capacity under assumptions of a uniform soil layer. They do not perform well where significant drainage to tiles is via macropores that connect to the surface. In addition, SWAT's daily curve number approach to simulating surface runoff means that SWAT's estimates of surface ponding, which is what drives the flow into the macropores, are likely to be unreliable.

SWAT 2012 simulates flux of nitrate from the soil into tile drains, but not DRP entry into tiles unless a code modification, such as presented in Kalcic et al. (2016) is applied. SWAT 2012 can also be used to simulate DRP entry to tiles through macropores, as further discussed in the next paragraph. Tetra Tech has also made modifications to the SWAT code to allow leaching of DRP into tile drains by assuming that the concentration potentially entering tile drains is equal to the leaching mass flux of P divided by the water percolating out of soil layer 1. Additionally, the current version of SWAT, SWAT+, routes DRP to tile drains through matrix flow in the soil layer having the tile, but does not simulate DRP through macropore flow.

Macropores are typically formed by decaying roots and burrowing animals. SWAT 2012 includes a provision for crack flow in vertisols, which are soils with a propensity to crack when dry, opening up large channels. Crack flow simulation in SWAT directly introduces water to subsurface soil layers but is not directly connected to the tile drain simulation. Vertisols are common in only some locations and not often included in SWAT models, although they were used in many of the models reviewed in this report. It has been suggested that crack/bypass flow can be used to represent the influence of macropores in regular soils; however, crack volume is simulated as inversely related to soil moisture, whereas macropores are present at all times. In sum, the SWAT approximation of macropores via the crack flow routine seems suboptimal and is likely to distort the partitioning of flow between surface and subsurface runoff, which also may adversely affect the soil erosion simulations.

Because of the lack of accurate macropore simulation, especially during wet soil conditions, it is possible that SWAT model simulations will underestimate the loading of DRP from tile-drained fields. The effects may differ depending on the type and placement of fertilization and the resulting amount of DRP that may be present in ponded surface water. If the introduction of DRP through macropores is underestimated this could in turn result in an over-estimation of the benefits of BMPs such as filter strips and field buffers that intercept surface runoff but not tile discharge.

3.0 ENSEMBLE MODELING EVALUATIONS

Ensemble-modeling of the Maumee River watershed has evaluated four to six separate SWAT models (Table 2). The models were independently developed, and use different assumptions (e.g., definition of hydrologic response units [HRUs], fertilizer application).[6] Summaries of model algorithms, model inputs, and spatial discretization are presented in Table 3, while summaries of land management operations and initial soil phosphorus levels are presented in Table 4. These models were developed using the same weather and point source data; however, the calibration periods varied by model. The models in the ensembles are also all affected by the representation of DRP transport discussed in Section 2.0.

The models' water quality simulations were calibrated and validated using streamflow data from the U.S. Geological Survey (USGS) gage on the Maumee River at Waterville (04193500) and using water chemistry data from the National Center for Water Quality Research (NCWQR) at Heidelberg University that were collected at the same location (i.e., gage 04193500). The SWAT model developed by the Ohio State University was also calibrated using data collected by USGS and NCWQR at the gages on the Tiffin and Blanchard rivers. Model calibration periods varied but the models were validated for 2005-2014.

Model performance was evaluated using multiple metrics, including representations of crop yields, soil nutrient content, hydrology, and water quality. Model performance for each model was deemed reasonable (Scavia et al. 2016, 2017). Model performance is presented in Table 5 and Table 6.

Sets of scenarios were developed for two modeling efforts that were summarized in four of the publications: Scavia et al. (2016, 2017) and Martin et al. (2019, 2021). For these models, results for each scenario from each model were compared. Key findings from scenario results are presented in Section 3.1.

One study used ensemble modeling to evaluate critical source areas (CSAs): Evenson et al. (2021). Key findings from the CSA predictions are presented in Section 3.2.

[6] Tabular comparisons of the inputs and assumptions of the SWAT models were presented in Appendix A2 by Scavia et al. (2016) and Appendix 1 by Martin et al. (2019).

Table 2. Summary of SWAT models in the ensemble modeling evaluations

Reference	Heidelberg University	LimnoTech	Ohio State University	Texas A&M University	University of Michigan	University of Toledo
Scavia et al. (2016, 2017) [a]	✓	✓	✓	✓	✓	--
Martin et al. (2019, 2021)	✓	✓	✓	--	✓	✓
Evenson et al. (2021) [a]	✓	✓	✓	--	✓	--

Notes:
Model inputs were changed and models were re-calibrated between Scavia et al. (2016, 2017) and Martin et al. (2019, 2021).
a. Scavia et al. (2016) and Evenson (2021) included a SPARROW model developed by USGS. SPARROW does not predict DRP.

Table 3. Summary of SWAT model development between Scavia et al. (2017) and Martin et al. (2021)

Model decision	Options	HU	LT	OSU	TAMU	UM	UT
Model/Sub-Model Algorithms							
SWAT code version	Scavia et al. (2017) Martin et al. (2021)	637 [a] 645	635 [a] 627	635 [a] 645	635 [a] --	635 [a] 635 [a]	-- 664
Tile drain routine	Old (SWAT_DRAIN) New (SWAT_HKdc)	-- ✓	-- ✓	-- ✓	✓ --	-- ✓	✓ --
Water table routine	Old New	-- ✓	✓ --	-- ✓	✓ --	-- ✓	✓ --
In-stream processes	On (QUAL2E) On - modified	✓ --	✓ --	✓ --	✓ --	✓ --	✓ --
Soil phosphorus model	Old New	✓ --	✓ --	-- ✓	✓ --	-- ✓	✓ --
Evapotranspiration method	Penman-Monteith Hargreaves	✓ --	-- ✓	✓ --	✓ --	✓ --	✓ --
Model Inputs							
Land use	NLCD CDL (2007-2012)	-- ✓	2001	-- ✓	-- ✓ [b]	2006 --	2001 ✓
Soils	SSURGO STATSGO	✓ --	-- ✓	✓ --	✓ --	✓ --	✓ --
Soil							
Crack flow	On (crack flow) Off (no crack flow)	n/a n/a	-- ✓	-- ✓	n/a n/a	✓ --	n/a n/a

Based upon: Martin et al. 2019 (Table A-1 in Appendix A) and Scavia et al. 2017 (Table A.2.1 in Appendix A2). When values were in conflict between publications, Martin et al. 2019 is shown in this table.

Model developers: HU = Heidelberg University; LT = LimnoTech; OSU = Ohio State University; TAMU = Texas A&M University; UM = University of Michigan; UT = University of Toledo.

Notes

CDL = Cropland Data Layer; HRU = hydrologic response unit; n/a = not available; NED = National Elevation Dataset; NLCD = National Land Cover Database; n/a = not available; SSURGO = Soil Survey Geographic Database; STATSGO = State Soil Geographic dataset;.

a. Revision 635 or 637 modified to allow more accurate flow of soluble phosphorus through subsurface drains.

b. CDL 2010-2011.

Table 4. Summary of SWAT models development – Land Management Operations and Soil Phosphorus

Model decision	Options	HU	LT	OSU	TA&MU	UM	UT
Spatial Discretization							
HRUs threshold	LU/Soil/Slope	50/25/0	5/10/0	5/20/30	5/10/0	0/10/0	5/10/0
Subbasins	Number of	374	203	1,482	391	358	97
HRU area	Average hectares	12,677	n/a	1,130	72	169	7,700
Land Management Operations							
Fertilizer application	County sales (2002) Tri-State Standards Ag Census Fert. Use [a]	-- ✓ --	-- ✓ --	-- ✓ --	-- -- ✓	-- ✓ --	✓ ✓ --
Manure application	Ag Census Fert. Use [a] County livestock Not included	-- -- ✓	-- ✓ --	-- ✓ --	-- -- ✓	-- ✓ --	✓ ✓ --
Crop rotations [b]	CS CSS CSW CWS CSWCSSW CSWH SS CC	CS CSS CSW CWS CSWCSSW -- SS CC	CS -- -- CWS -- -- SS CC	CS -- CSW CWS -- CSWH SS --	CS -- CSW -- -- -- SS CC	CS CSS -- -- CSWCSSW -- -- --	CS CSS CSW CWS -- -- -- CC
Areas with subsurface drainage	SP, P, or VP drained [c] HSG C or D <1% slope [e]	-- -- ✓	-- ✓ --	✓ -- --	-- -- ✓	✓ -- --	✓ -- --
Soil Phosphorus [f]							
Initial humic organic phosphorus	50-250 mg/kg	94.906 202	D D	D D	D --	D D	-- D
Initial labile phosphorus	5-100 mg labile phosphorus / kg soil	7.002 0.7	D D	10 0.5	34 --	1 1	-- D

Based upon: Martin et al. 2019 (Table A-1 in Appendix A) and Scavia et al. 2017 (Table A.2.1 in Appendix A2). When values were in conflict between publications, Martin et al. 2019 is shown in this table.

Model developers: HU = Heidelberg University; LT = LimnoTech; OSU = Ohio State University; TAMU = Texas A&M University; UM = University of Michigan; UT = University of Toledo.

Notes

D = default value; HSG = hydrologic soil group; kg = kilogram; LU = land use; mg = milligram.

a. Census of Agriculture yield data and fertilizer use Data (1990-2010)

b. C = corn; H = hay; S = soybean; W = winter wheat.

c. Agricultural land on somewhat poorly, poorly, or very poorly drained soils

d. Agricultural land or hay on hydrologic group C or D soils.

e. Agricultural lands with less than 1% slope.

f. For both soil phosphorus parameters, the top value is from Scavia et al. (2017) and the bottom value is for Martin et al. (2019). 'D' represents the default value.

Table 5. Summary of SWAT models development – Validation in Scavia et al. (2017)

Measure of model fit	Criteria for excellent fit	Multi-model average	HU	LT	OSU	TA&MU	UM	UT
Flow								
PBIAS	+/- 10%	6%	-7%	10%	10%	11%	6%	--
NSE	> 0.5	0.88	0.82	0.90	0.91	0.86	0.89	--
R^2	> 0.6	0.90	0.86	0.91	0.93	0.88	0.91	--
Total phosphorus								
PBIAS	+/- 25%	2%	37%	-6%	-7%	-22%	7%	--
NSE	> 0.4	0.69	0.64	0.82	0.73	0.56	0.70	--
R^2	> 0.5	0.75	0.74	0.82	0.75	0.71	0.70	--
Dissolved Reactive Phosphorus								
PBIAS	+/- 25%	14%	81%	1%	16%	13%	-13%	--
NSE	> 0.4	0.44	-0.22	0.71	51%	0.52	0.46	--
R^2	> 0.5	0.60	0.55	0.71	54%	0.70	0.51	--

Based upon: Scavia et al. 2017 (Table A.2.1 in Appendix A2).

Model developers: HU = Heidelberg University; LT = LimnoTech; OSU = Ohio State University; TAMU = Texas A&M University; UM = University of Michigan; UT = University of Toledo.

Note: NSE = Nash-Sutcliffe Efficiency; PBIAS = percent bias; R^2 = coefficient of determination.

Table 6. Summary of SWAT models development – Validation in Martin et al. (2019)

Measure of model fit	Criteria for good fit	Multi-model average	HU	LT	OSU	TA&MU	UM	UT
Flow								
PBIAS	+/- 10%	2.2%	2%	11%	-3%	--	1%	0.1%
NSE	> 0.65	0.89	0.88	0.91	0.99	--	0.94	0.83
Total phosphorus								
PBIAS	+/- 25%	-2.7%	-7%	-13%	19%	--	1%	-13%
NSE	> 0.65	0.70	0.73	0.77	0.71	--	0.61	0.66
Dissolved Reactive Phosphorus								
PBIAS	+/- 25%	5%	7%	-15%	-4%	--	7%	32%
NSE	> 0.65	0.67	0.77	0.67	0.73	--	0.69	0.50

Based upon: Martin et al. 2019 (Table A-1 in Appendix A).

Model developers: HU = Heidelberg University; LT = LimnoTech; OSU = Ohio State University; TAMU = Texas A&M University; UM = University of Michigan; UT = University of Toledo.

Note: NSE = Nash-Sutcliffe Efficiency; PBIAS = percent bias.

3.1 SCENARIO RESULTS

Different sets of sensitivity analyses and BMP implementation scenarios were developed and evaluated (Table 7 and Table 8). Some scenarios targeted specific practices while other scenarios simulated combinations of practices. Sensitivity analyses were performed to how different levels of a management activity affected TP and DRP losses, even if such an activity was unrealistic (e.g., elimination of all permitted point sources). Finally, scenarios with agricultural BMPs either randomly distributed BMPs throughout or targeted BMPs to cropland with the highest phosphorus loss.

Key findings for scenarios with certain BMPs are presented in Section 3.1.2 and achievement of the Annex 4 targets for various scenarios is presented in Section 3.1.3.

Throughout this section, load results are presented. Herein, "load" always refers to the average of the five SWAT models in the ensemble, though the individual SWAT models differed between Scavia et al. (2016, 2017) and Martin et al. (2019, 2021). When Scavia et al. (2016, 2017) is referenced, "load" represents the average of the five SWAT models over the 2005 through 2014 period (i.e., 10-year averages). When Martin et al. (2019, 2021) is referenced, "loads" represents the average of the five SWAT models for individual years.

Table 7. Summary of SWAT models sensitivity analyses and BMP implementation scenarios in the ensemble modeling (Scavia et al. 2016, 2017)

No.	Name	Short description
1	No point source discharges	All point source discharges removed.
2a-c	Cropland conversion to grassland (10%, 25%, 50%; targeted)	10%, 25%, and 50% of row cropland converted to switchgrass. Targeted fields with high TP losses and low crop yields. [a]
3	In-field practices (25%; random)	Four practices [b,c,d,e] . 25% of cropland randomly selected.
4	Nutrient management (25%; random)	Three practices [b,c,d]. 25% of cropland randomly selected.
5	Nutrient management (100%)	Three practices [b,c,d]. 100% of cropland selected.
6	Commonly recommended practices (100%; random)	One of four practices [b,d,f,g] applied to each randomly selected separate 25% of cropland.
7	Continuous no-till (50%; random)	Two practices [d,f]. 50% of cropland randomly selected.
8	Series of practices (50%; targeted)	Three practices [d,e,g]. Targeted 50% of fields with highest TP losses.
9	Series of practices (50%; random)	Three practices [d,e,g]. 50% of cropland randomly selected.
10	Diversified rotation (50%; random)	Alternative rotation [e]. 50% of cropland randomly selected.
11	Wetland and buffer strips (25%; targeted)	Two practices [g,h]. Targeted 25% of fields with highest TP losses.

Based upon: Scavia et al. 2017 (Table 2)

Notes

a. Managed for wildlife; no livestock grazing. Annual October harvest. Annual nitrogen fertilizer, and no phosphorus fertilizer.
b. 50% reduction of phosphorus fertilizer application.
c. Fall phosphorus fertilizer application.
d. Subsurface phosphorus fertilizer application.
e. Winter cereal cover crop, when winter wheat is not planted.
f. Continuous no-tillage.
g. Medium quality buffer strips.
h. Wetlands that treat half of the overland flow in a subwatershed. Targeted 25% of subwatershed with greatest TP loading rates.

Table 8. Summary of SWAT modeling sensitivity analyses and BMP implementation scenarios in the ensemble modeling (Martin et al. 2019, 2021)

No.	Name	Short description
1	No point source discharge	All point source discharges removed.
2	No manure application	All phosphorus from manure application removed.
3	25% phosphorus rate reduction	25% reduction of phosphorus fertilizer application.
4	Broadcast phosphorus	All fertilizer is broadcast without incorporation.
5	Broadcast and incorporated phosphorus	All fertilizer is broadcast with incorporation.
6	Subsurface applied phosphorus	All fertilizer is applied below a 1-centimeter soil depth.
7	Fall manure	All manure is applied in the fall.
8	Spring manure	All manure is applied in the spring.
9	Fall and spring manure	One-half of manure is applied in the fall and one-half in the spring.
10	Rate, replacement, timing	Inorganic fertilizer and manure application reduced 50%, applied in subsurface during the fall.
11	Cereal rye cover crop	Cereal rye cover crops planted on all cropland.
12	Controlled drainage	Drainage water management in all tile-drained fields.
13	Headwater wetland	Wetlands in all model subbasins receive one-half of the flow.
14	In-field and buffers (random)	58% cover crops + 50% subsurface placement+ 78% buffer strips.
15	In-field and buffers (targeted)	58% cover crops + 50% subsurface placement+ 78% buffer strips.
16	Likely adoption of in-field and buffers (targeted)	60% cover crops + 68% subsurface placement+ 50% buffer strips.
17	In-field and wetlands (targeted)	58% cover crops + 50% subsurface placement+ 78% wetlands.
18	In-field and controlled drainage (targeted)	50% cover crops+ 60% subsurface placement+ 50% no-tillage+ 15% drainage water management

Based upon: Martin et al. 2019 (Table 2)

3.1.1 Sensitivity Analyses

Generally, the evaluation of results of the sensitivity analyses indicated that elimination of certain sources did not meet Annex 4 targets because other sources still contributed significant in-stream TP and DRP loads. Additionally, the results of the sensitivity analyses of land use conversion indicated that significant conversion would be necessary to meaningfully reduce TP.

3.1.1.1 Elimination of a Source

The elimination of all point sources was simulated as a sensitivity analysis (Scavia et al. 2016; Martin et al. 2019, 2021). Scenario results from the ensemble modeling indicated that the elimination of all point sources' phosphorus loads generates only small decreases in TP and DRP loads reaching Lake Erie. For example, elimination of phosphorus from all point sources loads resulted in a 5.7% decrease in TP loads and 10.4% decrease in DRP loads for the March-July season (Martin et al. 2020, p.6). Thus, the elimination of all phosphorus from all point sources is insufficient to meet the Annex 4 targets (Martin et al. 2019, 2021).

Elimination of all manure land application to cropland was simulated in the second set of ensemble modeling (Martin et al. 2019, 2021). Similar to the scenario that eliminated all point sources, this scenario was a sensitivity analysis. Elimination of phosphorus from all manure land application resulted in a 7.2% decrease in TP loads and 7.7% decrease in DRP loads for the March-July season (Martin et al. 2020, p.6). Thus, the elimination of all phosphorus from manure land application is insufficient to meet the Annex 4 targets.

3.1.1.2 Land Use Conversion

In Scavia et al. (2016), scenarios were developed to simulate the conversion of cropland to grassland. Conversion was targeted to cropland with the lowest crop yields and the highest TP losses. The grassland was managed for wildlife, with no livestock grazing. The switchgrass was harvested annually each October, received annual nitrogen fertilizer, and received no phosphorus fertilizer. The objectives of these scenarios were to show that a considerable amount of farmland would need to be taken out of production to reduce phosphorus levels unless in-field and edge-of-field BMPs were implemented.

Conversion of 10%, 25%, or 50% of cropland to grassland considerably reduced TP loads (average of 5 models across 2005-2014) but did not reduce DRP loads as much, relative to baseline (Scavia et al. 2016). Although not documented in the reports it appears that assumptions about initial soil P and tile drainage were left unchanged in cropland converted to grassland.

3.1.1.3 Combined Sewer Overflows

CSOs were not incorporated into the ensemble modeling presented by Scavia et al. (2016, 2017). In the follow-up ensemble modeling, Martin et al. (2019, 2021) simulated CSOs using available CSO data, estimates from regressions of CSO data and precipitation, and gap-filling. The results from Martin et al. (2020) were similar to Scavia et al. (2017), which did not simulate CSOs. As such, CSOs appear to have minimal effect upon March-July seasonal total phosphorus loading relative to the sum of other sources.

3.1.2 Best Management Practices

Generally, model scenario results, as constructed, did not meet Annex 4 targets for TP and DRP load and FWMC, when averaging results across the 5 models. To achieve Annex 4 targets, suites of BMPs will need to be adopted at higher implementation rates than were simulated in the models.

This section presents summaries of key findings for certain BMPs.

3.1.2.1 In-Field Agricultural BMPs

Scenario results from Scavia et al. (2016) indicate that the two ‘series of practices’ scenarios result in the largest load reductions; the series of practices are: subsurface application of phosphorus fertilizer and cereal rye winter cover crop (aside from winter wheat).

Similarly, results in Martin et al. (2020) indicate that fertilizer application rate reductions, subsurface fertilizer placement, use of cover crops, and installation of headwater wetlands had the largest reductions in TP concentrations and loads and decreased DRP concentrations and loads (Martin et al. 2020, p. 6). Altering the timing of fertilizer application and changing all fertilizer application from broadcast to incorporation had inconsistent effects. The models generally showed that subsurface placement "had the greatest potential to reduce TP and DRP concentrations and loads" (Martin et al. 2020, p.6).

3.1.2.2 Edge-of-Field BMPs

Buffer strips and wetlands were simulated as edge-of-field BMPs (Scavia et al. 2016; Martin et al. 2019, 2021). In Scavia et al. (2016), the scenario with targeted buffer strip[7] and wetland installation[8] showed considerable reduction with the TP load but much less relative reduction with the DRP load. Similarly, in Martin et al. (2020) ensemble modeling showed that headwater wetlands had large reductions in TP concentrations and loads.

3.1.3 Target Achievement

In Scavia et al. (2016, 2017), March-July total loads for each model were averaged for each year, then the 10 years were averaged and compared with Annex 4 load targets. In Martin et al. (2019, 2021), March-July total loads and FWMCs for each model were averaged for each year, then each annual load or FWMC was compared with the Annex 4 load or FWMC target; simulated loads and FWMCs were evaluated using the 9-of-10 year target. The review presented herein focus more on the Martin et al. (2019,2020) target evaluations.

Scenario results generally indicated that achievement of the Annex 4 targets in 9-of-10 years will be difficult even with widespread adoption of a suite of BMPs.

While BMP scenarios for individual models resulted in the Annex 4 targets being met in 9-of-10 years, the average of model results for each BMP scenario never resulted in the Annex 4 TP load target being met in 9-of-10 years (Martin et al. 2019, 2021). Across results for the five models, the average number of years the TP target was met ranged from 5.8 to 6.8 years. TP FWMC results for a few scenarios for a few models met the Annex 4 FWMC target for TP.

None of the model scenarios averages of results met the Annex 4 target for DRP load in 9-of-10 years (Martin et al. 2019, 2021); the best performance for model DRP loads (average of 5 models) met the Annex 4 target for roughly half of the years (Martin et al. 2019). However, several scenarios met the DRP load target for 9-of-10 years in individual models. One of the BMP scenarios for one model was able to achieve the Annex 4 FWMC target for DRP (Martin et al. 2020).

3.1.3.1 Level of Implementation

Scenarios in two sets of ensemble modeling (Scavia et al. 2016, 2017; Martin et al. 2019, 2021) evaluated the levels of implementation necessary to achieve Annex 4 load targets. The authors compared different scenarios with differing rates of BMP implementation to draw conclusions. Note that Scavia et al. evaluated average results over 10 years (i.e., 10-year averages versus Annex 4 target loads), while Martin et al. evaluated achieving targets in 9 out of 10 years (i.e., individual year averages versus Annex 4 target loads).

---

[7] Buffer strips were targeted to the 25% of row cropland with the greatest TP losses.
[8] Wetlands were targeted to the 25% of subwatersheds with the highest TP losses.

In a sensitivity analysis, with the three model simulations of the conversion of 10%, 25%, and 50% of cropland to grassland, only the 50% conversion simulation met both the TP and DRP targets (average of 5 models across 2005-2014, comparing the 10-year average to the Annex 4 target loads) and neither target was met with the 10% conversion simulation (Scavia et al. 2016). These results indicate that without implementation of in-field and edge-of-field BMPs, to achieve Annex 4 TP and DRP targets, a quarter to half of cropland would need to be taken out of production (Scavia et al. 2017).

Several scenarios simulated different implementation rates of various in-field and edge-of-field suites of BMPs. A few scenarios can lead to Annex 4 load target achievement (when loads are averaged over the 2005-2014 period); "[h]owever, all the successful pathways require broad implementation of both common and less common practices" (Scavia et al. 2017, p. 129). Achievement of the necessary rates of BMP implementation will be a challenge. The authors further describe this effort as "daunting task and will require extensive changes in management and much greater investment of resources to achieve the required levels of implementation" (Scavia et al. 2017, p. 130).

3.1.3.2 Random versus Targeted Distribution

The distribution of BMPs across the watershed was evaluated in both sets of ensemble modeling (Scavia et al. 2016; Martin et al. 2019, 2021). Scenarios were developed for both a random distribution of specific BMPs across the watershed and a distribution of BMPs that targeted areas with high TP loss.

Scavia et al. (2016) evaluated TP and DRP target achievement (i.e., 10-year average compared with Annex 4 target loads) with two scenarios that used either random or targeted distribution for in-field and edge-of-field BMPs for 50% of cropland. As expected, the targeted distribution loads were smaller than the randomized distribution loads (Scavia et al. 2016). While both randomized and targeted distribution scenarios' TP loads met the Annex 4 target, when simulated loads were averaged over 2005-2014, only the targeted distribution scenario's DRP load met the Annex 4 target (Scavia et al. 2016); this study did not evaluate Annex 4 targets annually using the 9-of-10 years target

Similarly, targeting implementation of cover crops, riparian buffers, and subsurface fertilizer placement to existing areas with the highest phosphorus loads (versus random distribution) resulted in Annex 4 targets being met for more years (Martin et al. 2020; Scavia et al. 2017).

3.1.3.3 Dissolved Reactive Phosphorus versus Total Phosphorus

Achieving the DRP target will likely be more difficult than achieving the TP target. Scavia et al. (2016) found that the TP load target was achieved in 5 of 13 scenarios, whereas the DRP load target was achieved in only 3 scenarios and just barely at that; this analysis averaged the 10 years of seasonal loads.

Similarly, Martin et al. (2021, p. 7) concluded that "implementation of most scenarios would make less progress in meeting the DRP loading targes compared to TP". Unlike Scavia et al (2016) discussed above, Martin et al. (2021) compared the averages of the five models for individual years to the Annex 4 target loads with the 9-of-10 year goal. The authors further concluded that "the agreement among model results provide confidence in the finding that the bundled-management scenarios are more likely to reach the TP target than the DRP target" (Martin et al. 2021, p. 6).

3.2 CRITICAL SOURCE AREAS

Critical source areas (CSAs; i.e., "hotspots") were evaluated through the ensemble modeling evaluations (Evenson et al. 2021). For each model, the 20% of HUC12 subwatersheds (of the 252 HUC12 subwatersheds in the Maumee River basin) with the highest export of flow, TP, DRP, total nitrogen, and total suspended solids (TSS) were identified as CSAs. The CSAs between models were then evaluated statistically and graphically to determine patterns.

Generally, the ensemble modeling did not agree on the location of CSAs: "the overwhelming majority of HUC-12s identified as CSAs were identified as such by a minority of models" (Evenson et al. 2021, p. 5). This observation suggests that the models were not robust in their ability to identify CSAs, probably due to calibration mostly at the large-basin scale and not at the subbasin scale.

Ensemble modeling did agree for a small subset of CSAs: TP and TSS in the northwest corner of the Maumee River basin (i.e., in the upper St. Joseph River and Tiffin River watersheds) and DRP along the Maumee River mainstem (i.e., the lower Maumee River). Ensemble modeling also agreed on "a large subset of subwatersheds [that] were not simulated as a CSA by any of the models and that these areas could be removed for consideration for conservation and restoration action" (Evenson et al. 2021, p. 7).

The quantity of fertilizer application per HUC12 subwatershed was evaluated with the CSAs to determine if the quantity of fertilizer applied was correlated to CSA identification. The authors generally found that CSAs were more likely to be identified in areas with higher fertilizer application but that fertilizer application did not explain much of the variation of model outputs (Evenson et al. 2021, p. 6). They concluded that "fertilizer application rates alone were only weakly related to nutrient export and thus CSA location for most [of the SWAT] models" (Evenson et al. 2021, p. 5), indicating that other factors exert significant controls on the fraction of fertilizer nutrients that is exported to streams.

4.0 INDIVIDUAL SWAT MODELING

Tetra Tech also reviewed publications about SWAT modeling efforts for watershed in the Great Lakes basin and Gulf of Mexico basin. Six SWAT models that were not used in or related to the ensemble modeling efforts were developed to evaluate BMP scenarios (Table 9), and the these studies are summarized in the six subsections of Section 4.1. Six additional publications about the SWAT models for the Maumee River watershed, developed by teams of researchers from the University of Michigan and Ohio State University, some of which were later used for ensemble SWAT modeling, were also reviewed and summarized in Section 4.2. Finally, Tetra Tech's SWAT models for three watersheds in the Lake Erie basin are summarized in Section 4.3.

Table 9. Individual SWAT models with BMP scenarios

Citation	Watershed	Conservation cover	Conservation crop rotation	Conservation tillage	Cover crop	Cropland conversion	Fertilizer reduction	Filter strip	Grassed waterway	Nutrient management	Tillage (reduced or mulch)	Tillage (strip or no)	Two-stage ditch	WASCOB	Other
Cho et al. (2010)	Little River (Gulf of Mexico Basin)	--	--	✓	--	--	✓	--	✓	--	--	--	--	--	✓[a]
Christopher et al. (2017)	River Raisin (Western Lake Erie Basin)	--	--	--	--	--	--	--	--	--	--	--	✓	--	--
Hanief and Laursen (2019)	Grand River (Eastern Lake Erie Basin)	--	--	--	✓	✓	✓	✓	✓	--	--	--	--		✓[b]
Her et al. (2016)	St. Joseph River (Western Lake Erie Basin)	✓	✓	--	✓	--	--	✓	--	✓	✓	✓	--	✓	✓[c]
Merriman et al. (2019)	Upper East River (Lake Michigan Basin)	✓	✓	--	✓	--	--	✓	✓	✓	✓	✓	--	--	✓[d]
Wang et al. (2017)	Gully Creek (Lake Huron Basin)	--	--	✓	✓	--	--	--	--	✓	--	--	--	✓	✓[e]

Notes

SWAT = Soil and Water Assessment Tool; WASCOB = water and sediment control basin.

a. Contour farming, riparian forest cover, and terrace.

b. Bank stabilization.

c. Biomass planting and field border.

d. Waste (manure) storage facility.

e. Soil amendment and windbreak planting.

4.1 SWAT MODELS WITH BMP SCENARIOS

4.1.1 Little River (Cho et al. 2010)

The Little River experimental watershed is in Georgia and Florida, and the Little River discharges to the Gulf of Mexico. The watershed has low relief. The predominant soils are sands and sandy loams, both with high infiltration, that are underlain by an impermeable clay layer. The watershed is 41% row crop and pasture. The author does not discuss irrigation or tile-drainage. In recent years, riparian forests have been converted to wet pasture. The climate is humid subtropical.

A SWAT 2005 model was developed using AVSWAT-X to simulate 1996 through 2004. The model was calibrated for one small subwatershed and validated for the entire Little River watershed. Three types of BMPs were simulated: four crop management BMPs (grassed waterways, terraces, contour farming, and conservation tillage), one nutrient management BMP (whole-farm plan to reduce both nitrogen and phosphorus fertilizers by 30%), and three riparian forest cover BMPs (0%, current [88%], and 100%; simulated as filter strips). Scenarios were developed for three different schemes of BMP-distribution: random BMP distribution, BMP distribution by stream order, and prioritized to higher loading HRUs. Different rates of implementation were also simulated, from 11% to 41% of the area of the watershed.

Full implementation of the crop management BMPs resulted in reductions of 55% for sediment and 56% for TP, while full implementation of the nutrient management BMP resulted in reductions of <1% for sediment and 4% for TP.

For BMP distribution, "the study showed that using subwatershed nonpoint source pollutant load as a prioritization criterion resulted in the most rapid water quality improvement" (Cho et al. 2010, p. 463). Below about 50% implementation rate, random BMP distribution and BMP distribution by stream order were nearly indistinguishable. At greater than about 50% implementation rate, BMP distribution by stream orders showed more rapid improvement than random distribution "For both sediment and TP, the greatest rate of reduction was achieved when implementing [BMPs] on the highest priority 20% of the [watershed] upland cropped area without currently documented [BMPs]" (Cho et al. 2010, p. 470).

Increasing riparian forest cover from the current 88% to 100% would reduce sediment and TP loads by 21%. Increasing riparian forest cover to 100% shows the largest reductions with 1st and 2nd order streams (22% each), which the authors expected since much of the upland agriculture is directly drained by first and second order streams. For 3rd, 4th, and 5th order streams, the reductions were 4%, 1%, and 0.1% (respectively).

4.1.2 River Raisin (Christopher et al. 2017)

Empirical relationships between nutrient reduction and two-stage ditch implementation were paired with SWAT model results for the River Raisin watershed (southeast Michigan; Western Lake Erie Basin) to estimate watershed-scale load reductions for different rates of two-stage ditch implementation. The empirical relationship was developed using data from nine ditches in Indiana, Michigan, and Ohio that drain subwatersheds composed primarily of tiled corn-soybean fields. Two-stage ditches installed on former steep, incised ditches were paired with upstream reference reaches. Due to limited TP data, turbidity was used as a surrogate for TP based upon a statistical relationship between TP and turbidity.

The River Raisin SWAT model was developed by down-scaling a SWAT 2012 model for the Western Lake Erie basin developed earlier. The model was calibrated for 1990-1999 and validated for 2000-2006. For scenario development, the baseline SWAT model was run for two years (2009-2010).

Three different scenarios were simulated for the River Raisin watershed in SWAT: 25%, 50%, and 100% implementation of two-stage ditches in headwaters reaches. For this study, headwaters ditches discharge 2,000 liters/second or less (about 71 cfs or less), which account for about 75% of stream miles in the watershed. Baseline and scenario time-series of flow, nitrate, and TP were exported. The empirical relationships were then applied on the model reach scale.

The 25%, 50%, and 100% implementation of two-stage ditches on headwaters streams reduced annual TP loads for the River Raisin watershed by 12%, 20%, and 31%, respectively.

4.1.3 Grand River (Hanief and Laursen 2019)

The authors previously developed a SWAT model of the Grand River watershed (southeastern Ontario; Eastern Lake Erie Basin) for the 1996 through 2010 period, with 1996 through 2000 as the warm-up period, 2001 through 2005 as calibration, and 2006 through 2010 as validation. Due to limited sediment and nutrient data, the authors manually calibrated water quality.

About 28% of agricultural land in the Grand River watershed is tiled. Tile drainage was simulated with the following three assumptions: the depth to drain was 900 millimeter (mm), the time to drain the soil to field capacity was 24 hours, and the tile drain lag time was 3 hours.

Scenarios were run for the 2001 through 2010 period. Of all the BMPs, relative to baseline, 7.6-meter filter strips with conservation tillage resulted in the largest watershed-scale reductions (Table 10). The authors concluded that multiple BMPs need to be implemented in high-yield subwatersheds to achieve the Annex 4 target reductions (Hanief and Laursen 2019, p. 173).

Table 10. Hanief and Laursen (2019) BMP scenario results: watershed-scale load reductions

Scenario	Sediment	DRP	TP
Bank stabilization (increased channel trapping efficiencies)	38%	24%	36%
Cover crop (red clover planted in early fall, ploughed in May)	5%	11%	11%
Cropland conversion - 15% to forestland 10% to wetland	9% 2%	1% 1%	9% 4%
Filter strips with conservation tillage - 3-meter width 7.6-meter width	13% 23%	18% 48%	34% 50%
Grassed waterway	15%	2%	17%
Reduced fertilizer (by 20%) with conservation tillage	8%	33%	22%

Source: Hanief and Laursen 2019.

4.1.4 St. Joseph River (Her et al. 2016)

A SWAT model was developed for the St. Joseph River watershed (Maumee River watershed; Western Lake Erie Basin) with 39 model subbasins and 498 HRUs. Six agricultural management combinations were randomly distributed to the corn-soybean HRUs. Practice information for 10,028 BMPs implemented in 2005 through 2013 were identified. As not all practices affect water quality and some practices effects on water quality are not known, only 19 of the 201 types of BMPs were incorporated into the model.

The hydrology calibration and validation was for 1993 through 2009 using three USGS gages in the watershed. The water quality calibration and validation was for 2001 through 2009 using TP, TSS, and total nitrogen data collected by the Indiana Department of Environmental Management.

Field- and watershed-scale reductions of sediment and TP are presented in Table 11. Split nitrogen fertilizer application is omitted, as it has no impact on sediment and phosphorus loading. Summing all the practices across the St. Joseph River resulted in reductions of less than 5% for sediment and 6% for TP.

Total and soluble load reductions were evaluated on a monthly basis as well. Phosphorus loads increased during certain times for four BMPs:

- *Conservation crop rotation*: Increased TP loads in August and November and DRP in July through November
- *Cover crops*: Increased TP loads in July through October and DRP in August through March
- *Mulch-till*: Increased TP loads in June through December and DRP in November through June
- *No-till*: Increased TP loads in December through March and DRP in May through September and November through March

The authors found that the highest performing BMPs were not widely implemented and BMPs were not implemented in the areas where they would do the most benefit. “The two most effective practices at the field scale, conservation cover and biomass planting, were only implemented on 0.47% of the watershed, and so provided little watershed-scale load reduction (Her et al. 2016, p. 259). Conservation crop rotation and cover crops yielded the largest watershed-scale reductions because these two BMPs were most frequently implemented. Finally, the authors concluded (Her et al. 2016, p. 261):

Results from this study indicated that tangible improvement in water quality at a watershed scale would require wide implementation and appropriate placement of conservation practices, as well as significant field scale load reduction efficiency by the conservation practices.

Table 11. Her et al. (2016) BMP scenario results: field-scale reductions

BMP (3-digit NRCS practice code)	Field-scale		Watershed-scale [a]	
	Sediment	TP	Sediment	TP
Conservation cover (327)	99%	>99%	0.4%	0.4%
Conservation crop rotation (328)	17.%	20%	1.2%	1.9%
No-till (329 + 329A)	14%	1%	1.0%	0.2%
Mulch till (329B + 345)	33%	5%	0.4%	<0.1%
Cover crop (340)	74%	62%	1.1%	1.2%
Field border (386)	43%	39%	0.2%	0.2%
Filter strip (393)	41%	42%	0.3%	0.4%
Biomass planting (512)	99%	>99%	0.1%	0.1%
Nutrient management (590) - 10% - 20%	0% 0%	4% 8%	0%	0.3%
Water and sediment control basin (638)	47%	40%	0%	0.6%

Source: Her et al. 2019.

Notes

Field-scale reductions were rounded to the nearest percentage point and watershed-scale reductions were rounded to the nearest one-tenth percentage point.

a. Watershed-scale reductions were calculated as "load reduction as a percentage of total load aggregated across the whole watershed".

4.1.5 Upper East River (Merriman et al. 2019)

A SWAT model was developed for the Upper East River subwatershed of the Lower Fox River watershed that drains to Green Bay of Lake Michigan. The predominant crops are corn and alfalfa, which are about 61% of the land use in the subwatershed. The typical crop rotations are dairy (3-years corn silage then 3-years alfalfa silage) and grain (2-year corn-soybean). Only about 20% of cropland is tiled. Land-application of manure is an important source of nutrients in this subwatershed because the numbers of dairy cattle and dairy operations have been increasing, while cropland has been decreasing due to suburban expansion. Three CAFOs and a WWTP are within this subwatershed.

The Upper East River subwatershed model was developed using ArcSWAT2012 for ArcGIS 10.3 (revision 655b). The SWAT model was calibrated to 2000-2014 with a 4-year warmup period (2000-2014) and validated to 2015-2016. The model was calibrated to four USGS gages (including a gage at the subwatershed outlet) and validated with a fifth USGS gage. The authors selected the Hooghoudt and Kirkham tile drain equations as the method to simulate tile drainage. Assumptions included 1,300 mm soil depth to the impervious layer, 900 mm soil depth to drains, 25 mm tile diameter, and 12,000 mm between tiles. Tile drainage was then included as a soft calibration parameter.

The new soil phosphorus model[9] was used. Initial organic phosphorus concentration in soil was set to 1.25 mg/kg. Initial soluble phosphorus concentration in soil was set to 0.4 to 118.5 mg/kg, based upon soil test phosphorus records from the Brown County Land and Water Conservation District.

Baseline and seven scenarios were simulated. Field-scale results were calculated using a 13-year period (2004-2016). For individual BMPs, conservation cover had the largest reductions of sediment, TP, and DRP; however, taking significant areas of cropland out of production is not feasible. In dairy rotations, residue management (i.e., mulch-, reduced-, strip-, and no-till) increased soluble nutrient loads. In cash grain rotations, cover crops increased TP and DRP loads.

Watershed-scale results indicated that "[s]ediment and nutrient loads decreased at the HUC 12 outlet as scenarios increased the number and area of simulated BMPs" (Merriman et al. 2019, p. 630; Table 12). Generally, dairy rotations had larger reductions with sediment yields, while cash grain rotations had larger reductions in nutrient yields.

Table 12. Merriman et al. (2019) BMP scenario results: watershed-scale reductions

Scenario	No. of BMPs	No. of fields	Portion of HUC12	Sediment reduction	DRP reduction	TP reduction
Applied GLRI BMPs	183	180	7%	2%	1%	2%
All applied BMPs	1,312	777	32%	13%	6%	10%
All applied BMPs + planned GLRI BMPs	1,462	824	35%	n/a	n/a	n/a
All contracted BMPs	1,923	972	40%	28%	9%	20%
Hypothetical - Low	3,069	1,115	45%	35%	14%	26%
Hypothetical - Medium	3,527	1,261	51%	39%	16%	30%
Hypothetical - High	6,056	1,476	62%	48%	20%	36%

Source: Merriman et al. 2019.

The authors presented the average annual reductions in sediment, DRP, and TP loads by BMP combination for dairy and cash grain rotations. Across all BMP combinations, for cash grains, reductions ranged from 2% to 100% for sediment, -22% to 92% for DRP, and -19% to 96% for TP. For cash grain rotations, the BMP-combination with the greatest nutrient reductions was *cover crops + nutrient management planning (split manure application via injection) + reduced tillage*. Manure injection reduced DRP levels more than conventional manure incorporation.

Results from this study indicate that BMPs' phosphorus reduction "effectiveness does not easily translate from field to watershed scale; complex processes (landscape, routing, weather, soil interactions, in-stream

[9] This refers to an improved method of calculating the stable soil phosphorus pool from Vadas and White (2010). This was not in SWAT 2009 but was incorporated in SWAT 2012. Vadas and White state that the earlier method is likely to "underpredict soil total P" and "may underpredict solution P for several weeks after P is added to soils. This could result in underprediction of dissolved inorganic P loss in runoff soon after a P application to soils."

nutrient dynamics, etc.) control nutrient transport from the EOF [edge-of-field] to stream and HUC 12" (Merriman et al. 2019, p. 621).

4.1.6 Gully Creek (Wang et al. 2017)

Gully Creek is a small lakeshore tributary to Lake Huron in southeast Ontario, Canada. The watershed is two-thirds agricultural, including soybeans (30%), corn (22%), and winter wheat (11%). About a quarter of the watershed is woodland.

The Gully Creek watershed was studied as part of the Watershed Based BMP Evaluation (WBBE) program in 2010-2013 and as one of six pilot watersheds in the Great Lakes Agricultural Stewardship Initiative (GLASI) in 2015-2017. BMPs were inventoried and water quality was monitored at stations near the mouth and in the headwaters. Edge-of-field monitoring was performed at key locations.

A SWAT model was developed as part of the WBBE program and further refined during the GLASI program. 96 model subbasins were delineated. Landcover was fixed using 2011 data and 687 HRUs were developed. A special tool was developed to facilitate detailed scheduling of land management operations for agricultural HRUs. All cropland was simulated as tile-drained, with the following three assumptions: 900 mm depth to drain, 24-hours to drain soil to field capacity, and 1-hour tile drain lag time. A specially developed water and sediment control basin (WASCOB) module was used during initial modeling in the WBBE program. During the GLASI program, refinement of the SWAT model included incorporation of three modules (snow redistribution, frozen soil, and WASCOBs) from CanSWAT.

Flow, sediment, and nutrient data were collected in 2010-2017 and used to calibrate and validate the model. The model simulated 2001-2017, with 2001-2009 for model warm-up. Both visual graphical comparisons and statistical measures indicated reasonable model performance.

Sets of model scenarios were developed for the WBBE and GLASI programs, with foci on WASCOBs and land management BMPs. Model scenarios were run for 2001-2017 and, results were determined for a 15-year average (2002-2016). Watershed-scale sediment and TP reductions are presented in Table 13.

Multiple baseline scenarios were developed. For example, the WASCOB baseline scenario assumed no WASCOBs; each of the existing or hypothetical WASCOB scenarios' results were then compared with the baseline scenario. The baseline scenario for the land management scenarios included existing WASCOBs.

In general, the authors found that watershed-scale reductions in TP load (as measured at the watershed outlet) were small because the total area of the BMPs was small relative to the area of the watershed. Edge-of-field reductions at individual fields were larger than watershed-scale reductions; for example, with the 3 field with cover crops in 2013-2014 scenario, watershed-scale TP reduction was 0.9% while edge-of-field reductions for the three fields were 30.7%, 14.4%, and 16.8%.

The pollutant reductions for the GLASI scenario with 23 BMPs for 2015-2018 was less than the summation of the individual scenarios for those BMPs. The authors explained the smaller reductions as due to "interactions of different processes on the landscape and marginal decrease in pollutant reduction efficiencies as more BMPs were implemented" (Wang et al. 2017, p. 106).

Table 13. Wang et al. (2017) BMP scenario results: watershed-scale reductions

Scenario	Sediment	TP
WASCOB Scenarios		
Baseline (no WASCOBs)	--	--
10 WASCOBs [a]	5.6%	4.1%
3 WASCOBs [b]	3.5%	2.3%
8 WASCOBs in or near monitoring sites (built in 2012)	2.2%	0.9%
All 44 existing WASCOBs [a,b,c]	25.4%	21.6%
All 44 existing WASCOBs + 3 planned WASCOBs	25.9%	22.0%
Land Management Scenarios		
Baseline (no land management BMPs; existing WASCOBs)	--	--
3 fields with cover crops in 2013-2014	0.3%	0.9%
3 fields with cover crops in 2014-2015	1.5%	0.3%
6 fields with precision nutrient management in 2016-2017 [b]	0%	1.4%
3 fields with precision nutrient management and soil amendment with manure application and 3 fields with strip tillage in 2016-2017 [b]	0.9%	0.9%
8 fields with precision nutrient management, 2 fields as no-till, and 3 fields with vertical tillage in 2017-2018 [b]	1.1%	1.5%
Various BMPs in 23 fields in 2015-2018 [b,c]	1.6%	2.1%
2 fields with windbreaks (that also acted as filter strips)	0.8%	0.6%

Source: Wang et al. 2017.

Notes

a. WASCOBs installed as part of the Watershed Based BMP Evaluation program.

b. BMPs installed as port of the Great Lakes Agricultural Stewardship Initiative program.

c. These scenarios are cumulative of the BMPs in the preceding scenarios.

4.2 MAUMEE RIVER SWAT MODELS

Several teams of researchers have developed SWAT models for the Maumee River watershed. In this section four SWAT models of the Maumee River watershed, developed by researchers at the University of Michigan and Ohio State University, are summarized. Some of these models have evolved from models that were previously developed as part of the ensemble modeling effort of the Maumee River watershed.

4.2.1 Model Development and Evaluation of 25 BMP Scenarios (Kalcic et al. 2016)

The team of researchers developed the SWAT model after extensive stakeholder engagement that included an online survey and series of workshops (Kalcic et al. 2016). Survey respondents included representatives of agricultural producers, agricultural advisors, Soil and Water Conservation Districts, nongovernmental organizations, academia, and local, state, federal, and international agencies. A primary objective of stakeholder engagement was to identify important BMPs. The first set of workshops introduced modeling to the stakeholders and provided stakeholders the opportunity to suggest model improvements. The second set of workshops presented modeling results and provided stakeholders the opportunity to identify additional high priority model scenarios.

HRUs were defined with a single slope class and 10% threshold for soil class lumping. Corn-soybean and corn-soybean-winter wheat rotations were simulated. Commercial fertilizer and manure fertilizer land application were simulated. "Tile drainage was simulated on row cropland with very poorly, poorly, and somewhat poorly drained soils (Kalcic et al. 2016, p. 8138). The researchers "modified the SWAT 2012 Revision 635 source code to correct a bug preventing soluble P (a proxy for DRP) from flowing through tile drains" (Kalcic et al. 2016, p. 8138). Tile drain spacing was set to 15,000 mm. During calibration, the depth to drain (DDRAIN) was set to 1,000 mm and the depth to impervious layer was set to 1,500 mm.

The maximum soil crack volume (SOL_CRK) was set to 0.45 (range 0 to 1) and the new soil phosphorus subroutine was used (SOL_P_MODEL). During calibration, the initial labile phosphorus soil concentration (SOL_SOLP) was set to 1 mg/kg.

This SWAT model included 10,266 HRUs and 358 model subbasins. The model was calibrated for 2001-2005, validated for 2006-2010, and back-validated for 1981-2000; calibration and validation were at the Waterville gage. Daily and monthly flow and water quality simulations were very good for the calibration and validation (except nitrate) using percent bias (+/- 10% flow, +/- 25% water quality), the Nash-Sutcliffe Efficiency (>0.5 flow, >0.4 water quality), and the coefficient of determination (>0.6 flow, >0.4 water quality). Sediment was underestimated and DRP overestimated during the back-validation. The researchers noted that existing BMPs were not well simulated due to a lack of data, especially with existing filter strips and cover crops.

A total of 25 simulations (including sensitivity analyses and BMP implementation scenarios) across six groups of management options were run: nutrient placement, nutrient timing, cover crops, vegetated filter strips, combinations, and feasible scenarios. With the robust stakeholder engagement, researchers were able to develop scenarios with BMPs capable of achieving phosphorus reductions that the stakeholders considered desirable and feasible (Kalcic et al. 2016, p. 8142).

Nearly all scenarios predicted reductions in TP and DRP loads. The researchers concluded that subsurface placement of phosphorus fertilizer was the most effective single practice for reducing DRP loads. Cereal rye cover and filter strips reduced TP but were less effective with DRP (due in part to tile drains bypassing filter strips). The researchers further concluded that after implementing subsurface placement of phosphorus fertilizer, “adding more practices achieves modest and diminishing returns on conservation investment” (Kalcic et al. 2016, p. 8142).

4.2.2 Legacy Soil Phosphorus (Muenich et al. 2016)

The Maumee River watershed SWAT model described by Kalcic et al. (2016) was used to evaluate the role of legacy soil phosphorus. Refer to Section 4.2.1 for discussion of several key tile drainage and soil phosphorus parameters. Four scenarios were created (see list below). With each scenario, a 12-year warm-up period was simulated, followed by an 80-year analysis period. A single year’s temperature and precipitation were simulated for each of the 80-years. These scenarios are akin to a sensitivity analysis because the complete cessation of certain types of fertilizer practices is not feasible in the “real-world”.

- **Business as usual**: Same fertilization and management used in baseline model
- **No fertilizers**: Cessation of all inorganic and organic nitrogen and phosphorus fertilizer applications
- **No phosphorus fertilizers**: Cessation of all inorganic and organic phosphorus fertilizer applications
- **No inorganic phosphorus fertilizers**: Cessation of inorganic phosphorus fertilizer applications but continuing manure applications

In the *no phosphorus fertilizers* scenario, compared to the business as usual scenario, “spring DRP loads decreased to at or below the targets within 3-5 years, and continued to decline over time” (Muenich et al. 2016, p. 8149). TP loads also decreased but took much longer to meet targets. In the *no inorganic phosphorus fertilizer* scenario, TP and DRP loads decreased slowly and “hovered” near the targets for much longer.

Weather was an important factor in results. In the *business as usual* scenario, DRP targets are met in dry years. Under dry and average streamflow and precipitation conditions, targets were met quickly with the *no fertilizers* and *no phosphorus fertilizers* scenarios; when manure application continued (*no inorganic phosphorus fertilizers scenario*), it took 10 more years until the targets were met. Because inorganic and organic phosphorus fertilizer application will continue in the real-world, achievement of targets will take even longer than was simulated in SWAT.

Cessation of nitrogen and phosphorus fertilizer affected crop yields. Corn was immediately affected by the elimination of nitrogen fertilizer. Both corn and soybean began to be significantly affected by the lack of phosphorus fertilizers after about 25 years, when crop yields begin to significantly decrease. The authors concluded that this 25-year lag indicates that Maumee River watershed soils contain significant legacy phosphorus.

4.2.3 Additional BMP Sensitivity Analyses (Muenich et al. 2016)

Five additional sets of BMP sensitivity analyses were also simulated:

- **Reduction of phosphorus fertilizer application rate**: Inorganic phosphorus fertilizers were reduced to 75%, 50%, 25%, and 0% of the baseline.
- **Implementation of cereal rye winter cover crop**: Cereal rye winter cover crop on 25%, 50%, 75%, or 100% of agricultural land.
- **Varying effectiveness of vegetated filter strips**: Poor to good quality vegetated filter strips that receive drainage from 25%, 50%, 75%, or 100% of agricultural land.
- **Implementation of alternative row crops**: Crop rotations of continuous sunflower, continuous lentil, or sunflower-lentil rotations. High, medium, or no phosphorus fertilizer application.
- **Implementation of cellulosic biofuel crops**: Shawnee witchgrass and miscanthus, with and without manure land application.

The *reduction of phosphorus fertilizer application rate* simulation had minimal impact upon crop yields. In fact, the 0% inorganic phosphorus fertilizer simulation results indicated that "manure applications, combined with legacy [phosphorus], may be sufficient for plant growth for this 30 year period" (Muenich et al. 2016, p. 8151). The *implementation of cereal rye winter cover crop* simulation reduced erosion and particulate phosphorus but had no impact on DRP in tile water.

The *varying effectiveness of vegetated filter strips* simulation reduced TP and DRP losses. DRP targets were not met because considerable portions of DRP loads bypass the vegetated filter strips through subsurface tiles. With regards to vegetated filter strips, the authors noted that SWAT does not simulate vegetated filter strips when water rises above the banks and enters the filter strips (i.e., SWAT only simulates filter strip when surface runoff flows from fields to the filter strips to the waterway).

The *implementation of alternative row crops* simulation was similar to the baseline corn/soybean rotation; the authors concluded that fertilizer application was a more important factor than crop type, at least in SWAT modeling.

With the *implementation of cellulosic biofuel crop* simulations, "[c]ompared to the baseline, if all agricultural lands produced either switchgrass or miscanthus, DRP and TP loads would diminish greatly" (Muenich 2016, p. 8152). TP load targets were always met, while DRP targets were not always met due to manure land application.

4.2.4 Development of a Finer Scale Model (Apostel et al. 2021)

A new finer scale SWAT model of the Maumee River watershed was developed that was similar to the model described by Kalcic et al. (2016). Both models rely on the same spatial datasets, generally use the same algorithms, and have the same SWAT code (revision 635 with modification for DRP transport in tiles). The major differences include

- **Hydrologic response units**: The Apostel et al. (2021) SWAT model has 24,256 HRUs (average size of 74 hectares), including 18,018 agricultural HRUs with 12,676 tiled agricultural HRUs. The Kalcic et al. (2016) SWAT model had 10,266 HRUs (average size 1,385 hectares), including 870 agricultural HRUs with 645 tiled agricultural HRUs.

  The Apostel et al. (2021) model was developed to almost the field-scale, which is not common for SWAT models. Model run-time tripled for the Apostel et al. (2021) model.

- **Crop Management**: The Apostel et al. (2021) model includes 29 crop rotations of corn, soybean, winter wheat, alfalfa, and pasture. The Kalcic et al. (2016) model includes only 12 rotations, all with identical management operations. The Apostel et al. (2021) model includes organic fertilizer application based upon published research about animal feeding operations in Ohio, whereas the Kalcic et al. (2016) model based manure application off of county livestock estimates.
- **Soil phosphorus**: Initial soil phosphorus levels were based on soil test phosphorus measurements in the Apostel et al. (2021) model and were calibration parameters in the Kalcic et al. (2016) model.
- **Preferential flow**: Preferential flow through SWAT's soil crack flow routine is enabled in the Apostel et al. (2021) model to represent macropore loading into tile drains.
- **Snow parameters**: The Apostel et al. (2021) model used snowfall and snow depth from the Global Historical Climatology Network as calibration targets, while the Kalcic et al. (2016) model treated the snow parameters as calibration parameters.

In addition to calibration and validation at the Waterville gage on the Maumee River, the Apostel et al. (2021) and Kalcic et al. (2016) models were evaluated with edge-of-field monitoring data.

- **Discharge**: The models tended to overestimate surface discharge and tile discharge; however, the surface-to-tile discharge ratios between the models and edge-of-field datasets were close.
- **Total phosphorus loads**: "Simulated annual surface TP loads were significantly greater than observed (Wilcoxon p-values were $p<0.05$ for each paired group comparison), while simulated annual tile TP loads were significantly less than observed ($P<0.05$)" (Apostel et al. 2021).
- **Dissolved reactive phosphorus loads**: The Apostel et al. (2021) model overestimated annual surface DRP loads and underestimated annual tile DRP loads. The Kalcic et al. (2016) model displayed the opposite pattern.

The authors noted that SWAT does not have the capacity to simulate transport of particle-bound phosphorus through tile drains.

4.2.5 BMP Sensitivity Analyses with Finer Scale Model (Kast et al. 2021a)

The Maumee River watershed SWAT model described by Apostel et al. (2021) was used to evaluate phosphorus loading to waterways from surface runoff and subsurface drainage. Another objective was to assess model sensitivity to changing soil phosphorus concentrations and inclusion of irrigation and liquid manure fertilizer application.

SWAT model results indicate that the sources of DRP and TP were, in March-July:

Dissolved Reactive Phosphorus

- Inorganic fertilizers (58%)
- Soil (18%)
- Manure (12%)
- Point sources (12%)

Total phosphorus

- Soils (45%)
- Inorganic fertilizers (42%)
- Manure (8%)
- Point sources (5%)

Here, "soil" represents legacy stores within the soil matrix, much of which derives from past fertilizer application, while the other sources are for the current season.

Twenty-one simulations were developed (see list below) and evaluated at a monthly time-step for March through July during a 2005-2015 timeframe. These simulations are sensitivity analyses that are infeasible from the standpoint of BMP implementation.

- No point sources
- No phosphorus in manure application [a]
- No inorganic phosphorus fertilizers
- No phosphorus in manure application [a] and no inorganic phosphorus fertilizers
- No point sources, no phosphorus in manure application [a], and no inorganic phosphorus fertilizers
- No irrigation concurrent with liquid manure application
- Initial labile phosphorus concentration in agricultural soils set to 5 milligrams per kilogram
- Initial labile phosphorus concentration in agricultural soils reduced to 25%, 50%, and 75% of baseline or increased to 125%, 150%, 175%, or 200% of baseline
- Inorganic fertilizer and manure application reduced by 25%
- Inorganic fertilizer and manure application reduced by 25% and point sources removed
- Inorganic fertilizer and manure application reduced by 25%, all phosphorus from manure application removed [a], and point sources removed
- Inorganic fertilizer and manure application reduced by 25%, all phosphorus from inorganic fertilizer application removed, and point sources removed
- All point sources, inorganic phosphorus fertilizer and manure phosphorus applications removed [a]
- Manure application rates tripled
- Manure application rates tripled and all phosphorus in manure application removed [a]

a. Nitrogen in manure applications was maintained to support crop growth.

The authors conclusions are summarized below:

- **Sources**: Manure and point sources contributed similar masses of TP and DRP, while inorganic fertilizers and soils contributed considerably more TP and DRP.
- **Legacy soil phosphorus**: "In the calibrated baseline model, soil sources of phosphorus contributed between 39% and 49% of the annual TP and 19% and 27% of the annual DRP discharged from the watershed" (Kast et al. 2021a).
- **Delivery fractions**: The average delivery fractions for inorganic fertilizer (2.7% TP and 0.7% DRP) and manure (3.0% TP and 1.2% DRP) were similar. Delivery fractions increased in wetter years and decreased in drier years.
- **Irrigation**: "Removing the irrigation operation applied concurrently with liquid manure applications had a negligible impact on nitrogen and phosphorus loadings from the watershed" (Kast et al. 2021a).
- **Initial labile phosphorus concentrations**: Increasing or decreasing the initialized concentration of labile phosphorus within agricultural HRUs increased or decreased (respectively) TP and DRP loads but had more impact upon DRP loads.
- **Tripling manure application**: While maintaining baseline application of inorganic fertilizers, also "over-application of may contribute a disproportionately greater load of phosphorus than manure applied at lower rates" (Kast et al. 2021a).

4.2.6 BMP Targeting with Finer Scale Model (Kast et al. 2021b)

The Maumee River watershed model described in Apostel et al. (2021) was updated to include "conservation identities" based on farmer surveys to better represent BMP implementation. Scenarios were developed with BMP placement based either (1) solely on field-phosphorus losses or (2) field-level phosphorus losses and conservation identities of the farmers. Three levels of conservation identities (weak, moderate, and high) were developed using a measure of existing BMP adoption and measures of county-level results from farmer surveys. Two BMPs were used for scenario development: subsurface placement of inorganic phosphorus fertilizer and buffer strips. Five scenarios were developed for BMP adoption; one scenario was random-distribution and the other four scenarios targeted HRUs with the

- Greatest phosphorus loading rates
- Least phosphorus loading rates
- Greatest conservation identity
- Greatest phosphorus loading rates and greatest conservation identity

As expected, targeting BMP implementation to the HRUs with the greatest and least phosphorus loading rates yielded the highest and lowest efficiencies (respectively); here, efficiency is the phosphorus reduction to rate of BMP adoption. Targeting BMP implementation to the HRUs with the greatest conservation identity had similar efficiency to random BMP distribution. Targeting BMP implementation to the HRUs with the greatest phosphorus loading rates and greatest conservation identity yielded the second highest efficiency.

Kast et al. (2021b, p. 7) opines that the similar efficiencies of randomly distributing BMPs and targeting BMPs to HRUs with the greatest conservation identity may indicated that there is "little relation between farmers' conservation identity and runoff from their fields" because farmers do control the practices on their fields but have limited control over the landscape.

Due to several economic and social factors that impact farmers, simulated phosphorus reductions from models that only target BMP implementation to the HRUs with the greatest loading rates are not likely representative of potential phosphorus reductions because BMP adoption is likely overestimated for the Maumee River watershed. “An approach that accounts for behavioral factors in responses to program incentives is likely much more realistic than believing that all farmers are equally likely to implement conservation practices on their fields” (Kast et al. 2021b, p. 8).

4.2.7 Additional SWAT Limitations (Kalcic et al. 2016)

In addition to the limitations with SWAT model tracking of DRP (discussed in Section 2.0), the following additional limitations were identified in Kalcic et al. (2016).

- **Winter soil conditions**: SWAT does not restrict fertilizer applications to the soil surface and subsurface during frozen or saturated ground conditions.
- **Cover crops**: SWAT does not simulate the improvements to soil organic matter and infiltration capacity that result from cover crops. Thus, cover crops may affect phosphorus-loading in ways that the model does not account for.

4.3 TETRA TECH SWAT MODELS

Tetra Tech developed SWAT models for watersheds in the Lake Erie basin, under contract with U.S. EPA Region 5, to support Ohio EPA and the Indiana Department of Environmental Management (IDEM). Three SWAT models were developed to support TMDLs and one SWAT model was further enhanced to evaluate BMPs[10]. The three models developed to support TMDLs were not used to simulate BMPs or other scenarios.

Each of the models incorporated tile drainage. Lessons from Tetra Tech's SWAT model development provide additional perspective on strengths and weaknesses of SWAT modeling for tracking DRP and the potential effectiveness of different types of BMPs.

4.3.1 TMDL Models

Tetra Tech developed SWAT models for three watersheds to support Ohio EPA and IDEM with TMDL development:

- **Black River** (Ohio EPA 2021)[11]: ArcSWAT version 2012.10_1.15 (Winchell et al. 2013) and SWAT Editor version 2012.10_2.15.
- **Sandusky River** (Ohio EPA 2014): SWAT 2012 revision 591
- **St. Joseph River** (IDEM 2017)[12]: ArcSWAT version 2012.10_1.15 (Winchell et al. 2013) and SWAT Editor version 2012.10_2.18.

Modeling reports for the Sandusky River and St. Joseph River SWAT application are available on each agency's TMDL website; the report for the Black River application will be available once Ohio EPA public notices the final TMDL. Tile drainage was simulated in agricultural or hay HRUs[13] with cross-listed HSGs (i.e., A/D, B/D, and C/D). For the Black River and St. Joseph River SWAT model, tile drainage was also limited to HRUs with low slopes (<6% for the Black River and <5% for the St. Joseph River). Key model inputs related to tile drainage are presented in Table 14.

[10] The Sandusky SWAT model was later updated multiple times in 2017 through 2021, under contract with U. S. EPA or the Ohio Department of Natural Resources, to support the development of an Environmental Fluid Dynamics Code (EFDC) model of Sandusky Bay. These revisions of the Sandusky SWAT model were used to develop boundary condition inputs to the Sandusky Bay EFDC model. No BMPs or scenarios were simulated in these revisions of the Sandusky SWAT model.

[11] The SWAT model report was public noticed in December 2021 as part of the TMDL report. Ohio EPA will submit the TMDL report to U.S. EPA for approval in 2022.

[12] Ohio EPA will include the SWAT model report as part of a TMDL report that will be public noticed and submitted for U.S. EPA approval in 2022.

[13] Cultivated cropland (#82) and pasture/hay (#81) in the National Land Cover Database.

Table 14. Key model inputs in the Tetra Tech SWAT models

Parameter	Description	Black	Sandusky	St. Joseph
DEPIMP	Depth to impervious layer in soil profile (mm)	2,500	2,500	2,500
DDRAIN	Depth to subsurface drain (mm)	1,000	1,000	1,000
GDRAIN	Time to drain soil to field capacity (hour)	12	24	12
TDRAIN	Drain tile lag time (hour)	12	48	24

Sources: IDEM 2017; Ohio EPA 2014, 2021.

All three models were calibrated and validated by evaluating statistics with acceptance criteria and by visual evaluation of multiple charts of observed and simulated data. The statistics used for both flow and phosphorus were the Nash-Sutcliffe coefficient (NSE), root mean square error observations standard deviation ratio (RSR) and the magnitude of the relative average error (RE, also known as PBIAS)[14]. Flow calibration and validation also included nine additional statistics[15].

Flow calibration was typically at U.S. Geological Survey gages with long-term daily flow data. Water quality calibration was at monitoring sites operated by the National Center for Water Quality Research, with long-term data, or at monitoring sites operated by state agencies ,with various frequencies of data collection. Table 15 summarizes the nine flow calibration and validation statistics, and Table 16 summarizes key calibration and validation information for these three SWAT models.

[14] The acceptance criteria are presented in the table below.

Metric	Very good	Good	Satisfactory	Unsatisfactory
NSE	>0.75	0.64 - 0.75	0.50 - 0.65	≤0.50
RSR	≤0.50	0.51 - 0.60	0.61 - 0.70	>0.70
RE (flow)	≤ ±10%	≤ ±15%	≤ ±25%	≥ ±25%
RE (phosphorus)	≤ ±25%	≤ ±40%	≤ ±70%	≥ ±70%

[15] The nine statistics are: error in total volume (≤ 10%), error in 50% lowest flows (≤ 10%), error in 10% highest flows (≤ 15%), seasonal volume error in the summer (≤ 30%), seasonal volume error in the fall (≤ 30%), seasonal volume error in the winter (≤ 30%), seasonal volume error in the spring (≤ 30%), error in storm volumes (≤ 20%), and error in summer storm volumes (≤ 50%).

Table 15. Flow calibration and validation in the Tetra Tech SWAT models

Description	Accept. criteria	Black Calibration	Black Validation	Sand. Calibration	Sand. Validation	St. Joseph Calibration	St. Joseph Validation
No. of calibration/validation sites	--	1	1	5	4	5 [a]	5 [a]
Error in total volume	≤ ±10%	-1.0%	-4.6%	-3.9% to -1.1%	-8.7% to 1.6%	6.5% to 11%	0.8% to 9.7%
Error in 50% of lowest flows	≤ ±10%	-8.6%	-7.6%	-8.2% to 18.4%	-0.6% to 14%	-27% to 20%	-13% to 9.7%
Error in 10% of highest flows	≤ ±15%	-9.8%	-10%	-7.6% to -2.8%	-14% to -8.0%	-7.9% to 9.9%	-17% to -0.6%
Seasonal error volume - Summer	≤ ±30%	33%	60%	7.7% to 86%	46% to 66%	28% to 133%	64% to 124%
Seasonal error volume - Fall	≤ ±30%	-1.4%	-16%	1.9% to 14%	-0.4% to 5.2%	11% to 27%	6.7% to 26%
Seasonal error volume - Winter	≤ ±30%	-14%	-12%	-21% to -11%	-29% to -3.3%	-2.2% to 7.0%	-7.0% to -0.8%
Seasonal error volume - Spring	≤ ±30%	6.7%	1.8%	6.7% to 11.8%	-7.9% to 3.5%	-0.2% to 3.7%	-29% to -17%
Error in storm volumes	≤ ±20%	0.2%	0.4%	-8.3% to 3.7%	-9.5% to 9.0%	-17% to 5.3%	-26% to -2.7%
Error in summer storm volumes	≤ ±50%	28%	64%	-2.9% to 84%	30% to 65%	17% to 84%	35% to 72%

Sources: IDEM 2017; Ohio EPA 2014, 2021.

Note a: Four sites were used for both calibration and validation, one site was used only for calibration, and one site was used only for validation.

Table 16. Model calibration and validation Tetra Tech SWAT models

Description	Black	Sandusky	St. Joseph
Flow Calibration and Validation			
No. of flow calibration gages	1	5	5
Flow calibration statistics	NSE = 0.94 RSR = 0.24 RE = -0.8%	0.62 < NSE < 0.77	0.85 < NSE < 0.90 0.32 < RSR < 0.39 6.5% < RE < 11%
Flow calibration narrative	Very good	Satisfactory to very good	Very good
No. of flow validation gages	1	4	5
Flow validation statistics	NSE = 0.93 RSR = 0.26 RE = -4.6%	0.66 < NSE < 0.78	0.69 < NSE < 0.91 0.30 < RSR < 0.56 0.8% < RE < 9.7%
Flow validation narrative	Very good	Good to very good	Good to very good
Phosphorus Calibration and Validation			
No. of phosphorus calibration gages	1	3	12 [a]
Phosphorus calibration statistics	NSE = 0.80 RSR = 0.45 RE = 4.1%	0.64 < NSE < 0.76	0.76 < NSE < 0.84 0.47 < RSR < 0.49 -18% < RE < -6.3%
Phosphorus calibration narrative	Very good	Satisfactory to very good	Very good
No. of phosphorus validation gages	1	3	8 [a]
Phosphorus validation statistics	(too few observed data)	0.62 < NSE < 0.81	0.59 < NSE < 0.83 0.41 < RSR < 0.64 -10% < RE < 325%
Phosphorus validation narrative	--	Satisfactory to very good	Satisfactory to very good

Sources: IDEM 2017; Ohio EPA 2014, 2021.

Notes

NSE = Nash-Sutcliffe coefficient; RE = magnitude of relative average error (also known as PBIAS); RSR = root mean square error observations standard deviation ratio.

a. One or more monitoring sites was sampled by multiple entities. The data from multiple entities collected at the same site were pooled for calibration and validation.

For each watershed, the daily flow output from the calibrated and validated SWAT model was used to develop load duration curves (LDCs). Simulated daily nutrient and sediment loads were also plotted on the appropriate LDC figures; however, the state agencies only calculated reductions with observed loads.

SWAT model nutrient and sediment output was also used to support the source assessment. The LDC figures were evaluated to determine under what flow conditions the simulated nutrient and sediment loads exceeded the TMDL targets. The relative distribution of unit area loads and point source loads were evaluated, including assessment of nonpoint source loads with land use and land cover spatial data.

4.3.2 Drainage Water Management Evaluation Model (U.S. EPA 2017)

The SWAT model developed for the Sandusky River watershed to support Ohio EPA's TMDL effort was updated and improved for an integrated water management study for U.S. EPA. The SWAT model was used to evaluate the impacts of five BMPs (controlled drainage, bioreactors, saturated buffers, wetlands, and two-stage ditches) on surface and subsurface nutrient and sediment loading.

The Sandusky TMDL SWAT model (2012 revision 591) was updated to revision 637 and several upgrades were made to the model inputs[16]. In this upgraded Sandusky SWAT model, tile drainage was assumed for agricultural and hay HRUs with low slope (0-3%) and at least 50% hydric soils in fields no less than 5 acres.

Simulation of controlled drainage and saturated buffers required modification of the SWAT code.

- **Controlled drainage**: The SWAT 2012 code was modified to allow the depth of the drainage outlet to vary over the year. SWAT simulates tile drain flow based upon the height of the water table above the drain tile; therefore, the SWAT code was modified to simulate tile drain flow based upon the height of the water table above the weir.
- **Saturated buffers**: The subroutine *filter.f* that represents vegetated filter strips was modified to simulate saturated buffers. The first modification was to route tile flow (instead of surface flow) through the filter strip. The second modification was to apply the nutrient and sediment reduction to flow as well. The third modification was to route the reduced load and flow to shallow groundwater (in lieu of permanent loss). Load and flow in shallow groundwater can return to the stream.
- **Wetlands**: The SWAT 2012 code was modified to allow wetlands to receive tile flow (in addition to surface flow).
- **Two-stage ditches**: Grassed waterways, a BMP in SWAT, were used to simulate two-stage ditches.

[16] The digital elevation model was upgraded from 30-meter resolution to 10-meter resolution, minor land uses were aggregated (e.g., deciduous forest, evergreen forest, and mixed forest were aggregated to forest), similar soil types were aggregated, three slope classes (0-5%, 5-10%, and >10%) were reduced to two slope classes (0-3% and >3%), the simulation timeframe was modified, and gridded daily weather replaced point-based weather data.

The upgraded SWAT model was used to simulate four levels of BMP implementation (25%, 50%, 75%, and 100%) on suitable drainage area (i.e., field with tile drainage). The following conclusions were drawn after evaluating scenario results for those BMPs that affected phosphorus (i.e., not bioreactors):

- **Controlled drainage**: "The simulations suggest a net decrease in tile flow but show increases in surface runoff, lateral flow and groundwater flow. Nutrient loads associated with tile flow also decrease but those associated with other flow pathways generally increase." (U.S. EPA 2017, p. 20). The largest reductions (of the five BMPs) of DRP were simulated with controlled drainage; however, controlled drainages also resulted in a large increase in sediment-bound phosphorus.
- **Saturated buffers**: SWAT simulated reduced streamflow and DRP. Recharge to the shallow aquifer increased, resulting in higher shallow aquifer outflow.
- **Wetlands**: One wetland was simulated for each model subbasin, and the wetland was sized to be 1% of the area of the model subbasin. Multiple scenarios were run with between 25% and 100% of the upstream drainage area routed through the subbasin wetland. For the different scenarios, annual runoff decreased between 1% and 12% and TP loads decreased between 2% and 12% as measured at the outlet of the Sandusky River.
- **Two-stage ditches**: SWAT simulated reductions with nutrient loads.

Combinations of BMPs were also simulated, and the simulations relied on a GIS-based opportunity analysis. Results "suggest that increases in nitrogen and sediment bound phosphorus under the controlled drainage practice may be offset using wetlands and two-stage ditches" (U.S. EPA 2017, p. 48). At the subbasin-level, all the combinations reduced DRP loads. With regards to sediment-bound phosphorus, at the subbasin-level, most scenarios suggested an increase or small decrease; only one combination decrease sediment-bound phosphorus appreciably (controlled drainage + bioreactor + wetland + two stage ditch). At the outlet of the Sandusky River, only one scenario (controlled drainage + bioreactor + wetland + two stage ditch) predicted reductions in flow and TP.

5.0 REFERENCES

Apostel, A., M. Kalcic, A. Dagnew, G. Evenson, J. Kast, K. King, J. Martin, R. L. Muenich, and D. Scavia. 2021. Simulating internal watershed processes using multiple SWAT models. *Science of the Total Environment* 759: 143920. https://www.sciencedirect.com/science/article/abs/pii/S0048969720374519. Accessed March 2, 2022.

Cho, J., G. Vellidis, D.D. Bosch, R. Lowrance, and T. Strickland. 2010. Water quality effects of simulated conservation practice scenarios in the Little River Experimental watershed. *Journal of Soil and Water Conservation* 65(6): 463-473.

Christopher, S.F., J.L., Tank, U.H., Mahl, H. Yen, J.G. Arnold, M.T. Trentman, S.P. Sowa, M.E. Herbert, J.A. Ross, M.J. White, and T.V. Royer. 2017. Modeling nutrient removal using watershed-scale implementation of the two-stage ditch. *Ecological Engineering* 108: 358-369.

Evenson, G.R., M. Kalcic, Y. Wang, D. Robertson, D. Scavia, J. Martin, A. Apostel, N. Aloysius, J. Kast, H Kujawam, C. Boles, T. Redder, R. Confesor, T. Guo, R.L. Muenich, A.T. Dagnew, A. Murumkar, and M. Brooker. 2021. Uncertainty in critical source area predictions from watershed-scale hydrologic models. *Journal of Environmental Management* 279: 111506.

GLWQA (Great Lakes Water Quality Agreement). 2015. *Recommended Phosphorus Loading Targets for Lake Erie*. Annex 4 Objectives and Targets Task Team Final Report to the Nutrients Annex Subcommittee. May 11, 2015.

Hanief, A. and A.E. Laursen. 2019. Meeting updated phosphorus reduction goals by applying best management practices in the Grand River watershed, southern Ontario. *Ecological Engineer*ing 130: 169-175.

Her, Y., I. Chaubey, J. Frankenberger, and D. Smith. 2016. Effect of conservation practices implemented by USDA programs at field and watershed scales. *Journal of Soil and Water Conservation* 71 (3): 249-266.

IDEM (Indiana Department of Environmental Management). 2017. *St. Joseph River Watershed Indiana TMDLs*. Prepared for U.S. EPA Region 5 and IDEM; prepared by Tetra Tech. August 8, 2017. https://www.in.gov/idem/nps/resources/total-maximum-daily-load-reports/st-joseph-river-lake-erie-watershed-tmdl/. Accessed March 15, 2022.

Kalcic, M.M., C. Kirchoff, N. Bosch, R.L. Muenich, M. Murray, J.G. Gardner, and D. Scavia. 2016. Engaging Stakeholders To Define Feasible and Desirable Agricultural Conservation in Western Lake Erie Watersheds. *Environ. Sci. Technol*. 50 (15): 8135-8145. http://scavia.seas.umich.edu/wp-content/uploads/2016/08/Kalcic-et-al.-2016.pdf. Accessed February 25, 2022.

Kast, J.B., A.M. Apostel, M.M. Kalcic, R.L. Muenich, A. Dagnew, C.M. Long, G. Evenson, and J.F. Martin. 2021a. Source Contribution to Phosphorus Loads from the Maumee River Watershed to Lake Erie. *Journal of Environmental Management* 279 (111803). https://www.sciencedirect.com/science/article/pii/S030147972031728X. Accessed March 1, 2022.

Kast, J.B., M. Kalcic, R. Wilson, D. Jackson-Smith, N. Breyfogle, and J. Martin. 2021b. Evaluating the efficacy of targeting options for conservation practice adoption on watershed-scale phosphorus reductions. Water Research 201 (2021): 117375.

Lu, S., H.E. Andersen, H. Thodsen, G.H. Rubæk, and D. Trolle. 2016. Extended SWAT model for dissolved reactive phosphorus transport in tile-drained fields and catchments. *Agricultural Water Management*, 175: 78-90.

Martin, J.F., M.M. Kalcic, N. Aloysius, A.M. Apostel, M.R. Brooker, G. Evenson, J.B. Kast, H. Kujawa, A. Murumkar, R. Becker, C. Boles, T. Redder, R. Confesor, T. Guo, A. Dagnew, C.M. Long, R. Meunich, D. Scavia, Y. Wang, and D. Robertson. 2019. *Evaluating management options to reduce Lake Erie algal blooms with models of the Maumee River watershed*. The Ohio State University.

Martin, J.F., M.M. Kalcic, N. Aloysius, A.M. Apostel, M.R. Brooker, G. Evenson, J.B. Kast, H. Kujawa, A. Murumkar, R. Becker, C. Boles, R. Confessor, A. Dagnew, T. Guo, C.M. Long, R.L. Muenich, D. Scavia, T. Redder, D.M. Robertson, and Y. Wang. 2021. Evaluating management options to reduce Lake Erie algal blooms using an ensemble of watershed models. *Journal of Environmental Management* 280: 111710.

Muenich, R.L., M. Kalcic, and D. Scavia. 2016. Evaluating the impact of legacy P and agricultural conservation practices on nutrient loads from the Maumee River Watershed. *Environ. Sci. Technol*. 50 (15): 8146-8154. https://scavia.seas.umich.edu/wp-content/uploads/2016/08/Muenich-et-al.-2016.pdf. Accessed March 10, 2022.

Merriman, K.R., P. Daggupati, R. Srinivasan, and B. Hayhurst. 2019. Assessment of site-specific agricultural Best Management Practices in the Upper East River watershed, Wisconsin, using a field-scale SWAT model. *Journal of Great Lakes Research* 45: 619-641.

Ohio EPA (Ohio Environmental Protection Agency). 2014. *Total Maximum Daily Loads for the Sandusky River (lower) and Bay Tributaries Watershed*. Final Report. Prepared for U.S. EPA Region 5 and Ohio EPA; prepared by Tetra Tech. May 28, 2014. https://epa.ohio.gov/divisions-and-offices/surface-water/reports-data/sandusky-river-and-sandusky-bay-tributaries-watersheds. Accessed March 15, 2022.

-----. 2021. *Black River Watershed TMDLs*. Draft Report. Prepared for U.S. EPA Region 5 and Ohio EPA; prepared by Tetra Tech.

Radcliffe, D.E., D.K. Reid, K. Blombäck, C.H. Bolster, A.S. Collick, Z.M. Easton, W. Francesconi, D.R. Fuka, H. Johnsson, K. King, M. Larsbo, M.A. Youssef, A.S. Mulkey, N.O. Nelson, K. Persson, J.J. Ramirez-Avila, F. Schmieder, and D.R. Smith. 2015, Applicability of models to predict phosphorus losses in drained fields: A review. *J. Environ. Qual.* 44: 614-628. https://doi.org/10.2134/jeq2014.05.0220.

Scavia, D., M. Kalcic, R. Logsdon-Muenich, N. Aloysius, C. Boles, R. Confesor, J. DePinto, M. Gildow, J. Martin, J. Read, T. Redder, D. Robertson, S. Sowa, Y. Wang, and H. Yen. 2016. *Informing Lake Erie Agriculture Nutrient Management via Scenario Evaluation*. Final Report. University of Michigan, Water Center. Ann Arbor, Michigan.

http://graham.umich.edu/media/pubs/InformingLakeErieAgricultureNutrientManagementviaScenarioEvaluation.pdf. Accessed December 15, 2021.

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

U.S. EPA (United States Environmental Protection Agency). 2017. *Integrated Water Management Strategy: Integrated Water Quality and Water Quantity Benefits in Select Watersheds Draining to Lake Erie and Saginaw Bay*. Prepared for U.S. EPA Region 5; prepared by Tetra Tech. March 31, 2017.

Williams, M. R., K.W. King, W. Ford, A.R. Buda, and C.D. Kennedy. 2016, Effect of tillage on macropore flow and phosphorus transport to tile drains, *Water Resour. Res.* 52: 2868–2882. doi:10.1002/2015WR017650.

Yang, W., Y. Liu, J. Simmons, A. Oginskyy, K. Mckague, R. Wilson, A. Gutteridge, T. Mcpherson, H. Brock, and K. Monk. 2013. *SWAT Modelling of Agricultural BMPs and Analysis of BMP Cost Effectiveness in the Gully Creek Watershed*. University of Guelph. Guelph, Ontario, Canada. December 15, 2017. https://www.abca.ca/downloads/2017_GLASI_SWAT_Modelling_Report_Gully_Creek_Watershed_Web_RE.pdf. Accessed February 2, 2022