



Assessing nutrient and sediment load reduction potential of vegetation by utilizing the nutrient tracking tool at the field and watershed scale in a Great Lakes priority watershed

Ethan D. Pawlowski¹ · Diana L. Karwan¹ · Randall K. Kolka²

Received: 13 June 2022 / Accepted: 10 April 2023 / Published online: 18 May 2023
© The Author(s) 2023, corrected publication 2024

Abstract

Purpose Nutrient and sediment pollution of surface waters remains a critical challenge for improving water quality. This study takes a user-friendly field-scale tool and assesses its ability to model at both the field and watershed scale within the Fox River Watershed (FRW), Wisconsin, USA, along with assessing how targeted vegetation implementation could attenuate nutrient and sediment exports.

Methods To assess potential load reductions, the nutrient tracking tool (NTT) was used with a scoring system to identify areas where vegetation mitigation could be implemented within three selected FRW sub-watersheds. A corn soybean rotation, an implementation of a 10-m-vegetated buffer, a full forest conversion, and tiling were modeled and assessed. The corn–soybean results were aggregated and compared to watershed level gauge data in two sub-watersheds. Edge-of-field data was compared to modeled results using multiple parameterization schemes.

Results The agricultural areas that scored higher and were untilled showed greater potential nutrient and sediment export reduction (up to 80 to 95%) when vegetation mitigation was implemented in the model. Field-scale results aggregated to the watershed scale showed disparities between modeled and measured phosphorus exports but modeled sediment exports fell within observed gauge data ranges. Field-specific parameter adjustments resulted in more accurate modeled results compared to measured edge-of-field export data but needed further refinement.

Conclusion Targeted mitigation using a vegetation-based scoring system with the NTT model was shown to be a helpful tool for predicting nutrient and sediment reductions. Using a field-scale model aggregated to the watershed scale presents tradeoffs regarding processes found beyond the edge of field.

Keywords Nutrient tracking tool · Phosphorus · Sediment · Nitrogen · Field scale · Watershed scale

1 Introduction

Soil and water degradation from agriculture remains an important concern for sustainable agriculture and global food security. Chemical fertilizer use paired with common

agricultural practices have created large nutrient imbalances in agricultural soils that can be transported from the field in soluble and particulate forms to downstream waterbodies (Tilman et al. 2002, 2011; Alexander et al. 2008; Thompson et al. 2011). Eutrophication caused from the excess delivery of phosphorus (P), nitrogen (N), and sediment loads are particularly detrimental to aquatic ecosystem health, drinking water supplies, and outdoor recreation. These eutrophic areas are expected to increase globally and throughout the Great Lakes (Selman et al. 2008; Dodds et al. 2009).

Revegetation, herein defined as the process of replanting vegetation and rebuilding the soil in disturbed sites, provides an effective mitigation strategy to reduce non-point source nutrient and sediment loads from agricultural areas (Reddy

Responsible editor: Alexander Koiter

✉ Ethan D. Pawlowski
pawlo026@umn.edu

✉ Diana L. Karwan
dlkarwan@umn.edu

¹ University of Minnesota Twin Cities, St. Paul, MN 55108, USA

² USDA Forest Service, Northern Research Station, Grand Rapids, MN 55744, USA

et al. 1999; Powers et al. 2012; Schulte et al. 2017; Keller and Fox 2019). Additional vegetation promotes increased biodiversity above and below the soil surface as well as the accumulation of organic matter which aids in aggregate and macropore formation. These features improve water infiltration rates and storage that reduce nutrient and sediment loads stemming from overland flow. The soil aggregates are less likely to erode and increase nutrient use efficiency by an intact network of plant, microbial, and mycelium (Tilman et al. 2002; Schulte et al. 2017; Rubin and Görres 2021). The key to obtaining these benefits is through strategic and targeted implementation that utilizes physical characteristics of individual fields such as contours, soil erodibility, areas of marginal production, and drainage characteristics to efficiently place vegetation to intercept sediment and nutrients (Asbjørnsen et al. 2014; Schilling and Jacobson 2014; Schulte et al. 2017; Lenhart et al. 2018; Carstensen et al. 2020; Jiang et al. 2020; Prosser et al. 2020). These benefits are achievable through the incorporation of a wide range of different vegetative species (Prosser et al. 2020); to narrow this scope, we focus on native forest and prairie plant communities and consider other vegetative management later in the discussion. There are added benefits to using native plant communities besides reducing sediment and nutrient erosion; the Science-based trials with Row-crops Integrated with Prairie Strips (STRIPS) experiment in Iowa showed that using native prairie communities had significantly greater benefits to native insect, pollinator, and bird species abundance than non-native species planted with less biodiversity such as in some cover cropping systems (Schulte et al. 2017). Even with using native species, a targeted approach matching physical site characteristics with plant species that will thrive in those conditions is pivotal for successful sediment and nutrient reduction (Prosser et al. 2020). A major tradeoff of using all native vegetation is the reduction in crop production acreage. However, studies have shown that even modest land conversions of 10% have significant benefits on water quality and soil health with further increases in biodiversity following native perennial vegetation establishment (Schulte et al. 2017). Thus, the assessment and identification of areas where the conservation of, or land conversion to, native vegetation is important for water quality management (Selman et al. 2008; Dodds et al. 2009).

Utilizing physical landscape characteristics is pivotal to identifying areas for revegetation of native plant communities. Using the landscape slope and hydrologic drainage class provides a critical first step in identifying areas at high risk of nutrient and sediment loss under traditional row-cropping rotations. Hydrologic drainage class divides soils into four main groups (A, B, C, and D) based on runoff potential and infiltration capacity. Landscape slope and hydrologic drainage class form the basis for the USDA Soil Vulnerability Index (SVI; Lohani et al. 2020) and the GIS-based

Agricultural Conservation Policy Framework (ACPF; Porter et al. 2016), two tools that were designed to identify areas susceptible to nutrient and sediment losses in agricultural settings. The SVI additionally incorporates the soil erodibility *K*-factor, coarse fragments and organic carbon content which improved the index site differentiability in areas with heterogeneous slopes and soils but are less important than slope and hydrologic drainage class in areas with greater homogeneity (Lohani et al. 2020). The ACPF identifies areas of probable overland flow using hydrologic routing (based on soils and slopes) and notes that proper validation of the slope and drainage characteristics are vital to proper on-site mitigation planning (Porter et al. 2016). Other common models that predict nutrient and sediment export often rely on the curve number approach (Neitsch et al. 2002; Texas A&M AgriLife Research 2013) which is an empirical approach that uses slope, hydrologic drainage class, and land cover characteristics (Ponce and Hawkins 1996). These examples use a host of physical characteristics to estimate sediment and nutrient export processes but the utilization of slope and drainage characteristics to effectively model those export processes are the key parameters behind the SVI, ACPF, and other models.

Models have been used to simulate conservation practice implementation effectiveness at reducing nutrient and sediment export (US Environmental Protection Agency 2018). These models work at varying spatial scales from watersheds to individual fields and have vastly different data needs, treatments of hydrologic routing, degrees of sophistication, incorporation of management decisions, and ease of use (US Environmental Protection Agency 2018). To accurately represent revegetation with native species, an appropriate model must be able to incorporate vegetative changes and agricultural practices, as well as model nitrogen, phosphorus and sediment dynamics. The Soil Water Assessment Tool (SWAT) and the Agricultural Policy Environmental eXtender Model (APEX) are two models that satisfy these requirements, but at differing scales—watershed and field, respectively (Texas A&M AgriLife Research 2013; USDA ARS Grassland Soil and Water Research and Texas A&M AgriLife Research 2018). Models designed for a specific spatial scale may have difficulty simulating results relevant to another scale because the key processes could differ, or processes may even be absent at different spatial scales. Conservation practices are often applied and implemented to achieve field-scale results based on individual management decisions, making field-scale models appropriate for simulating those decisions at a local scale (Sharpley et al. 2009; Tomer and Locke 2011). Despite implementation at the field scale, water quality outcomes of practices are often evaluated at the watershed scale with water quality monitoring efforts and is the scale at which mitigation plans such as total maximum daily loads (TMDLs) are written.

We focused on using a field-scale model for this research to match conservation practices that are implemented within individual agricultural fields. The nutrient tracking tool (NTT) provides a user-friendly web-based graphical user interface to run the Agricultural Policy Environmental eXtender (APEX; Texas A&M AgriLife Research 2013). The APEX model has been applied and evaluated across the globe and shown to be an effective field-scale model for estimating the export of nutrients and sediment (Keller and Fox 2019 and references within). Nutrient tracking tool results have primarily been verified in watersheds flowing into Lake Erie (Saleh et al. 2015), but the model is in various stages of verification across the USA. Despite limited validation, a user-friendly model was a key attribute in deciding to move forward with using the NTT. Conservation practice implementation and investment from private landowners is increased when the process and benefits are openly communicated and easy to understand with input contributed from both the private landowner and the conservationist practitioner (Brant 2003; Rose et al. 2018; Walsh et al. 2019). Although not a part of this research, the NTT provides a way for landowners to be actively involved in the modeling process in the future resulting in reduced barriers to conservation practice implementation (Brant 2003).

In this study, we evaluated the NTT prediction of field-scale nutrient and sediment load reductions due to revegetation efforts in select sub-watersheds within the Lower Fox River Watershed (FRW), Wisconsin (WI), USA. We perform preliminary analysis to focus revegetation implementation

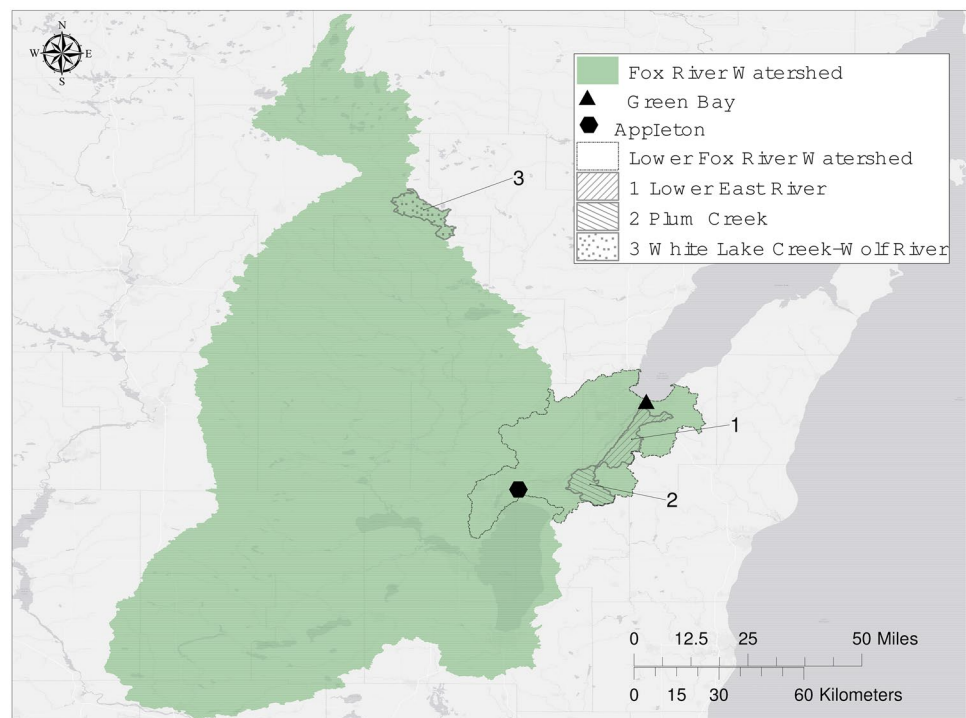
on agricultural lands with low infiltration capacity and high slopes, following the framework developed by Keller and Fox (2019). These lands are generally less productive and more suitable to target for conservation practice implementation. Upon identification of areas with high slope and poor drainage, the NTT was used to simulate and evaluate the load reductions resulting from the addition of buffer strips between agricultural crops and adjacent waterways, conversion to forest, and the effects of tiling on load reductions. Model outputs utilizing traditional and observed cropping systems were juxtaposed with measured edge of field and watershed gage export data. Considerations for widespread NTT application in the Lower FRW and other Great Lakes Priority Watersheds are also discussed.

2 Methods

2.1 Site description

The FRW is 1,701,397 ha in size and located in North-east WI (Fig. 1). According to the 2016 National Land Cover Dataset (NLCD), the area's primary land uses are comprised of approximately 36% cultivated cropland, 40% undeveloped land, and 8% urban/suburban (MRLCC 2016). The climate is temperate with average high temperatures around 27 °C in July and August, and an average high of −5 °C in January. In Green Bay, WI, at the FRW outlet, the average annual precipitation is 750 mm with an average

Fig. 1 Fox River Watershed (FRW) and the Lower FRW. Three sub-watersheds are highlighted and include (1) the Lower East River (LER), (2) Plum Creek (PC), and (3) White Lake Creek–Wolf River (WLCWR)



annual snowfall deposition of 1295 mm (as snow) during the colder months; the monthly rainfall peaks in June with 99 mm (U.S. Climate Data 2021). Peak river discharge correlates with the snowmelt period each spring (Graczyk et al. 2011; Jacobson 2012). Past glaciation is reflected in the watershed's topography, soils, and geomorphic features (Hadley and Pelham 1976; Luczaj 2011, 2013). Glacial features range from outwash plains to lacustrine plain deposits, but ground and end moraines are predominant. Drainage characteristics range from well drained in the sandy outwash plains to very poorly drained in the lacustrine deposits common near the watershed outlet (Provost et al. 2001; Santy et al. 2001; Natural Resources Conservation Service 2008a, b). Tile drainage is largely unknown in the greater watershed because tiles are implemented on private property, but attempts have been made to quantify tiling in areas with excessive nutrient and sediment pollution (S. Kussow, personal communication, 08/04/2020).

2.2 Modeling framework

The Lower FRW has been designated an USEPA priority watershed and was assigned targeted P and sediment load reductions through the Total Maximum Daily Load (TMDL) program to meet water quality goals beginning in 2012. Three sub-watersheds, the Lower East River (LER), Plum Creek (PC), and White Lake Creek–Wolf River (WLCWR) were used for the nutrient and sediment load quantification using the NTT (Fig. 1). The LER and PC have been identified as major contributors to the P and sediment loads impairing the Lower FRW. Agricultural and developed land uses cover 50% and 36%, respectively, of the 11,600 ha LER sub-watershed (Online Resource Fig. S1; MRLCC 2016). Cultivated crop land covers 76% of the 9,030 ha PC watershed, with the remainder comprised of 9% developed land and 10% as vegetated land uses such as forests and wetlands (Online Resource Fig. S2). Both sub-watersheds are situated in an area defined by glacial lacustrine deposits (Hadley and Pelham 1976) with lower infiltration rates and high erodibility relative to other soils in the watershed (Soil Survey Staff 2020b). The WLCWR sub-watershed is located outside of these deposits and 84% of its land cover is in forested, shrubby, or wetland vegetation (Online Resource Fig. S3). The WLCWR was included as a stark contrast to compare NTT results from the more developed landscapes found in the LER and PC and to show the benefits of conserving areas not presently in agricultural land use.

2.2.1 Priority scoring areas for revegetation

The entire FRW was scored to identify areas where revegetation, such as reforestation, could improve water quality based on soil and slope properties (Fig. 2; Keller and Fox

2019). Our scoring system, based upon an extended numerical range of Keller and Fox (2019), assigned a value (1–9) to each 10-m-grid cell based on slope and soil hydrologic group rating as published in the USDA Soil Survey Geographic database (SSURGO; Table 1). Slope was calculated between grid cells using a 10-m U.S. Geological Survey DEM (U.S. Geological Survey 2018). Higher slopes and poor infiltration were assigned a higher component score (Table 1). The total priority score for each 10-m grid cell was computed as the sum of the hydrologic group and slope component scores. Although other soil properties may have been useful, infiltration capacity paired with slope are essential components of other vulnerability tools (Porter et al. 2016; Lohani et al. 2020), and thus provide a simple but effective way to identify soils vulnerable to nutrient and sediment export.

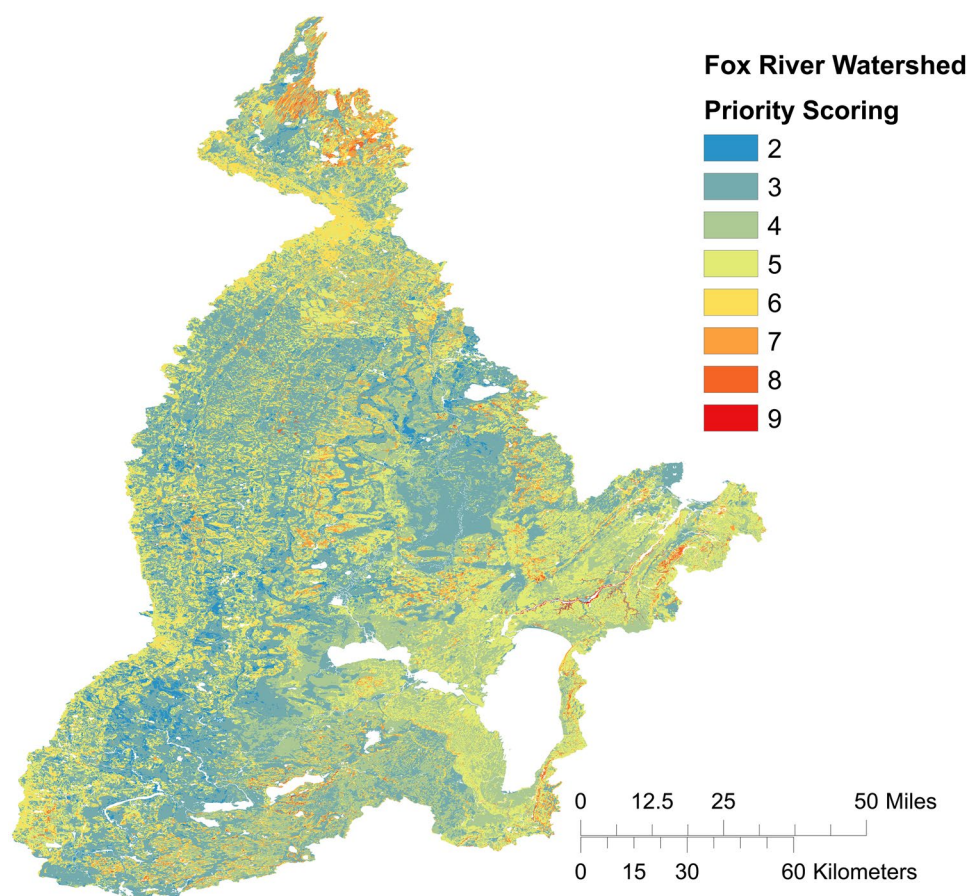
2.2.2 Scored area selection criteria for use with the NTT

The NTT was used to evaluate potential nutrient load reduction to waterways from agricultural fields and forested land within targeted FRW sub-watersheds. Waterways were identified using the National Hydrography Dataset (U.S. Geological Survey 2020). Only areas in agricultural land use according to the most recent NLCD (Online Resource Figs S1, S2 and S3; MRLCC 2016) with a total priority score greater than or equal to six (“Sect. 2.2.1”) that met all the following criteria were identified as high priority for revegetation and modeled within the NTT:

- In a row crop land use per the 2016 NLCD
- Within 500-m of stream channels
- Aggregated scoring area greater than 0.4 ha (40 grid cells) in size

Untiled areas meeting the above criteria were aggregated based on grid cell score. Priority areas were allowed to be larger or smaller than an agricultural field defined by landownership boundaries. The untiled high priority areas were converted to a vector file for import to the NTT. Untiled fields were evaluated separately from tiled fields. The above criteria were used to select tiled fields, but the tiled fields were not scored. Tiled fields were treated as one hydrological unit due to the connection of subsurface drainage over an unknown region of varying priority scores. Tiled fields were mapped as a vector file using GIS software by Outagamie County Land Conservation Department by identifying fields with tile inlets based on aerial imagery; hence, tile presence and proximity to the stream channel was known (S. Kussow, personal communication, 08/04/2020), but tile outlet location, network connection, function, and condition within individual fields were unknown.

Fig. 2 Priority scoring results for the FRW based on slope and soil hydrologic drainage class



2.2.3 Modeled nutrient and sediment load reductions utilizing the NTT

Upon selection based on scoring and location, field vector files for both tilled and untilled areas (referred to as fields hereafter) were uploaded and identified within the spatial interface of NTT and a field list was created. In the NTT interface, the APEX model is paired with U.S. Department of Agriculture (USDA) soil property databases and climate data from the Parameter-elevation Regressions on Independent Slopes Model (PRISM; PRISM Climate Group 2004; Soil Survey Staff 2020a) to simulate sediment and nutrient export for the spatial extent of each field.

Following field identification, the NTT uses the USDA SSURGO database to obtain soil properties for each individual field (Soil Survey Staff 2020a). The NTT ranks and selects the three-soil series with the greatest percent coverage within the field and will adjust the area of the

three selected soils to encompass the whole field based on a weighted average. The NTT then uses the physical and chemical characteristics from the USDA SSURGO database with the 10-m DEM to calculate an average value of each characteristic for each soil series. Both physical and chemical data such as porosity, field water capacity, saturated hydraulic conductivity, erosivity, and slope are used by equations within the model (described and referenced below). Once the weather data and management scenarios are defined (explained below) the NTT will use APEX to simulate exports for each soil series and use a weighted average based on spatial extent to generate the final field exports.

NTT generated results by incorporating PRISM (Parameter-elevation Regressions on Independent Slopes Model; PRISM Climate Group 2004) data to model field runoff and water quality constituent export within NTT on a daily basis. The PRISM combines spatial climate knowledge with point data from weather stations, digital elevation models,

Table 1 Scores using % slope and hydrologic soils group to indicate areas for revegetation

% Slope	> 25	15–25	7–15	4–7	2–4	< 2	Hydrologic soil group	C, D, C/D	B, B/D, A/D	A
Component Score	6	5	4	3	2	1		3	2	1

and other spatial data with regression to generate annual, monthly, daily, and event-based climatic estimates that are interpolated at an approximately 800-m grid cell size (PRISM Climate Group 2004). The interpolated climatic values of each 800-m grid are assigned to the NTT fields based on spatial overlap.

APEX incorporates a large array of empirical and theoretically based equations to generate export estimates. Descriptions of these equations were reviewed in Gassman et al. (2009). Equations that represent all aspects of the hydrologic cycle, erosion, plant growth, aspects of livestock management, nutrient cycles, and routing are used and represented within the model (Gassman et al. 2009 and references therein). Management practices are also incorporated and modify the model output. These equations have an influence over the temporal discretization used by the model resulting in daily calculations. Although the APEX output files are available to download, the NTT interface summarizes this output and has it available for download as annual averages.

Just like the base version of APEX, the NTT is designed to evaluate the effectiveness of several conservation practices, including reforestation, on nutrient export from fields to surface waters (Saleh et al. 2015). Four main scenarios were set up:

- (Scenario 1) a corn–soybean rotation (the business-as-usual case).
- (Scenario 2) a corn–soybean rotation with 10-m vegetated buffer along waterways
- (Scenario 3) conversion to a red pine plantation (currently the only forest option)
- (Scenario 4) elimination of tiling where present

Tillage, fertilization, planting, and harvesting practices were incorporated into these scenarios as described in Keller and Fox (2019). Tillage consisted of chisel plowing in the spring, field cultivating prior to planting, and disking post-harvest. Fertilizer was applied in the spring; planting occurred in early May for corn and mid-May for soybeans with harvests set for October (Online Resource Table S1). The tiled fields used the same scenarios with the addition of tiling specified for each field in NTT. One downside of NTT version 20–2 used in this study is that red pine plantation is the only included forest type, which may not be suitable for all locations (Saleh et al. 2011). A subset of forested land within the WLCWR watershed was simulated with the NTT to simulate the conversion of forested land to row crop agriculture using the scenarios listed above. The simulation results from the WLCWR are labeled as low priority regardless of scoring criteria because the 2016 NLCD dataset classifies the land cover as forested, shrubby, or wetlands.

We used the NTT to simulate annual field export of total nitrogen (TN), total phosphorus (TP), and sediment over a

35-year period (1982–2018). NTT model results break the TN into runoff N, tile drain N, organic N, and subsurface N components. TP is separated into organic, soluble reactive P (orthophosphate), and tile drain P loads. Each of the individual components makes up the total N or P values. Sediment export is defined as surface erosion and not broken down into other components.

A randomly selected subset of 15 fields were selected from tiled, high priority, and mid-priority (score of 4 and 5, “Sect. 2.2.1”) as areas for sensitivity testing of NTT results to model parameters. Sensitivity testing was conducted on P export results from the NTT among the fields in the PC watershed. Mid-priority areas were included in order to complete the watershed-scale evaluation (described below) and were added to the sensitivity analysis after completing that evaluation. A mid-priority area is one that may have a relatively steep slope or low infiltration capacity that resulted in a score of 4 or 5 but did not represent combination of both low infiltration and steep slopes such as areas ranked as high priority (scores of 6 or greater). The soluble P runoff coefficient and the P enrichment ratio coefficient for routing were left at the default values of 20 ($0.1 \text{ m}^3 \text{ Mg}^{-1}$) and 0.78 respectively. Other parameters that were included in the sensitivity analysis were the soluble P runoff exponent, the P enrichment ratio exponent for routing, and the soluble P leaching equilibrium dissociation constant value. These parameters are used in the NTT model equations to generate the nutrient export results. Parameters were adjusted to maximize P export from NTT simulations at the field scale. Using the changed parameter values, the NTT simulated new export values for the chosen fields that were compared to the original values. The sensitivity analysis was used to check if unadjusted parameter settings caused the large discrepancy between the high and mid-priority aggregated modeled NTT exports and the sub-watershed data used during the watershed-scale evaluation (described below). Nitrogen was excluded from this exercise because TN watershed export data is unavailable for PC.

2.3 Evaluation of NTT output at the watershed and field scale

Verifying the modeled results with export data at both the field and watershed scale acts as a proof of concept that any modeled reductions from revegetation mitigation using the NTT may be an effective mitigation strategy for the FRW. The comparison between NTT modeled field exports and published watershed loads additionally checks the magnitude of the NTT model results against the measured data. Field-scale data was recently published (Komiskey et al. 2021), and the LER and the main branch of PC have annual sediment and P export

estimates at the watershed scale in association with USGS gauges (040851378 and 04084911) from 2004 to 2006 and 2012 to 2017 (Graczyk et al. 2011; Jacobson 2012; U.S. Geological Survey 2016a, b). The data collection period for these studies is significantly shorter than the average annual sediment and nutrient exports estimated from the NTT over the 35-year simulation period. The published watershed export values, or loads, from the LER and PC were compared against the manually aggregated NTT 35-year-simulated average annual sediment and P total load results in the tiled and high priority areas under the corn–soybean rotation scenario (business as usual) with the addition of mid-priority areas (scores of 4 and 5), which represented the remainder of the agricultural area in row crops located near a stream channel or drainage network. Changing the high priority, mid-priority, and tiled unit area loads into total loads and aggregating across the watershed area makes the values comparable to the published measured loads for each watershed.

For the field-scale evaluation, NTT parameters were adjusted along with the cropping system to better match the field-scale characteristics and management practices (Online Resource Table S2). The comparison was completed on a single edge-of-field monitoring site (ESW3) that had both surface and tile drain datasets that were collected from 2014 to 2019 (Komiskey et al. 2021). The NTT was run separately for the surface and subsurface export from the field because the tiled portion drains approximately 10 ha and the surface drains approximately 2.8 ha. The parameters were largely set based on previous work using the ArcApex model (Kalk 2018) and involved adjusting settings to resemble the cropping and soil systems found in the field along with incorporating a 5-year training or model warm up period prior to the 2014 to 2019 simulation. Model predictions improved with the increased temporal range. The NTT was run with the parameter values used for the watershed-scale evaluation of the NTT output and collated but the model was run with the 5-year training period and for 2014 to 2019 so that it could be compared to Kalk (2018) in addition to the measured data. After completing this initial evaluation, two additional steps were attempted to improve results. First, instead of having the initial soil P estimated by the model, the soil P was changed to the measured value (47.5 g Mg^{-1} Bray P; Kalk 2018) and the NTT was run for the same time period. Second, a longer model run of 35 years was simulated with the original parameter settings. Nash Sutcliffe efficiency (NSE) was calculated for the discharge and phosphorus values based on annual variation between modeled and observed values. A NSE value of 0.5 was used as the criteria for determining acceptable versus poor model performance.

3 Results

3.1 Priority scoring

The 10-m grid cell score based on slope and hydrologic drainage class were computed for the FRW (Fig. 2). Scoring resulted in 208,093 ha or 12% of the watershed classed as high priority areas (score ≥ 6), 700,420 ha or 41% of the watershed area as mid-priority areas (score equal to 4 or 5), and 672,888 ha, or 40% of the watershed area as low priority (score equal to 2 or 3; Fig. 2). Within the FRW, a total of 51,865 ha or 3% of the total watershed area were both scored high priority in this analysis and were classified as cultivated cropland in the 2016 NLCD.

3.2 Results of NTT field-scale simulations

3.2.1 Lower East River

The LER encompasses approximately 11,300 ha of which 43% is in row crop agriculture (Online Resource Figure S1). Of the land in row crop agriculture, tiled fields account for 28% and 12% was determined to be high priority and untilled using the scoring system (Online Resource Fig. S4), and 3% were scored as less than high priority. In total, 30 tiled fields and 214 high priority fields were simulated with the NTT. Tiled fields in a corn–soybean rotation showed an average annual loss of $(34.56 \pm 14.31) \times 10^{-3} \text{ Mg TN ha}^{-1}$, $(1.13 \pm 0.32) \times 10^{-3} \text{ Mg TP ha}^{-1}$, and $0.97 \pm 0.55 \text{ Mg sediment ha}^{-1}$ (Fig. 3; Online Resource Tables S3 and S4). High priority fields showed the greatest nutrient and sediment export ranges compared to the tiled fields (Fig. 4) as expected due to the greater quantity of fields modeled. The conversion of tiled lands to a red pine plantation decreased loads the greatest. The forest conversion resulted in 85.7, 94.7, and 89.2% reductions for TN, TP, and sediment, respectively. The untilled high priority fields showed a similar trend with forest conversion corresponding to 85.3% TN, 91.8% TP, and 80.9% sediment reductions. Buffer strips worked better at reducing sediment compared to nutrients but did not provide as great of reductions as the forest conversion scenario. The buffer strips in high priority untilled fields reduced TN, TP, and sediment by 8.9, 22.9, and 57.6%. Tiled systems reduced TN, TP, and sediment by 0.5, 6.0, and 41.5% respectively (Fig. 3; Online Resource Tables S3 and S4).

3.2.2 Plum Creek main branch

The agricultural land in Plum Creek is 31% tiled and 2% scored as high priority (Online Resource Fig. S5). The NTT

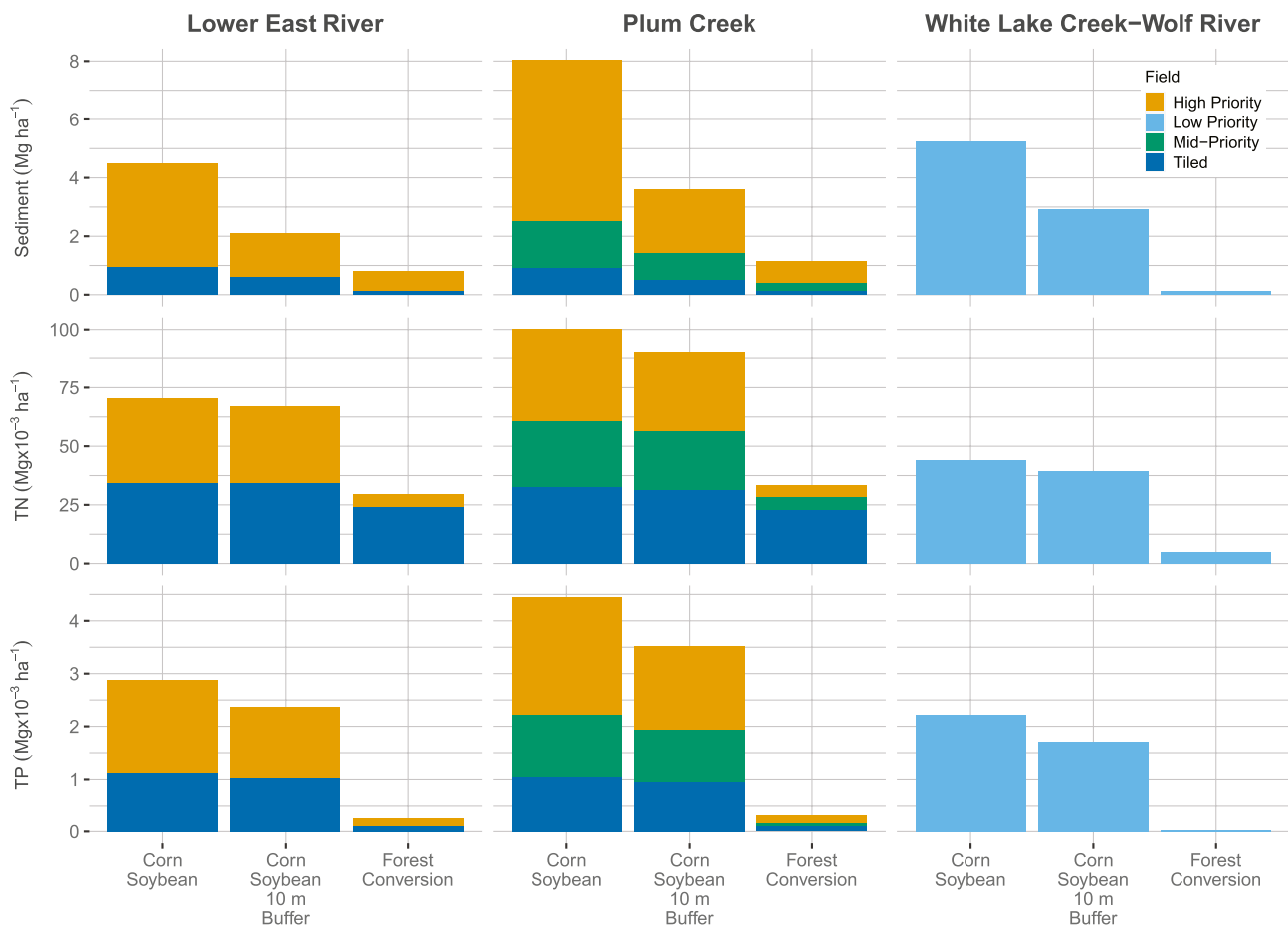


Fig. 3 Nutrient tracking tool (NTT) modeling results for sediment, total nitrogen (TN), and total phosphorus (TP) in the Lower East River (LER), Plum Creek (PC), and White Lake Creek–Wolf River (WLCWR) for all, high priority, low priority, mid-priority, and tiled fields. The values rep-

resent the average annual unit area export of all fields in the respective watersheds simulated for 35 years. The LER had 30 tiled and 214 high priority fields; PC had 47 tiled, 110 high priority, and 104 mid-priority fields; WLCWR had 5 low priority fields evaluated

was used to simulate 47 tiled, 110 high priority, and 104 mid-priority fields. The corn–soybean rotation in tiled fields were projected to yield $(32.70 \pm 14.07) \times 10^{-3}$ Mg TN ha^{-1} , $(1.05 \pm 0.30) \times 10^{-3}$ Mg TP ha^{-1} , and 0.93 ± 0.55 Mg sediment ha^{-1} in annual losses. The corn–soybean rotation in the high priority fields resulted in $(39.79 \pm 17.04) \times 10^{-3}$ Mg TN ha^{-1} , $(2.25 \pm 0.72) \times 10^{-3}$ Mg TP ha^{-1} , and 5.50 ± 2.48 Mg sediment ha^{-1} losses on an annual basis and had the greatest range in values compared to other management scenarios, tiled, and mid-priority fields (Fig. 4). Similar to the LER, tile removal and conversion to forest resulted in the greatest nutrient and sediment reductions with 85.4, 95.2, and 86.2% reductions for TN, TP, and sediment, respectively (Fig. 3; Online Resource Tables S5, S6, and S7). Forest conversion in the non-tiled high priority areas showed similar results with 87.2% TN, 92.9% TP, and 86.2% sediment reductions. Tile drainage that circumvents buffer strips increased TN and TP loads by 27.2% and 6.7% versus fields that incorporated buffers without tile drainage. Sediment export was

reduced by 18.8% in the tiled scenario versus when tiles were removed (Fig. 3; Online Resource Table S5).

3.2.3 White Lake Creek–Wolf River

The WLCWR has 84% of the 10,522 ha in forested and wetland vegetation (Online Resource Fig. S3). Only five of these low priority fields were simulated with the NTT, but the 5 fields were spatially comparable to the other sub-watersheds. The modeled forested landscapes in the WLCWR result in less nutrient and sediment losses to streams compared to conversion to row crop agriculture. Converting forested land uses to a corn–soybean rotation resulted in 803%, 9,800%, and 3,767% increases in TN, TP, and sediment exports to waterbodies with modeled fluxes changing from $(4.87 \pm 1.82) \times 10^{-3}$ Mg TN ha^{-1} , $(0.02 \pm 0.02) \times 10^{-3}$ Mg TP ha^{-1} , and 0.13 ± 0.09 Mg sediment ha^{-1} to $(44.05 \pm 17.05) \times 10^{-3}$ Mg TN ha^{-1} , $(2.22 \pm 0.76) \times 10^{-3}$ Mg TP ha^{-1} , and 5.25 ± 2.06 Mg sediment ha^{-1} exported annually (Fig. 3; Online Resource Table S8).

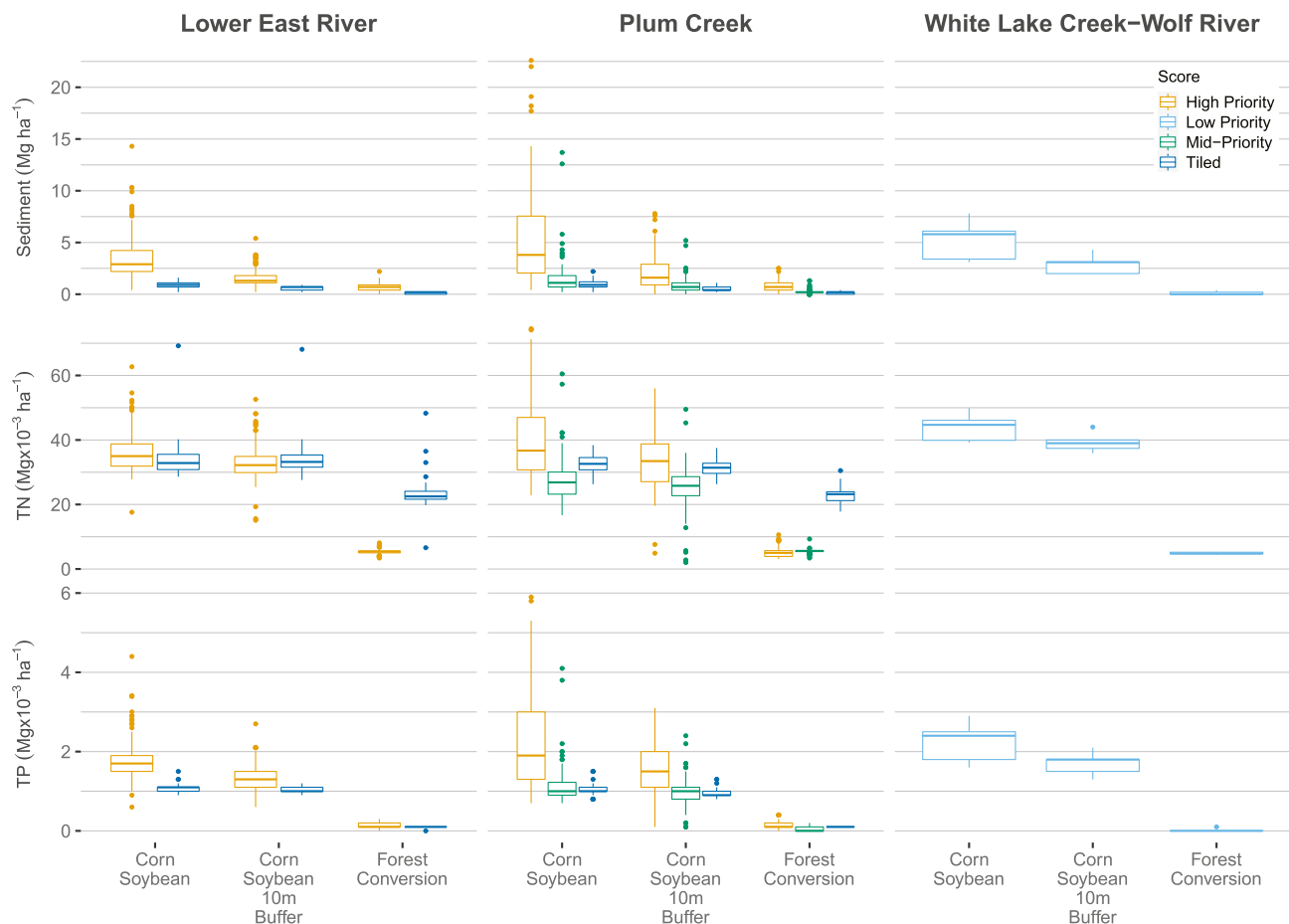


Fig. 4 Nutrient tracking tool (NTT) modeling results for sediment, total nitrogen (TN), and total phosphorus (TP) using corn soybean, forest conversion, and corn soybean with 10-m buffer management scenarios in the Lower East River (LER), Plum Creek (PC), and White Lake Creek–Wolf River (WLCWR) for all high priority, low priority, mid-

priority, and tiled fields. The LER had 30 tiled and 214 high priority fields; PC had 47 tiled, 110 high priority, and 104 mid-priority fields; WLCWR had 5 low priority fields evaluated. The distribution of values represents the average annual results of each field simulated for 35 years

3.3 Watershed- and field-scale evaluation of NTT

Overall, the NTT produced sediment exports that fell within observed ranges at the watershed scale, but TP was underestimated in the LER using the 35-year simulation with default parameter values and the corn–soybean management scenario. According to the NTT simulations, the LER tiled and high priority fields exported 2.74 ± 0.85 Mg yr⁻¹ of TP and $3,981 \pm 2,159$ Mg yr⁻¹ of sediment (Table 2). Total nitrogen was not measured at the watershed outlet, so it was excluded from the evaluation. From 2004 to 2006, the measured annual TP export reported at the LER USGS gaging station (gauge number 040851378) averaged 35.11 Mg yr⁻¹ and ranged from 17.28 to 62.60 Mg yr⁻¹ (Graczyk et al. 2011; U.S. Geological Survey 2016a). The Lower East River measured sediment export (Table 2) from the same timespan averaged 9,350 Mg yr⁻¹ and ranged from 2,422 to 20,593 Mg yr⁻¹ (Graczyk et al. 2011; U.S. Geological Survey 2016b). Comparing the NTT field export to the watershed yield indicates

the cumulative modeled field sediment export is on the same order of magnitude as the measured yield at the watershed outlet; however, annual variation in the watershed outlet is considerable, as the annual average has a standard deviation of 9,824 Mg yr⁻¹, or 105%. Comparison of the TP values indicates the NTT 35-year average annual field export is well under the TP yield measured at the watershed outlet, with the NTT field export prediction approximately one order of magnitude lower than the annual average measured at the watershed outlet over 2004–2006.

In PC, the tiled, high priority, and mid-priority fields combined were predicted to annually export 6.23 ± 1.89 Mg yr⁻¹ of TP and $8,051 \pm 4,433$ Mg yr⁻¹ of sediment based on NTT simulations (Table 2). Both the TP and sediment modeled field export fell within the ranges of observed exports at the PC gauge from 2012 to 2017 (USGS gauge 04084911): 6.12 to 18.69 Mg yr⁻¹ TP and 3,157 to 11,649 Mg yr⁻¹ of sediment (U.S. Geological Survey 2016b).

Table 2 Sub-watershed comparison of total phosphorus (TP) and sediment export from modeled fields in the Lower East River, Plum Creek, and Edge-of-field ESW3 to USGS data at the watershed and field scales respectively. Watershed-scale modeled results represent the total annual load for each of the respective scoring areas based on the

35-year-simulated average annual export. Field-scale modeled results for tile effluent and surface runoff using the two parameter schemes represent the total annual load based on the 5-year simulated average annual export. Measured data is the estimated average annual load for each of the watersheds and field ESW3 over the listed sampling periods

Modeled						Measured		
Sub-watershed	Material Exported	Tiled	High priority	Mid-priority	Total	USGS gauge	Annual mean	Range
Lower East River	TP (Mg yr ⁻¹)	1.61	1.18		2.74	4085119 (Sampled 2004-2006)	35.11	17.28–62.60
	Sediment (Mg yr ⁻¹)	1,605	2,385		3,989		9,350	2,422–20,593
Plum Creek	TP (Mg yr ⁻¹)	2.33	0.55	3.36	6.24	4085119 (Sampled 2012-2017)	12.43	6.12–18.69
	Sediment (Mg yr ⁻¹)	2,230	1,194	4,627	8,051		6,605	3,157–11,649
Modeled						Measured		
Edge-of-field ESW3	Default parameters	Kalk (2018) parameters				ESW3 (Sampled 2014-2019)		
Tile effluent	Discharge (m ³ yr ⁻¹)	14,980	17,268				17,169	
	TN (Mg yr ⁻¹)	2.94	3.25				5.3×10^{-2}	
	TP (Mg yr ⁻¹)	5.84×10^{-5}	5.84×10^{-5}				3.72×10^{-3}	
	Sediment (Mg yr ⁻¹)	-	-				477	
Surface runoff	Discharge (m ³ yr ⁻¹)	3,237	3,400				5,131	
	TN (Mg yr ⁻¹)	0.1	0.1				0.1	
	TP (Mg yr ⁻¹)	7.4×10^{-3}	7.9×10^{-3}				2.1×10^{-2}	
	Sediment (Mg yr ⁻¹)	10.52	11.56				7.73	

The measured surface and tile drain discharge, TN, TP, and sediment export from the edge-of-field ESW3 location were checked against the NTT model output using the NSE value. The observed tile drainage had an annual average discharge of 17,169 m³ and tile export of 527×10^{-3} Mg TN, 3.72×10^{-3} Mg TP, and 477×10^{-3} Mg sediment along with an annual average surface runoff discharge measured at 5,131 m³ and an annual surface export of 83.0×10^{-3} Mg TN, 20.6×10^{-3} Mg TP, and 7.73 Mg of sediment between 2014 and 2019 (Table 2; Komiskey et al. 2021). Using NTT model parameter values from the watershed wide comparison (referred to as default values), the average annual surface runoff was estimated at 3,237 and 3,400 m³ using the adjusted parameter values from Kalk (2018). The surface hydrologic output had an NSE value of -0.05 and 0.04 respectively. When the model training period was increased from 5 to 14 years using the default parameters the NSE increased by 0.09 but was well below the 0.5 acceptable value. The average annual tile drainage discharge using the adjusted parameter values was 17,268 m³ and 14,980 m³ using the default parameters. The tile drainage using the adjusted parameter values had an NSE of -1.26 and -1.47 using the watershed wide parameter settings. Although both the adjusted parameter value and default parameter discharges were close to the observed average annual discharge, the negative NSE values resulted from calibration

difficulties to capture the annual tile discharge variation. Total nitrogen surface and tile exports were undervalued when using the default and updated parameter changes and parameters need further refinement. The parameters were not explicitly changed for N export and NSE values reflected that with -4.77 and -4.78 for the default and adjusted parameters, respectively, for the surface export. The organic N values in surface exports were improved from -1.92 for the default parameters to -1.75 using adjusted parameter values and were further improved using the default settings with a longer training period (NSE -1.10). Total phosphorus showed similar trends as TN. There were no changes in NSE values between the tile exports using either the default or adjusted parameter settings (NSE of -3.84). However, the adjusted parameters increased the NSE by for surface exports from -2.02 to -1.89. Using a longer training period increased the NSE to -0.25. The NSE results were also improved when the soil P value was manually adjusted based on soil P test results (Kalk 2018). Although TP results were poor in terms of the NSE, using default parameter settings with a long training period resulted in a NSE of 0.32 for soluble reactive phosphorus (SRP). Sediment export yielded a NSE value of -1.14 and -1.54 with the default and adjusted parameters respectively. Values below 0 for the NSE suggest that the observed mean does a better job of predicting discharge or export than the model.

3.4 Sensitivity to model parameters

Adjusting model parameters to maximize TP export and limit the discrepancy with the measured watershed data resulted in minimal changes within the high priority and mid-priority sites in PC. When parameters were changed to maximize P export in tiled fields, the tile drain P export averaged an increase of 867% and increased the TP export between 19 to 74% with an average increase of 48% across all tiled PC fields when P parameters were changed. Even when a 48% increase in TP export is used to compare the aggregated modeled results with the gauge data for PC, the modeled results are still on the lower end of the range observed at the gauge. Despite these parameter changes and the resultant increases, the nutrient and sediment export were within the original simulation parameter setting margin of error (Online Resource Table S3 to S7).

4 Discussion

4.1 Field- vs. watershed-scale discrepancies

The field-scale results were poor compared to the observed USGS edge-of-field location at ESW3, based on the NSE values and other work using ArcAPEX at this site. Kalk (2018) used three edge-of-field locations for model calibration (including ESW3) and found a NSE value of 0.67 for discharge after calibration at the ESW3 location with a value over 0.5 deemed acceptable for the calibration. Despite that success when using the same parameter values in this study, the NSE value for discharge was significantly lower (NSE 0.04). The U.S. Geological Survey dataset was collected on an event basis following edge-of-field sampling protocols (Komiskey et al. 2021). Having adjusted for differences in timescales, a key difference between the NTT and ArcAPEX research was that the NTT used PRISM data for the weather modeling whereas the work using ArcAPEX used rain gauge data collected at the field location. The PRISM precipitation estimates are derived from weather station data that informs 800-m grid cells that may vary from local data collected at the ESW3 location and could have resulted in some of the discrepancies in modeled water volumes. Precipitation can vary significantly from what was observed at a weather station and a field a few kilometers away. If hydrologic differences exist or are inaccurate, then the nutrient and sediment export cannot be calibrated appropriately within the NTT model.

Kalk's calibrated ArcAPEX model did lose performance in the subsequent years and locations, and the authors noted a number of discrepancies (Kalk 2018). At ESW3, 95% of TP was sediment bound and APEX overestimated the sediment load by 18% (Kalk 2018) which was similar to our

results in this study, indicating almost no DP or tile drain P was shown in the NTT model outputs. This lack of partitioning has dramatic effects on DP and P lost through tile drains. Across all six locations evaluated from within the Lower FRW in Kalk (2018), they found DP and tile drained P were quite different than measured values despite calibrating the model to the surface and tile hydrologic volumes. Similarly, tile drained P was essentially estimated at zero by the NTT in this study which differed from the U.S. Geological Survey field measurements. One change that did make TP export more accurate in both this study and Kalk (2018) was the inclusion of measured Bray P from field sediment samples. The NTT will automatically set this value based on the soil characteristics and farming management scenarios and, in the case of Kalk (2018), the soil P was estimated at roughly half the value as the measured P at the site. Therefore, APEX and the NTT are undervaluing TP stored and/or historically applied in these systems. Even with the default soil P settings, the TP values had better NSE scores when using a longer model training period and should be utilized in future modeling efforts in the FRW. Updating the underlying equations could be one remedy to the undervalued P exports. The P equations are based on partitioning and mineralization research from the 1980s (Knisel 1980; Jones et al. 1984) that likely does not reflect the increased dissolved P loads more recently being observed across the USA (Dodds et al. 2009; Stoddard et al. 2016). However, this change is outside the control of the average NTT and APEX user.

One other large challenge in the calibration is how to handle the presence of tiling and snowmelt in the NTT and underlying APEX system. The surface drainage area in the observed field was approximately 2.8 ha, whereas the estimated tile drainage area was approximately 10 ha. Without explicit knowledge of tile networks, the accurate simulation of agricultural fields with tile drainage poses a challenge due to unknown hydrologic connectivity which is critical in systems that experience snowmelt. Kalk (2018) noted that the APEX model had a difficult time predicting the influence of snowmelt during their calibration because APEX uses solar radiation to predict snowmelt which is a simplification of the melting process. Relying solely on solar radiation to simulate snowmelt may not work well in the upper Midwest where the largest nutrient and sediment loads typically occur during the spring snowmelt period (Huisman et al. 2013; Dolph et al. 2019). Overall, APEX and the NTT need further refinement to accurately model field-scale exports within the Lower FRW, including but not limited to field data such as tile networks, nutrient levels from soil tests, fertilization practices, management choices, and the ability to choose specific vegetation for buffers and forested areas to inform model calibration and assess model output.

At the watershed scale, aggregating undervalued field-scale results could compound into the large discrepancies

observed for TP at the LER and PC (Table 2). If APEX is undervaluing the DP and tile drain export, aggregating those results to the watershed scale could drastically reduce the estimated nutrient loads at the watershed outlets. However, without field data from other locations for verification, it is impossible to know if this undervaluation of loads is field-specific. Incorrect soil P data may also contribute to the discrepancies between the observed and simulated TP export values. Kalk (2018) found that soil P was estimated by APEX at approximately 50% of the measured Bray P value which was similar to this study. Additionally, when aggregating to an entire watershed one shortcoming of using field-scale models is the lack of accounting for storage and transport processes within the stream channel. Phosphorus in particular can be held in varying sediment fractions with differing capacities to release SRP to the water column (Lannergård et al. 2020). Associations with iron and manganese oxides may be reduced by microbiota in anoxic conditions and release P. In systems with high calcium concentrations, secondary mineral formation of apatite which binds calcium and P in a mineral structure may essentially eliminate any possible SRP release because the mineral has low solubility. Additionally, biological processes paired with hydrologic fluxes may lengthen the time both P and sediment are stored within the stream channel (Dolph et al. 2019). However, aggregated results for sediment were within observed ranges at the watershed scale despite having NSE values below zero (an indication that the observed mean predicted sediment export better than the modeled results at the field scale), indicating that modeled values were close to actual field-scale exports but not a replacement for actual field data. These discrepancies between the field- and watershed-scale results emphasizes tradeoffs between field-scale accuracy and the accuracy needed for a generalized watershed-scale aggregation when a model cannot be finely tuned to every field within a watershed due to lack of site knowledge and resources.

Despite these results, the NTT and the underlying APEX model have been found to be effective models in agricultural systems at the field scale (Gassman et al. 2009; Saleh et al. 2015) and is an excellent model for comparing and developing conservation practices with individual landowners (Saleh et al. 2011). The more a model is calibrated to an individual farming system, the more likely it is to accurately predict nutrient and sediment losses. Shortcomings can arise when model parameters are applied across varying sites and when applying a field-scale model at the larger watershed spatial scale. The NTT was not designed to account for hydrologic, sediment, and nutrient dynamics beyond the field or at a high temporal resolution. While some daily data can be extracted from the underlying APEX files, most of the nutrient and sediment export data is extractable as an average monthly or annual value. This lack of temporal

resolution should be taken into consideration and further validated before using the NTT because of the mounting evidence that extreme events account for most of the annual nutrient and sediment export within watersheds (Dolph et al. 2019). Channel processes and hydrogeological routing are also absent from the NTT and could be a source of disparities between aggregated field-scale model results and watershed outlet data particularly in systems with built up P reservoirs. Other models that include better subsurface hydrologic routing, nutrient and sediment storage parameters, channel processes, and the ability to incorporate land uses other than agriculture may be more effective at modeling nutrient and sediment export at larger scales and in more diverse landscapes. The USEPA summarizes 30 different models ranging from simple to difficult and two modeling systems that can be used for watershed protection (USEPA 2018). Some of these models, such as the NTT, address the field scale and require moderate knowledge of the field system, whereas others only work at the watershed scale (e.g., SPATIally Referenced Regression on Watershed attributes or SPARROW). Other models are capable of modeling at both scales such as SWAT (Soil and Water Assessment Tool).

The greatest benefit of the NTT is the intuitive and easy-to-use interface. Even though limitations exist, it is a more accessible model to a variety of users (e.g., landowners and local conservation district employees) versus a more sophisticated model such as SWAT which requires advanced training. Moving forward, utilizing the NTT's APEX output or using APEX to model mitigation strategies at the field scale in conjunction with agricultural management systems as an input to a model such as SWAT for watershed assessments could be an effective way to model and refine water quality goals (Wang et al. 2011). Model choice should also take the availability of data for validation into consideration. Data found for this study encompassed a spatial extent beyond what is found at the field scale. Although the aggregated data comparison for the LER and PC fell within measured ranges observed at USGS gage locations for sediment, in-channel and watershed-scale processes are absent from the modeling effort and could be a source of error in the LER, PC, and other sub-watersheds. These potential sources of error bring to question the suitability of using the NTT, or other field-scale planning tools, in future work to model nutrient and sediment export at the watershed or larger spatial scale. Despite a lack of field-scale data, the goal to assess potential benefits of revegetation as a mitigation strategy was able to be accomplished using the NTT and existing datasets.

As a field-scale model, the NTT is an excellent tool for understanding the interplay between agriculture and conservation practices that can be tailored to individual farming systems. To aggregate data over a broad spatial extent, such as a watershed, meant that individual fields were simplified (Keller and Fox 2019) by making assumptions about

the cropping systems to complete the analysis because specific management characteristics were unknown. Future work with the NTT could be improved by diversifying farm management practices and adding complexity (e.g., including manure application, no-till) and tailoring model parameters to individual fields. Tailoring model parameters would likely reduce disparities between results and field-scale observational data.

4.2 Management considerations

4.2.1 Incorporating vegetation

The addition of buffers decreased the modeled N, P, and sediment loads from the fields versus the typical corn–soybean rotation. Our ability to modify buffers based on individual field physical characteristics such as topography and landscape setting is another limitation of the NTT but could provide further load reductions. Within the NTT, a buffer width may be specified but the user interface does not allow for the modification, the addition of multiple buffers on a field, or specific site selection. Although not used in this study, one way to work around this limitation is to divide a field into smaller sub-areas. One or multiple subareas could be designated as a buffer with full vegetative coverage. Users could then use the NTT routing function to have the cropped sub-areas flow into the buffered sub-areas to assess buffer placement on the nutrient and sediment exports. Buffer design that varies buffers with topography and drainage characteristics such that areas where the majority of nutrients and sediment are being exported from the field surface (i.e., areas of convergent overland flow) are proven to be more effective at reducing nutrient and sediment loads (Dosskey et al. 2002; Helmers et al. 2011; Zhou et al. 2014; Schulte et al. 2017; Jiang et al. 2020). Total buffer area can also be spread out into multiple strips within a field and still yield reductions as long as the design allows for the greatest contact between the buffer and the flow bearing nutrients and sediment from the field (Hernandez-Santana et al. 2013).

Buffers are a commonly used best management practice, so field-scale designs are paramount to their success. If landowners are unwilling to segment a field using vegetated strips, conservation practices may still be optimized by placing vegetated features near the channel (Hansen et al. 2021). Placing vegetation between the fields and the stream would further delay routing and be pivotal in reducing nutrient and sediment loads entering the stream channel (Kreiling et al. 2021). Nutrient and sediment storm event data from PC shows hysteresis patterns where transport to streams is delayed in the upper portion of the watershed suggesting mobilization from fields (Rose and Karwan 2021) and provides evidence that vegetation could further delay and largely reduce delivery to the stream channel.

Despite many positives, buffer effectiveness varies spatially, temporally, and by species composition (Schulte et al. 2017). Increased buffer thickness at concentrated flow points can drastically improve the nutrient and sediment retention (Dosskey et al. 2002). Buffers can act as nutrient sources especially in cold climates (Vanrobaeys et al. 2019) and may also increase the available soil P pool. Deep-rooted perennial plants have been shown to increase labile P previously immobilized in soils (Stutter et al. 2009). Harvesting buffers could help offset these negatives by physically removing nutrients tied up in biomass. Removal of one kind of vegetation within a buffer at a time (e.g., a buffer with trees and grasses where only the trees or grass is removed) can also be used to minimize reduced nutrient and sediment capture efficiency stemming from vegetation loss (Jiang et al. 2020; Liu et al. 2021).

Being able to assess different vegetation communities for both buffer and forest conversion management scenarios would be an improvement to the NTT model version 20–2 framework. Vegetation choices for forest conversion and buffers were unavailable for this research and would be a useful addition to future versions of the NTT software. Nutrient losses following agricultural land conversion to forest or other regionally appropriate vegetation could be improved with a broader variety of forest and endemic vegetation cover types. Specifically, the conversion of agricultural land to red pine plantation may not be as beneficial to nutrient and sediment export as other forest types or forested wetlands historically found in the area. In many locations within the Great Lakes region, red pine may not be a suitable species. Within the NTT, the fields are assumed to be managed for agricultural products and the NTT was not designed to simulate land conversion between agriculture and specific forest or wetland vegetation communities. In these forested systems, the NTT might not fully parameterize the nutrient and water quality tradeoffs under the full range of other forest and wetland vegetation types. Incorporating vegetation suitability from soil survey databases into the scoring system may be one alternative to work around these limitations in the future. If the forest conversion scenario were expanded to include open or forested wetlands, the field-scale results have the potential to yield more realistic results at the watershed scale from a mitigation standpoint. Both upland forests and forested wetlands retain nutrients, which may provide a means for economic income through tree harvesting, and these landscape features dampen the effects of extreme hydrologic events by lowering the magnitude and increasing the duration of elevated discharge observed in stream networks (Kovacic et al. 2006; Grant et al. 2008; Salemi et al. 2012; Zhu et al. 2020). Future work assessing the strategic placement of vegetated areas could have a dramatic effect on nutrient and sediment loads observed in the stream channel. An example of vegetated systems strategically placed

on a large scale are the Everglades Stormwater Treatment Areas (STAs) which were created to intercept and slow the nutrient and sediment loads flowing from agricultural areas to the Everglades. These wetlands have succeeded in reducing nutrients and sediment in addition to providing flood mitigation (Moustafa et al. 2012; Mitsch et al. 2015; Pietro and Ivanoff 2015). A similar scenario could work as a long-term mitigation strategy within the FRW if perennial upland forest, prairie, or wetland vegetation is implemented at a larger scale.

The NTT does allow for the simultaneous incorporation of other management practices despite limitations on complexity such as with buffer design. The scenarios we tested were buffer installation and land conversion to a forested plantation system. Buffer installation combined with other conservation practices such as implementation of no-till or the use of cover crops could further reduce nutrient and sediment loads and increase buffer effectiveness. These practices behave relatively similar to buffers where carbon storage, water infiltration, water storage, nutrient retention, and erosion protection are field characteristics that are generally improved with adoption along with improvements to overall soil health and biodiversity (Kaspar and Singer 2011; Poeplau and Don 2015; Myers et al. 2019). Cover crops in particular reduce farm costs when combined with other cropping practices and provide many economic benefits to the landowner (Myers et al. 2019). On the other side of the spectrum, the conversion of cropped fields to forests or wetland vegetation, albeit an effective way to reduce nutrient and sediment loads, is unlikely to be implemented as a management option without further economic incentives. In spite of the simplification to not incorporate other practices to compare aggregated modeled results to the loads observed at the LER and PC watersheds, future work incorporating more specific cropping systems will likely improve model results and close the gap between the observed P export and the modeled export.

4.2.2 Tile drainage

One benefit of working in the LER and PC watersheds was the inclusion of tile drainage in the management scenarios for the Lower FRW. Future work in larger watersheds will need to overcome data limitations regarding tile drainage extent. No large-scale records exist of tile networks because they are implemented on private property. Available data did not allow for the validation of the modeled tile exports in the LER and PC watersheds; however, the NTT sensitivity analysis indicated weaknesses to the model and/or the scenario in these watersheds. By changing parameters affecting P loads, the tile drain P increased from an average of about 5% to upwards of 34 to 51% of the TP load which is in agreement with other studies following similar fertilization

scenarios (Williams et al. 2016). However, those increases were within the margin of error for the scenario outcomes using the original model parameter values. Even if all fields within LER and the PC watersheds were assumed to be tiled, a 50% increase in the TP load would bring the modeled estimates for the LER to 4.11 Mg P yr⁻¹ and PC to around 9.36 Mg P yr⁻¹. The estimate would bring PC to within 50% of the maximum observed export based on the gage data (USGS gauge 04084911, U.S. Geological Survey 2016b), but LER is still much lower than measured (U.S. Geological Survey 2016a).

Despite the dissimilarity between the gauge data and the model, the model results show that the inclusion of tile drainage reduces the effects of vegetation as a nutrient and sediment mitigation strategy and increases nutrient export which is in general agreement with the literature (Royer et al. 2006; Carstensen et al. 2020; Klaiber et al. 2020; Schilling et al. 2020; Hanrahan et al. 2021; Thalmann 2021). Tile drains tend to increase dissolved N and P losses from fields, but in comparison to undrained fields increased surface losses may have a net zero effect on TP export depending on the agricultural management practices used in the field (Royer et al. 2006; Klaiber et al. 2020; Thalmann 2021). The presence of tile drains allows for targeted mitigation of the drainage systems (Schilling et al. 2020). Terminating tile drains prior to reaching buffers or rerouting drainage so that tile effluent must flow through vegetated buffers, wetlands, or bioreactors are common mitigation strategies that yield nutrient and sediment reductions despite the presence of subsurface drainage (Jaynes and Isenhardt 2014; Carstensen et al. 2020; Mendes 2020; Gordon et al. 2021). Thus, when planning mitigation using vegetation, knowledge of subsurface drainage is critical to the mitigation's effectiveness from a load reduction viewpoint.

5 Conclusion

The NTT was used to successfully simulate nutrient and sediment export from three sub-watersheds within the FRW. The modeling demonstrated that conversion or conservation of vegetated areas within watersheds is an effective management strategy to reduce nutrient and sediment export from fields and that the targeted placement of vegetation can result in 80 to 95% load reductions. By using slope and hydrologic drainage class, conservation measures in high priority areas were shown to have greater sediment, TN, and TP reductions than mid-priority scoring areas within the PC watershed when buffer and forest conversion scenarios were compared to the corn–soybean rotation.

Future work will be needed to build datasets to further validate NTT results at both the field and watershed scales

to meet load reduction requirements of sub-watersheds in the FRW. Modeled field hydrologic discharge along with nutrient and sediment export were not accurate compared to measured field-scale means using generalized parameter settings. The field-scale modeling exercise showed that field-scale data is pivotal to accurately modeling field exports with the NTT. Detailed management practices and inputting Bray P measurements increased the field-scale accuracy. When scaling up to watershed exports using the generalized parameter settings worked well for estimating sediment loads. The aggregated values of TP were undervalued but sediment fell within range of observed annual fluxes at the sampling locations. Although modeled results fell within observed ranges for sediment export at the watershed scale, the NTT and its user-friendly interface are likely better suited for use on an individual field basis rather than at a larger aggregated watershed scale. The NTT was designed to focus on individual fields and could prove to be a powerful tool to help agricultural producers implement changes and model scenarios to reduce nutrient and sediment loads into downstream waterbodies by playing to its strengths.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11368-023-03521-y>.

Acknowledgements Dr. Diana L. Karwan received additional support from the Minnesota Agricultural Experiment Station. Margaret McGann and Zoe Brown provided assistance in compiling gauge export data.

Funding Funding was provided by the USDA Forest Service from the Great Lakes Restoration Initiative.

Data availability The data that supports and was generated by this study are available upon request from the corresponding author.

Declarations

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alexander RB, Smith RA, Schwarz GE et al (2008) Differences in phosphorus and nitrogen delivery to the Gulf of Mexico from the Mississippi River Basin. *Environ Sci Technol* 42:822–830. <https://doi.org/10.1021/es0716103>
- Asbjornsen H, Hernandez-Santana V, Liebman M et al (2014) Targeting perennial vegetation in agricultural landscapes for enhancing ecosystem services. *Renew Agric Food Syst* 29:101–125. <https://doi.org/10.1017/S1742170512000385>
- Brant G (2003) Barriers and strategies influencing the adoption of nutrient management practices (Technical Report 13.1). U.S. Department of Agriculture, Natural Resources Conservation Service
- Carstensen MV, Hashemi F, Hoffmann CC et al (2020) Efficiency of mitigation measures targeting nutrient losses from agricultural drainage systems: a review. *Ambio* 49:1820–1837. <https://doi.org/10.1007/s13280-020-01345-5>
- Climate Green Bay, Wisconsin (2021) In: U.S. Clim. Data. <https://www.usclimatedata.com/climate/green-bay/wisconsin/united-states/uswi0288>. Accessed 8 Jan 2020
- Dodds WK, Bouska WW, Eitzmann JL et al (2009) Eutrophication of U. S. freshwaters: analysis of potential economic damages. *Environ Sci Technol* 43:12–19. <https://doi.org/10.1021/es801217q>
- Dolph CL, Boardman E, Danesh-Yazdi M et al (2019) Phosphorus transport in intensively managed watersheds. *Water Resour Res* 55:9148–9172. <https://doi.org/10.1029/2018WR024009>
- Dosskey MG, Helmers MJ, Eisenhauer DE et al (2002) Assessment of concentrated flow through riparian buffers. *J Soil Water Conserv* 57:336–343
- Gassman PW, Williams JR, Wang X et al (2009) The Agricultural Policy/Environmental eXtender (APEX) model: an emerging tool for landscape and watershed environmental analyses (Technical Report 09-TR 49). Center for Agricultural and Rural Development, Iowa State University
- Gordon BA, Lenhart C, Peterson H et al (2021) Reduction of nutrient loads from agricultural subsurface drainage water in a small, edge-of-field constructed treatment wetland. *Ecol Eng* 160:106128. <https://doi.org/10.1016/j.ecoleng.2020.106128>
- Graczyk DJ, Robertson DM, Baumgart PD, Fermanich KJ (2011) Hydrology, phosphorus, and suspended solids in five agricultural streams in the Lower Fox River and Green Bay Watersheds, Wisconsin, Water Years 2004–06. *Sci Investig Rep* 2011–5111
- Grant GE, Lewis SL, Swanson FJ et al (2008) Effects of forest practices on peak flows and consequent channel response: a state-of-science report for western Oregon and Washington. USDA For Serv - Gen Tech Rep PNW-GTR 1–82. <https://doi.org/10.2737/PNW-GTR-760>
- Hadley DW, Pelham J (1976) Glacial deposits of Wisconsin: Sand and gravel resource potential. In: Wisconsin Geological and Natural History Survey. https://ngmdb.usgs.gov/ngm-bin/pdp/zui_viewer.pl?id=19927. Accessed 8 Jan 2020
- Hanrahan BR, King KW, Williams MR (2021) Controls on subsurface nitrate and dissolved reactive phosphorus losses from agricultural fields during precipitation-driven events. *Sci Total Environ* 754:142047. <https://doi.org/10.1016/j.scitotenv.2020.142047>
- Hansen AT, Campbell T, Cho SJ et al (2021) Integrated assessment modeling reveals near-channel management as cost-effective to improve water quality in agricultural watersheds. *Proc Natl Acad Sci USA* 118:1–25. <https://doi.org/10.1073/pnas.2024912118>
- Helmers MJ, Zhou X, Asbjornsen H et al (2011) Sediment removal by prairie filter strips in row-cropped ephemeral watersheds. *J Environ Qual* 41:1531–1539. <https://doi.org/10.2134/jeq2011.0473>
- Hernandez-Santana V, Zhou X, Helmers MJ et al (2013) Native prairie filter strips reduce runoff from hillslopes under annual row-crop systems in Iowa, USA. *J Hydrol* 477:94–103. <https://doi.org/10.1016/j.jhydrol.2012.11.013>
- Huisman NLH, Karthikeyan KG, Lamba J et al (2013) Quantification of seasonal sediment and phosphorus transport dynamics in an agricultural watershed using radiometric fingerprinting techniques. *J Soils Sediments* 13:1724–1734. <https://doi.org/10.1007/s11368-013-0769-0>
- Jacobson MD (2012) Phosphorus and sediment runoff loss: management challenges and implications in a Northeast Wisconsin

- agricultural watershed [Masters Thesis, University of Wisconsin-Green Bay].
- Jaynes DB, Isenhardt TM (2014) Reconnecting tile drainage to riparian buffer hydrology for enhanced nitrate removal. *J Environ Qual* 43:631–638. <https://doi.org/10.2134/jeq2013.08.0331>
- Jiang F, Preissendanz HE, Veith TL et al (2020) Riparian buffer effectiveness as a function of buffer design and input loads. *J Environ Qual* 49:1599–1611. <https://doi.org/10.1002/jeq2.20149>
- Jones CA, Cole CV, Sharples AN, Williams JR (1984) A simplified soil and plant phosphorus model. *Soil Sci Soc Am J* 48:800–805
- Kalk FS (2018) Evaluation of the APEX model to simulate runoff, sediment, and phosphorus loss from agricultural fields in Northeast Wisconsin [Masters Thesis, University of Wisconsin-Green Bay].
- Kaspar TC, Singer JW (2011) The use of cover crops to manage soil. Publ from USDA-ARS/UNL Fac 1382. <https://doi.org/10.2136/2011.soilmanagement.c21>
- Keller AA, Fox J (2019) Giving credit to reforestation for water quality benefits. *PLoS ONE* 14:1–18. <https://doi.org/10.1371/journal.pone.0217756>
- Klaiber LB, Kramer SR, Young EO (2020) Impacts of tile drainage on phosphorus losses from edge-of-field plots in the Lake Champlain Basin of New York. *Water (switzerland)* 12:1–15. <https://doi.org/10.3390/w12020328>
- Knisel WG (1980) CREAMS, A field scale model for chemicals, runoff, and erosion from agricultural management systems. U.S. Department of Agriculture, Science and Education Administration, conservation research report number 26. <https://www.tucson.ars.ag.gov/unit/publications/PDFfiles/312.pdf>
- Komiskey MJ, Stuntebeck TD, Loken LC et al (2021) Nutrient and sediment concentrations, loads, yields, and rainfall characteristics at USGS surface and subsurface-tile edge-of-field agricultural monitoring sites in Great Lakes States (ver. 1.0, May 2021) U.S. Geological Survey data release. <https://doi.org/10.5066/P9LO8O70>
- Kovacic DA, Twait RM, Wallace MP, Bowling JM (2006) Use of created wetlands to improve water quality in the Midwest-Lake Bloomington case study. *Ecol Eng* 28:258–270. <https://doi.org/10.1016/j.ecoleng.2006.08.002>
- Kreiling RM, Bartsch LA, Perner PM et al (2021) Riparian forest cover modulates phosphorus storage and nitrogen cycling in agricultural stream sediments. *Environ Manage* 68:279–293. <https://doi.org/10.1007/s00267-021-01484-9>
- Lannergård EE, Agstam-Norlin O, Huser BJ et al (2020) New insights into legacy phosphorus from fractionation of streambed sediment. *J Geophys Res Biogeosciences* 125. <https://doi.org/10.1029/2020JG005763>
- Lenhart CF, Smith DJ, Lewandowski A et al (2018) Assessment of stream restoration for reduction of sediment in a large agricultural watershed. *J Water Resour Plan Manag* 144:1–13. [https://doi.org/10.1061/\(ASCE\)WR.1943-5452.0000908](https://doi.org/10.1061/(ASCE)WR.1943-5452.0000908)
- Liu R, Thomas BW, Shi X et al (2021) Effects of ground cover management on improving water and soil conservation in tree crop systems: a meta-analysis. *Catena* 199:105085. <https://doi.org/10.1016/j.catena.2020.105085>
- Lohani S, Baffaut C, Thompson AL et al (2020) Performance of the Soil Vulnerability Index with respect to slope, digital elevation model resolution, and hydrologic soil group. *J Soil Water Conserv* 75:12–27. <https://doi.org/10.2489/JSWC.75.1.12>
- Luczaj JA (2011) Preliminary geologic map of the buried bedrock surface, Brown County, Wisconsin (WOFR2011-02). Wisconsin Geol Nat Hist Surv Open-File Rep 2011-02:1
- Luczaj JA (2013) Geology of the Niagara escarpment. *Geosci Wisconsin* 22
- Mendes LRD (2020) Edge-of-field technologies for phosphorus retention from agricultural drainage discharge. *Appl Sci* 10. <https://doi.org/10.3390/app10020634>
- Mitsch WJ, Zhang L, Marois D, Song K (2015) Protecting the Florida Everglades wetlands with wetlands: can stormwater phosphorus be reduced to oligotrophic conditions? *Ecol Eng* 80:8–19. <https://doi.org/10.1016/j.ecoleng.2014.10.006>
- Moustafa MZ, White JR, Coghlan CC, Reddy KR (2012) Influence of hydropattern and vegetation on phosphorus reduction in a constructed wetland under high and low mass loading rates. *Ecol Eng* 42:134–145. <https://doi.org/10.1016/j.ecoleng.2012.01.028>
- Multi-Resolution Land Characteristics Consortium (U.S.) (2016) National land cover dataset (NLCD)
- Myers R, Weber A, Tellatin S (2019) Cover crop economics opportunities to improve your bottom line in row crops. *SARE Tech Bull* 1–24
- Natural Resources Conservation Service (2008a) Rapid watershed assessment Lower Fox River Watershed (WI) HUC:04030204. Natural Resources Conservation Service, Wisconsin State Office, Madison, WI
- Natural Resources Conservation Service (2008b) Rapid watershed assessment Wolf River Watershed (WI) HUC: 04030202. Natural Resources Conservation Service, Wisconsin State Office, Madison, WI
- Neitsch SI, Arnold JG, Kiniry JR et al (2002) Soil Water and Assessment Tool user's manual version 2000. Texas water resources institute, College Station, Texas, TWRI Report TR-192. <https://swat.tamu.edu/media/1294/swatuserman.pdf>
- Pietro KC, Ivanoff D (2015) Comparison of long-term phosphorus removal performance of two large-scale constructed wetlands in South Florida, U.S.A. *Ecol Eng* 79:143–157. <https://doi.org/10.1016/j.ecoleng.2014.12.013>
- Poeplau C, Don A (2015) Carbon sequestration in agricultural soils via cultivation of cover crops - a meta-analysis. *Agric Ecosyst Environ* 200:33–41. <https://doi.org/10.1016/j.agee.2014.10.024>
- Ponce VM, Hawkins RH (1996) Runoff curve number: has it reached maturity? *J Hydrol Eng* 1:11–19
- Porter SA, Tomer MD, James DE, Boomer KMB (2016) Agricultural Conservation Planning Framework ArcGIS® Toolbox user's manual version 2.2 date of release: 06/2017. 82
- Powers SM, Johnson RA, Stanley EH (2012) Nutrient retention and the problem of hydrologic disconnection in streams and wetlands. *Ecosystems* 15:435–449. <https://doi.org/10.1007/s10021-012-9520-8>
- PRISM Climate Group (2004) No title. In: Oregon State Univ. <http://prism.oregonstate.edu>
- Prosser RS, Hoekstra PF, Gene S et al (2020) A review of the effectiveness of vegetated buffers to mitigate pesticide and nutrient transport into surface waters from agricultural areas. *J Environ Manage* 261:110210. <https://doi.org/10.1016/j.jenvman.2020.110210>
- Provost S, Hyatt L, McLennan R et al (2001) The state of the Upper Fox River Basin (Publication WT-665-2001). A report by the Wisconsin Department of Natural Resources in cooperation with the Upper Fox River Partnership Team. https://nelson.wisc.edu/wp-content/uploads/2001_DNR_State_of_Fox_River_Basin.pdf
- Reddy KR, Kadlec RH, Flaig E, Gale PM (1999) Phosphorus retention in streams and wetlands: a review. *Crit Rev Environ Sci Technol* 29:83–146. <https://doi.org/10.1080/10643389991259182>
- Rose DC, Sutherland WJ, Amano T et al (2018) The major barriers to evidence-informed conservation policy and possible solutions. *Conserv Lett* 11:1–12. <https://doi.org/10.1111/conl.12564>
- Rose LA, Karwan DL (2021) Stormflow concentration–discharge dynamics of suspended sediment and dissolved phosphorus in an agricultural watershed. *Hydrol Process* 35:1–16. <https://doi.org/10.1002/hyp.14455>
- Royer TV, David MB, Gentry LE (2006) Timing of riverine export of nitrate and phosphorus from agricultural watersheds in Illinois: implications for reducing nutrient loading to the Mississippi River. *Environ Sci Technol* 40:4126–4131. <https://doi.org/10.1021/es052573n>

- Rubin JA, Görres JH (2021) Potential for mycorrhizae-assisted phytoremediation of phosphorus for improved water quality. *Int J Environ Res Public Health* 18:1–24. <https://doi.org/10.3390/ijerph18010007>
- Saleh A, Gallego O, Osei E et al (2011) Nutrient Tracking Tool—a user-friendly tool for calculating nutrient reductions for water quality trading. *J Soil Water Conserv* 66:400–410. <https://doi.org/10.2489/jswc.66.6.400>
- Saleh A, Gallego O, Osei E (2015) Evaluating nutrient tracking tool and simulated conservation practices. *J Soil Water Conserv* 70:115–120. <https://doi.org/10.2489/jswc.70.5.115A>
- Salemi LF, Groppo JD, Trevisan R et al (2012) Riparian vegetation and water yield: a synthesis. *J Hydrol* 454–455:195–202. <https://doi.org/10.1016/j.jhydrol.2012.05.061>
- Santy KL, Bougie C, Lychwick T et al (2001) Lower Fox River Basin integrated management plan (Publication WT-666-2001). A report by the Wisconsin Department of Natural Resources in cooperation with the Lower Fox River Basin Partnership Team. <https://dnr.wi.gov/water/basin/lowerfox/lowerfox.pdf>
- Schilling KE, Jacobson P (2014) Effectiveness of natural riparian buffers to reduce subsurface nutrient losses to incised streams. *CATENA* 114:140–148. <https://doi.org/10.1016/j.catena.2013.11.005>
- Schilling KE, Streeter MT, Vogelgesang J, et al (2020) Subsurface nutrient export from a cropped field to an agricultural stream: implications for targeting edge-of-field practices. *Agric Water Manag* 241:. <https://doi.org/10.1016/j.agwat.2020.106339>
- Schulte LA, Niemi J, Helmers MJ et al (2017) Prairie strips improve biodiversity and the delivery of multiple ecosystem services from corn–soybean croplands. *Proc Natl Acad Sci USA* 114:11247–11252. <https://doi.org/10.1073/pnas.1620229114>
- Selman M, Greenhalgh S, Diaz R, Sugg Z (2008) Eutrophication and hypoxia in coastal areas: a global assessment of the state of knowledge. *Water Resour Inst* 1–6
- Sharpley AN, Kleinman PJA, Jordan P et al (2009) Evaluating the success of phosphorus management from field to watershed. *J Environ Qual* 38:1981–1988. <https://doi.org/10.2134/jeq2008.0056>
- Soil Survey Staff (2020a) Soil Survey Geographic (SSURGO) Database for Wisconsin. United States Department of Agriculture, Natural Resources Conservation Service. Available online at <https://gdg.sc.egov.usda.gov>. Accessed 2020
- Soil Survey Staff (2020b) Web soil survey. United States Department of Agriculture. Available online at <https://websoilsurvey.nrcs.usda.gov/>. Accessed 2020
- Stoddard JL, Van Sickle J, Herlihy AT et al (2016) Continental-scale increase in lake and stream phosphorus: are oligotrophic systems disappearing in the United States? *Environ Sci Technol* 50:3409–3415. <https://doi.org/10.1021/acs.est.5b05950>
- Stutter MI, Langan SJ, Lumsdon DG (2009) Vegetated buffer strips can lead to increased release of phosphorus to waters: a biogeochemical assessment of the mechanisms. *Environ Sci Technol* 43:1858–1863. <https://doi.org/10.1021/es8030193>
- Texas A&M AgriLife Research (2013) APEX - Agricultural Policy/Environmental eXtender Model (Version 0806). Blackland Research and Extension Center, Temple, Texas.
- Thalmann LE (2021) Phosphorus and nitrogen losses in runoff from fields with and without tile drainage. Graduate College Dissertations and Theses. 1355. <https://scholarworks.uvm.edu/graddis/1355>
- Thompson SE, Basu NB, Lascrain J et al (2011) Relative dominance of hydrologic versus biogeochemical factors on solute export across impact gradients. *Water Resour Res* 47:1–20. <https://doi.org/10.1029/2010WR009605>
- Tilman D, Balzer C, Hill J, Befort BL (2011) Global food demand and the sustainable intensification of agriculture. *Proc Natl Acad Sci USA* 108:20260–20264. <https://doi.org/10.1073/pnas.1116437108>
- Tilman D, Cassman KG, Matson PA et al (2002) Agricultural sustainability and intensive production practices. *Nature* 418:671–677
- Tomer MD, Locke MA (2011) The challenge of documenting water quality benefits of conservation practices: a review of USDA-ARS's conservation effects assessment project watershed studies. *Water Sci Technol* 64:300–310. <https://doi.org/10.2166/wst.2011.555>
- U.S. Geological Survey (2018) USGS NED 1/3 arc-second n40w098 1 x 1 degree IMG. In U.S. Geol Surv
- U.S. Geological Survey (2020) National Hydrography Dataset for hydrological unit (HU) 4–0403
- U.S. Geological Survey (2016a) National Water Information System data available on the World Wide Web (USGS Water Data for the Nation). https://waterdata.usgs.gov/nwis/uv?site_no=040851378. Accessed 8 Jan 2020
- U.S. Geological Survey (2016b) National Water Information System data available on the World Wide Web (USGS Water Data for the Nation). https://waterdata.usgs.gov/wi/nwis/uv/?site_no=04084911&PARAMeter_cd=00065.00060. Accessed 8 Jan 2020
- US Environmental Protection Agency (2018) Nutrient and sediment estimation tools for watershed protection. Report EPA 841-K-18–002
- USDA ARS Grassland Soil and Water Research L, Texas A&M AgriLife Research (2018) SWAT- Soil and Water Assessment Tool
- Vanrobaeys JA, Owens PN, Lobb DA et al (2019) Seasonal efficacy of vegetated filter strips for phosphorus reduction in surface runoff. *J Environ Qual* 48:880–888. <https://doi.org/10.2134/jeq2018.12.0452>
- Walsh JC, Dicks L V., Raymond CM, Sutherland WJ (2019) A typology of barriers and enablers of scientific evidence use in conservation practice. *J Environ Manage* 250:109481. <https://doi.org/10.1016/j.jenvman.2019.109481>
- Wang X, Kannan N, Santhi C et al (2011) Integrating APEX output for cultivated cropland with SWAT simulation for regional modeling. *Trans ASABE* 54:1281–1298
- Williams MR, King KW, Baker DB et al (2016) Hydrologic and biogeochemical controls on phosphorus export from Western Lake Erie tributaries. *J Great Lakes Res* 42:1403–1411. <https://doi.org/10.1016/j.jglr.2016.09.009>
- Zhou X, Helmers MJ, Asbjornsen H et al (2014) Nutrient removal by prairie filter strips in agricultural landscapes. *J Soil Water Conserv* 69:54–64. <https://doi.org/10.2489/jswc.69.1.54>
- Zhu X, Liu W, Chen J et al (2020) Reductions in water, soil and nutrient losses and pesticide pollution in agroforestry practices: a review of evidence and processes. *Plant Soil* 453:45–86. <https://doi.org/10.1007/s11104-019-04377-3>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.