Environmental Modelling and Software 148 (2022) 105269

Contents lists available at ScienceDirect

Environmental Modelling and Software

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

A modified response matrix method to approximate SWAT for computationally intense applications

Shaobin Li [1], Kevin Wallington [1], Sundar Niroula [1], Ximing Cai [*]

Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, 61801, USA

ARTICLE INFO

Keywords:
Surrogate modeling
Nutrient transport
Watershed modeling
Soil and water assessment tool (SWAT)
Management decisions

ABSTRACT

The computational burden of running a semi-distributed hydrological model numerous times, such as for optimization applications, can be exorbitant. This study provides a surrogate model to estimate streamflow, nutrient, and sediment export under spatially distributed management decisions. Specifically, we surrogate the Soil and Water Assessment Tool (SWAT) using a modified response matrix (RM) approach. A traditional RM approach applied to SWAT assumes hydrological responses are approximated by linear functions of management decisions, and falls short in accounting for in-stream and reservoir processes. Here, we explain and illustrate three key modifications that address interaction effects between co-located conservation practices and in-stream and reservoir processes affecting nutrient and sediment loads. The modified RM approach provides excellent estimation (Nash-Sutcliffe Efficiency, NSE >0.95) for streamflow and nutrient export throughout the stream network and provides very good estimation (NSE >0.85) for sediment export at most, though not all, points in the stream network.

1. Introduction

Many process-based hydrological watershed modeling tools have been developed to evaluate the effectiveness of agricultural conservation practices, such as the Soil and Water Assessment Tool (SWAT), Water Quality Analysis Simulation Program (WASP), and Storm Water Management Model (SWMM) (Babbar-Sebens et al., 2015; Daniel et al., 2011; Sinshaw et al., 2019). These models simulate a multitude of processes, such as runoff, infiltration, channel and reservoir routing, sedimentation, and nutrient dynamics. Due to their complexity, watershed models can be computationally expensive; for example, it can take several minutes to hours to simulate hydrology and nutrient loads with models such as SWAT and MODFLOW (Arnold et al., 2012; Peterson et al., 2016; Zhang et al., 2009). This computational load becomes cumbersome when the model must be run thousands of times, such as when searching for optimal management decisions. The need for reduced computational burden multiplies further when the watershed model is just one among multiple integrated models within the optimization framework. In an integrated modeling framework, individual process-based models also typically cannot communicate directly, given their independent development.

Surrogate models, which capture statistical relationships between inputs and outputs, can ease the burdens of computation and integration, though at some cost of model fidelity (Razavi et al., 2012). Various methods, such as artificial neural networks (ANN), support vector machines (SVM), and response matrices (RM), have been used to surrogate process-based models (Cai et al., 2015; Housh et al., 2014; Li et al., 2021b; Zhang et al., 2009), each with its own limitations. For example, most applications of ANN and SVM for surrogating watershed models have only considered a small number of inputs and outputs. Zhang et al. (2009) used ANN and SVM to surrogate a SWAT model for parameterization, mapping combinations of 16 parameters to model effectiveness at predicting total basin runoff. Cai et al. (2015) used SVM to surrogate a SWAT model to optimize decision-making under climate uncertainty, mapping four management decisions to four measures of basin-scale drought impact. Training an SVM or ANN to surrogate spatially distributed decision-making (e.g., 50 conservation practices per subwatershed * 40 subwatersheds = 2000 decision variables) and spatially distributed and temporally refined outputs (e.g., 120 months * 40 subwatersheds = 4800 monthly outputs) would require an exorbitant

* Corresponding author. Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, 2030 Civil Engineering Hydrosystems Lab, Urbana, IL, 61801, USA.

E-mail address: xmcai@illinois.edu (X. Cai).

[1] co-senior author.

https://doi.org/10.1016/j.envsoft.2021.105269
Received 2 June 2021; Received in revised form 28 September 2021; Accepted 6 December 2021
Available online 14 December 2021
1364-8152/

number of simulations.

The traditional RM method is a suitable surrogate method for generating a large set of spatially distributed and dynamic hydrological responses under spatially distributed agricultural management decisions. RM methods approximate hydrological responses as a linear function of a (potentially large) set of distributed management decisions and have been used to surrogate hydrological models, such as groundwater models (Gorelick, 1983; Maddock, 1972; Yoon et al., 2021), integrated surface water and groundwater model (Seo et al., 2018), and watershed models (Gorelick et al., 2019; Housh et al., 2014, 2015; Shafiee-Jood et al., 2018). Although those studies do not explicitly name their approaches as "response matrix"; however, their approaches hold the core idea of response matrix, assuming that hydrological responses are approximated by linear functions of management decisions. The linear nature of the RM method enables the use of linear programming for optimization applications, saving considerable computational resources. For example, Housh et al. (2014) applied an RM approximation within a mixed-integer linear programming model for biofuel development considering more than 10,000 decision variables. Even in optimization applications that do not use linear programming, RM approximations have been adopted for large spatial problems because assumed independence between spatially distributed decisions (a condition of linearity) provides a path to estimate overall outcomes with substantially fewer simulations. For example, Gorelick et al. (2019) and Gramig et al. (2013) coupled a genetic algorithm with RM approximations that map management decisions to watershed sediment or nutrient loss. However, watershed modeling applications of the traditional RM method have not explored potentially co-located conservation practices and have sparingly addressed the impacts of in-stream source-model processes. In a rare example of treating in-stream processes within an RM framework, Femeena et al. (2018) loosely coupled SWAT landscape outputs with an exponential decay model for in-stream nutrient processes. The authors suggest that more efforts should be invested when loosely coupling SWAT results to better consider in-stream processes. Co-located practices and in-stream processes introduce nonlinearities and interactions between decisions, and thus they may reduce the effectiveness of RM-based approaches. For studies concerning diverse conservation practices or watershed export responses, the traditional RM method based on linear approximation must be validated or modified (Femeena et al., 2018).

The overall goal of this study is to provide insights from our experience that converts a distributed hydrological model (SWAT) to a RM-based surrogate model and demonstrate how to surrogate SWAT in a reasonable way that addresses the interaction effects between co-located conservation practices and in-stream and reservoir processes affecting nutrient and sediment loads. The specific study objectives are to provide a revised RM method, compare it with traditional RM, and discuss the application and limitations of the revised RM. For the present context, the key assumption of the RM surrogate model is that, given a known weather scenario, watershed hydrological responses may be reasonably represented by linear functions of agricultural management decisions. The validity of and modifications required for satisfying this assumption are the cornerstones of the discussion in this paper. In the sections to follow, we first provide relevant background regarding SWAT and the traditional response matrix method (Section 2). In Section 3, we explore the consequences of co-located conservation practices (Section 3.1) and in-stream and reservoir processing (Sections 3.2 and 3.3) for the validity of the traditional RM method and, accordingly, illustrate modified RM approaches. We also discuss how these findings and adjustments are shaped by our specific research questions and hydrological model (i.e., SWAT), so as to increase the transferability of the process and findings.

2. Background

2.1. Modeling a testbed watershed in SWAT

SWAT is a semi-distributed hydrologic model designed to evaluate and predict the impacts of agricultural management practices on water, sediments, and pollutants (Arnold et al., 2012). SWAT discretizes the watershed into user-defined subwatersheds and, within each subwatershed, into hydrologic response units (HRUs) which share a common land use, soil type, and slope. For each HRU, the model independently (1) implements management practices, (2) applies external weather forcing (uniform within subwatersheds), and (3) simulates surface and sub-surface hydrology, plant growth, nutrient transformations, and nutrient and sediment transport. In SWAT2012, HRU water, nutrient, and sediment yields are aggregated at the subwatershed level and routed directly to the subwatershed-level stream reach (within each subwatershed the stream network is consolidated to a single reach). Note that, for SWAT+, the latest version of SWAT model currently available, HRU yields are aggregated at the landscape unit (LSU) level and routed through floodplains into "channels" before entering the subwatershed-level stream reach. Finally, SWAT simulates flow, nutrient, and sediment routing as well as simplified water quality transformations for every stream reach.

Here we use a SWAT2012 model set for the Upper Sangamon River Watershed (USRW), located in central Illinois, USA. The USRW is predominantly operated for corn and soybean rotation (80% of the total 3680 km^2 watershed area) and is extensively tile-drained (USDA NASS, 2019). The Lake Decatur dam, in the downstream half of the watershed (see Fig. 1), supplies municipal and industrial water to the city of Decatur and nearby bioethanol producers (Fitzpatrick et al., 1987). Sediment deposition in Lake Decatur reduces active storage, impairs water quality, and requires costly dredging to remediate. Residential and bioethanol facility wastewaters are treated by the Sanitary District of Decatur (SDD) and discharged downstream of Lake Decatur. Notably, phosphorus (P) discharge from SDD constitutes over 70 percent of P loads at the watershed outlet (680 Mg P/year SDD discharge versus 940 Mg P/year total watershed export), in part due to significant trapping of upstream P loads in the Lake Decatur reservoir.

The SWAT model used in this study is calibrated for flow, crop yield, sediments, nitrate, and total phosphorus (TP) at multiple sites during the period from 2003 to 2012. The daily flow is calibrated at four sites: Fisher, Monticello, Decatur and the watershed outlet. Flow calibration shows satisfactory model performance (Moriasi et al., 2007) at all sites with Nash-Sutcliffe efficiencies (NSE) ≥ 0.65 and percent bias within $\pm 10\%$. The model then follows the calibration of annual crop yield for both corn and soy in Macon County with a reasonable model performance. Furthermore, the model shows satisfactory performance after the calibration of monthly sediments, nitrate and TP at Monticello, Decatur, and Wyckles Bridge (downstream of SDD) with NSE ranging from 0.63 to 0.83 and bias mostly within $\pm 10\%$. Lastly, the model parameters are readjusted with a global run for flow, crop yield, sediments, nitrate, and TP at all the calibration sites. The model performs reasonably well at these sites in the validation period (2013–2018).

2.2. Application of the traditional response matrix method

Due to the nature in which SWAT discretizes and simulates watershed processes (described above), there is no interaction between HRUs until runoff mixes within the stream reach. Thus, the "pre-stream" portion of SWAT (i.e., the land-surface model processes which proceed routing into and through the stream network) is inherently compatible with the independence criterion of a linear model (e.g., a response matrix), if the decision variables are modeled at the HRU level. To reduce computational load in optimization applications, a user might desire to aggregate decisions to the subwatershed level – that is, to decide the percent of each subwatershed on which each conservation

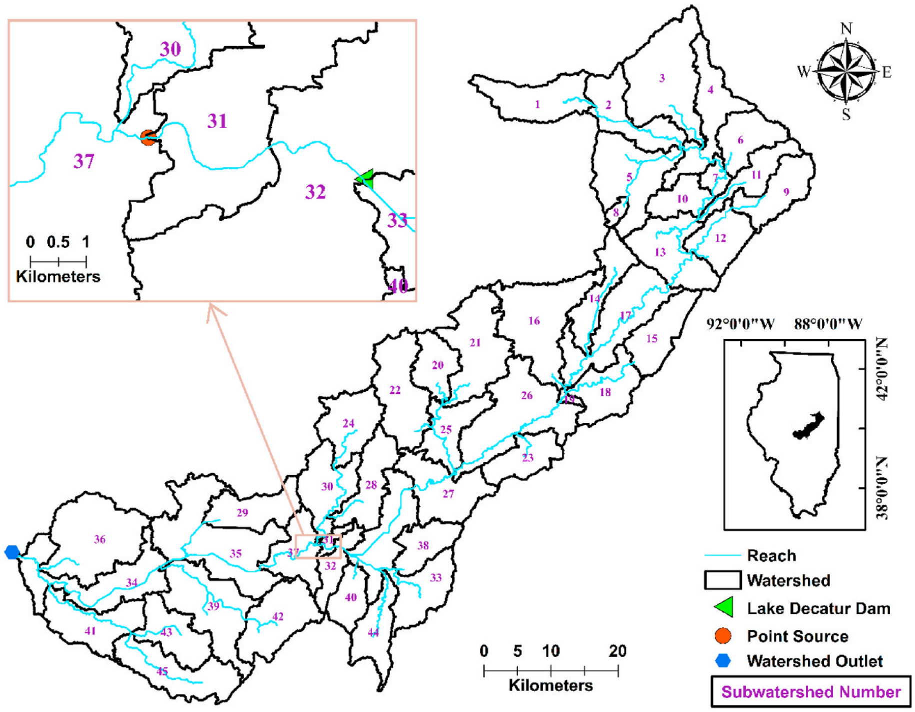

Fig. 1. Map of the Upper Sangamon River Watershed (USRW) with 45 subwatersheds.

practice is applied rather than which practices are applied on each HRU (Housh et al., 2014). Our study follows Housh et al. (2014) in this regard.

The traditional response matrix approach, formally presented below following (Housh et al., 2014), takes advantage of HRU independence in SWAT to approximate land surface yields as the product of agricultural practice land area fractions and their siloed impacts (that is, their impacts when adopted in isolation, without other conservation practices). Housh et al. (2014) applied linear production functions to determine the total nitrate and runoff contribution from subwatersheds, which follows the same method of the response matrix, that is, approximating hydrological responses using linear functions of management decisions. The primary shortcomings of this traditional approach, expanded upon in Section 3, are that it (1) does not clearly describe how to account for interaction effects between co-located conservation practices, and (2) does not account for interaction effects and other sources of nonlinearity originating within the stream network.

The steps for implementing the traditional RM method are as follows. Step 1, for each conservation practice under consideration, run a SWAT model for the watershed where the practice is adopted on all agricultural HRUs. Step 2, construct response matrices: for each conservation practice, hydrologic variable (e.g., nutrients, streamflow), and month, construct a single response matrix where elements along the diagonal are the SWAT landscape yields for each subwatershed ($\boldsymbol{Y}_{m,t}$). Step 3, construct decision vectors for each conservation practice ($\boldsymbol{F}_m$) where the elements are the area fractions that are allocated to that particular practice at a subwatershed level. Step 4, calculate the agricultural land area ($\boldsymbol{A}$) for each subwatershed (or depending on the modeled practice, perhaps the total or urban area, where appropriate). Step 5, construct a connectivity matrix ($\boldsymbol{W}$) describing the upstream-downstream relationships of all subwatersheds, with off-diagonal elements $w_{i,j|i\neq j}$ equal to one if subwatershed j is upstream of subwatershed i and zero otherwise ($w_{i,i} = 1\ \forall i$). Step 6, apply Equation (1) to estimate landscape yield during month t across all subwatersheds. Step 7, apply Equation (2) to estimate in-stream loads at the outlet of each subwatershed by summing its own yield and all upstream yields. Because SWAT models are typically calibrated and analyzed at the monthly scale, we present the RM method as a surrogate for monthly SWAT outputs, following Housh et al. (2014). For a model considering $\boldsymbol{M}$ possible conservation practices in $\boldsymbol{N}$ subwatersheds, the subwatershed landscape yield during month t ($\boldsymbol{q}, \boldsymbol{p}, \boldsymbol{ni}, \boldsymbol{s} \in \boldsymbol{R}^N$) is calculated as the sum-of-products of each practice's area allocations and response matrices:

$$\boldsymbol{q}_t = \sum_{m\in M} diag\left(\boldsymbol{Y}_{m,t}\right)_Q \cdot diag(\boldsymbol{A}) \cdot \boldsymbol{F}_m \quad \forall t \in T \tag{1}$$

where the function $diag(\cdot)$ converts the vector argument (with elements $\boldsymbol{Y}_i$ *or* $\boldsymbol{A}_i$) into a diagonal matrix (with elements $\boldsymbol{X}_{ii} = \boldsymbol{Y}_i$ *or* $\boldsymbol{A}_i$); $\boldsymbol{Y}_{m,t} \in \boldsymbol{R}^N$ are response vectors of SWAT subwatershed yield outputs for conservation practice m during month t; $\boldsymbol{A} \in \boldsymbol{R}^N$ is a vector containing the total agricultural area of each subwatershed; and $\boldsymbol{F}_m \in \boldsymbol{R}^N$ are decision vectors indicating the fraction of subwatershed areas allocated to conservation practice m. Under the traditional formulation, the in-stream loads at the outlet of each subwatershed ($\boldsymbol{Q}, \boldsymbol{P}, \boldsymbol{Ni}, \boldsymbol{S} \in \boldsymbol{R}^N$) are the sum of all upstream landscape yields:

$$\boldsymbol{Q}_t = \boldsymbol{W} \cdot \boldsymbol{q}_t \quad \forall t \in T \tag{2}$$

where $\boldsymbol{W} \in \boldsymbol{R}^{N\times N}$ is a connectivity matrix with off-diagonal elements $w_{i,j|i\neq j}$ equal to one if subwatershed j is upstream of subwatershed i and zero otherwise ($w_{i,i} = 1\ \forall i$). Equations (1) and (2) are presented for the case of estimating flow ($\boldsymbol{Q}$), but are applied likewise for any modeled output variable, e.g. phosphorus ($\boldsymbol{P}$), nitrogen ($\boldsymbol{Ni}$), or sediment ($\boldsymbol{S}$).

This application of the RM method estimates the impact of partial practice adoption in a subwatershed according to the *aggregate impact* of complete adoption across the whole subwatershed. In SWAT, the impacts might be distinct to slopes and soil types (i.e. HRUs) within the subwatershed. Therefore, the subwatershed-level RM estimates most accurately represent a hypothetical SWAT scenario where equal parts of each HRU within a subwatershed adopt a practice. For example, the RM model decision to apply cover crops on 20% of agricultural land in a subwatershed corresponds to the SWAT implementation of cover crops on 20% of the land in each agricultural HRU in the subwatershed. In actual SWAT use, practices are assigned at the HRU level and for entire HRUs. This dissonance cannot be resolved for RM models framed at the subwatershed-level, since the framework does not contain requisite HRU information. The trade-off for this dissonance at the subwatershed level is the aforementioned benefit of reduced decision variables when optimizing. For applications where soil- and slope-specific decisions are more desirable than reduced decision variables, all methods discussed herein may be applied by allowing $\boldsymbol{N}$ to be the number of HRUs and using HRU-level SWAT outputs. Moreover, all results and discussion of modifications to the RM method presented below remain applicable, none are specific to the subwatershed-level application.

Fig. 2 demonstrates the output-specific effectiveness and limitations of the traditional RM method for approximating streamflow, nitrate, total phosphorus, and sediment at the monthly scale at the outlet of the USRW. As shown in Fig. 2a and b, the traditional RM method effectively reproduces SWAT streamflow and nitrate outputs. However, Fig. 2c and d illustrate that the traditional RM method does *not* effectively reproduce SWAT phosphorus or sediment outputs. The sediment Nash-Sutcliffe Efficiency (NSE) indicates that the traditional RM surrogates SWAT far worse than simply the long-term mean (NSE $\ll$ 0), and the percent bias (P-bias) indicates that sediment is vastly overestimated. While the phosphorus NSE and P-bias indicate the traditional RM may be an effective surrogate in general, the RM estimates systematically overestimate P load during periods of low flow, which in some settings are the most critical periods for nutrient management. In Section 3, we demonstrate how interactions between co-located conservation practices and nonlinearities within the stream network contribute to traditional RM shortcomings and how we modify the RM approach to better approximate SWAT phosphorus and sediment outputs.

3. Modifying the response matrix approach

In this section, we illustrate the primary shortcomings of the traditional RM approach and propose three modifications to the RM approach to better approximate SWAT-simulated phosphorus and sediment export. While we provide brief comment on approximating flow and nitrogen, we focus the discussion on phosphorus and sediment because these variables are not adequately approximated by the traditional RM approach.

3.1. Modification 1: Dealing with interaction effects between co-located conservation practices

When multiple conservation measures are applied at the same location, the total impact is not likely to equal the sum of its parts, though the impacts are typically complementary (Boreux et al., 2013; Chaubey et al., 2010; Illinois Environmental Protection Agency et al., 2015). For instance, implementing conservation tillage may affect the nutrient reductions achieved from reducing fertilizer applications (Jarvie et al., 2017). To the extent that these interaction effects are captured in SWAT, the traditional RM approach – where impacts of siloed practices are simply added – would not accurately capture the total impact of co-located conservation practices.

Therefore, we test a modified the RM approach where each unique combination of conservation practices is simulated to generate unique response matrices for each combination – rather than just simulating and generating response matrices for the siloed, individual practices. Likewise, the land allocation decision vector of the RM model is reformulated to contain fractions allocated to each possible conservation practice *combination*. Under this modification, according to the syntax of Equation (1), $\boldsymbol{M}$ now represents the set of all possible combinations of conservation practices.

Here, we illustrate the interaction effects between conservation practices by comparing traditional response matrix model estimates and SWAT-simulated landscape yield for a case of co-located practices. The baseline management practices for the USRW are two-year corn-soybean rotations, conventional tillage, 207 kg/ha diammonium phosphate application preceding corn years, no cover crops, and no vegetative filter strips. We compare estimates for the case of co-located filter strips, cover crops, and fertilizer reduction throughout the entire watershed. For this case, the traditional RMs are derived from SWAT-simulated impacts of siloed filter strips, siloed cover crops, and siloed fertilizer reduction (see scenarios 1–3 in Table 1). On the other hand, the modified RMs would be derived directly from the SWAT-simulated impact of all three practices implemented together (scenario 4 in Table 1); that is, for the reasons

Table 1
Conservation practice scenarios simulated with SWAT to generate response matrices.

Scenario	Cover crop	Fertilizer reduction percentage	Filter strips
Baseline		0	
1		0	x
2	x	0	
3		30	
4	x	30	x

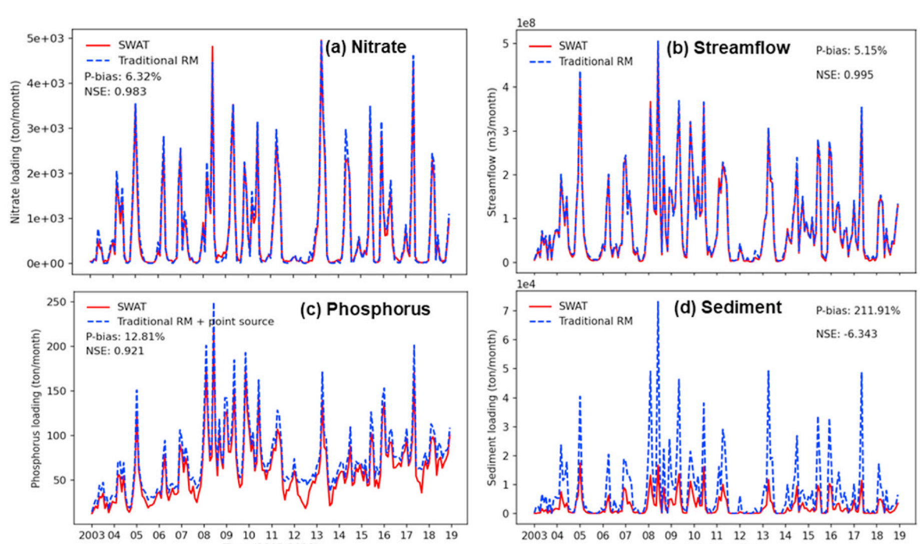

Fig. 2. Simulation results of in-stream loads of nutrient and sediment and streamflow at the outlet of Upper Sangamon River Watershed (USRW) using the traditional response matrix method compared to the direct use of SWAT: (a) nitrate; (b) streamflow; (c) phosphorus; (d) sediment.

discussed in Section 2.2, the modified RM estimate exactly matches the landscape yield (i.e. "pre-stream") simulated in SWAT.

Fig. 3 shows the estimated annual sediment and phosphorus reductions at the USRW outlet if the described, co-located conservation practices had been applied during the years 2003–2018. If there were no interaction effects among conservation practices, the traditional RM approach would provide the same estimate as the SWAT simulation. Instead, Fig. 3a and c indicate that, on average, the siloed, traditional RM approach overestimates the SWAT-simulated sediment and phosphorus yield reductions by 19.75 percent and 16.6 percent, respectively. Fig. 3b and d further reveal that the overestimation is greatest in years with high water yield (and thus also high sediment and phosphorus yield). The combined impact of filter strips, cover crops, and fertilizer reductions is thus demonstrably less than the sum of their siloed impacts according to SWAT simulations, and this emergent SWAT outcome is captured by the modified RM approach but not the traditional approach. Per Fig. 2a, interaction effects do not seem meaningful for nitrogen, perhaps implying that interaction effects in SWAT are primarily mediated by sediment loss, to which phosphorus is more tightly coupled than nitrogen.

3.2. Modification 2: Dealing with impacts of in-stream and reservoir processes for flow and nutrient

If flow and nutrient processes in each SWAT stream reach or water body (e.g. reservoir) behave as a linear system, then the traditional response matrix method can be modified by incorporating linear in-stream and water body processing effects within the channel network connectivity matrix (see Section 2.2). The degree to which a linear approximation holds for these processes dictates the potential effectiveness of any RM-based method.

3.2.1. SWAT in-stream and reservoir processes for flow and nutrients

3.2.1.1. Streamflow. For streamflow, SWAT's in-stream and water body (e.g. reservoir) simulation routines include routing, seepage, and evaporation. SWAT simulates seepage and evaporation losses as the product of a seepage or evaporation rate coefficient (k_Q, [mm/hr]), the exposed area (i.e. wetted stream area or surface area, A, [m^2]), and the residence time in the reach or water body (ΔT, [hr]). That is, they share the general form:

$$\Delta \boldsymbol{Q} = k_Q \cdot A \cdot \Delta T \tag{3}$$

where $\boldsymbol{Q}$ is the volume of water in the reach. Since both residence time and wetted perimeter/surface area are (nonlinear) functions of inflow themselves, seepage and evaporation are each the product of rate constants and potentially nonlinear functions of flow:

$$\Delta \boldsymbol{Q} = k_Q \cdot A \cdot \Delta T = k_Q \cdot f_1(\boldsymbol{Q}) \cdot f_2(\boldsymbol{Q}) \tag{4}$$

The effectiveness of a modified RM method for flow thus depends upon the accuracy of a linear approximation such as:

$$\sum \boldsymbol{Q} = \sum_i \left(k_{Q,i} \cdot f_{1,i}(Q) \cdot f_2(Q) \right) \approx \widehat{k}_Q \cdot \boldsymbol{Q} \tag{5}$$

where $\widehat{k}_{Q,i}$ [-] is the aggregate fraction of flow lost due to seepage (i = 1) and evaporation (i = 2) at typical conditions. For the scale of the present application, $\boldsymbol{\Delta Q}$ is a vector of water lost [m^3] from each reach, applied for each month of analysis. In addition, an RM-based approach is not able to account for changes in stream or reservoir storage between time steps; while assuming zero storage change is likely reasonable for annual-scale estimates, the assumption is plausibly problematic for monthly-scale estimates. For the USRW though, accurate estimation by the traditional RM approach (see Fig. 2) implies that monthly storage change is negligible and that $\widehat{\boldsymbol{k}}_Q$ is a zero vector, i.e. seepage and evaporation are also negligible.

3.2.1.2. Phosphorus. On the other hand, while somewhat similarly formulated in SWAT, phosphorus is not adequately estimated by the traditional RM approach. For stream reaches, SWAT simulates phosphorus settling, mineralization, exchange with algae, and release from benthos. These processes are formulated either as a (1) first-order reaction (i.e. $\Delta P = k_P P$, where P is the mass of phosphorus in the reach and k_P is the rate coefficient) or (2) zeroth-order reaction with dependence on flow depth or algae concentration (i.e. $\Delta P = k_P \cdot f(depth)$ *or* $\Delta P = k_P \cdot f(algae)$). Similar to flow though, the amount of phosphorus transformed or exchanged during transport also depends on the residence time within the reach; therefore, these processes take the generalized forms:

$$\Delta \boldsymbol{P} = k_{P,P} \cdot \Delta T \cdot \boldsymbol{P} \tag{6a}$$

$$or \quad \Delta \boldsymbol{P} = k_{P,Q} \cdot f_3(depth, algae) \cdot \Delta T \tag{6b}$$

where $k_{P,P}$ [hr^{-1}] and $k_{P,Q}$ [units vary] are the rates of phosphorus lost within the stream for a given process. Note that *in SWAT* phosphorus and sediment transport are completely de-coupled once they have reached the stream network. While (in reality and in SWAT) sediment lost from the landscape exercises greatly influences phosphorus lost from the landscape, sediment deposition and erosion have no effect on phosphorus transport in SWAT (though the two are indeed coupled in

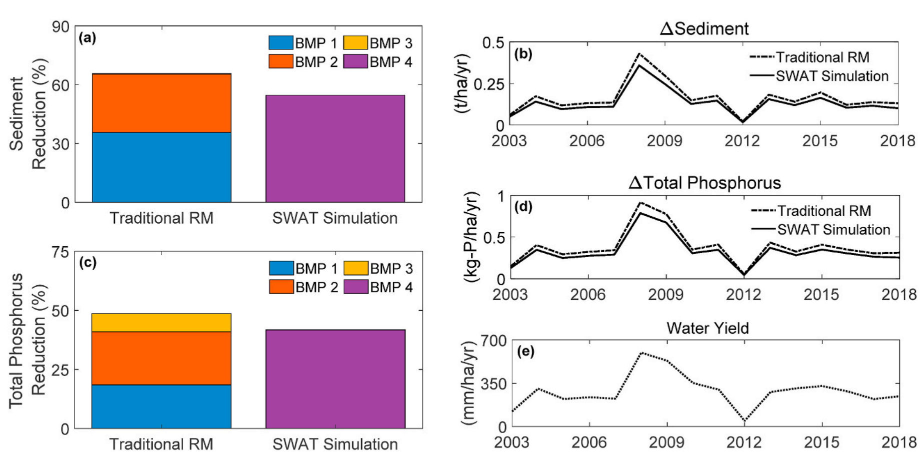

Fig. 3. Comparison of landscape loss reductions from co-located conservation practices, according to traditional RM estimate versus SWAT simulations. (a) Average annual reduction in sediment yield; (b) Annual reduction in sediment yield; (c) average annual reduction in phosphorus yield; (d) annual reduction in phosphorus yield; (e) annual water yield.

reality). Since we aim to surrogate SWAT, and only indirectly to estimate real-world loads, we do not include sediment as a predictor for in-stream phosphorus processing.

Therefore, we test a linear regression with one phosphorus-dependent term and one flow-dependent term. We expect that (1) the phosphorus-dependent term should increase in magnitude as incoming phosphorus load increases, due to the first-order reaction basis, and (2) the flow-dependent term should increase in magnitude as incoming flow *decreases*, due to the inverse relationship between flow and residence time. Therefore, a linear approximation around which to build the modified RM formulation might be:

$$\Delta \boldsymbol{P} \approx \widehat{k}_{P,P} \cdot \boldsymbol{P} + \widehat{k}_{P,\frac{1}{Q}} \cdot \boldsymbol{Q}_{inv} \tag{7a}$$

$$\boldsymbol{Q}_{inv} = \boldsymbol{W} \cdot \boldsymbol{q}_{inv} \tag{7b}$$

$$\boldsymbol{q}_{inv\,t} = \sum_{m \in M} diag\left(\boldsymbol{Y}_{m,t}\right)_{\frac{1}{Q}} \cdot diag(\boldsymbol{A}) \cdot \boldsymbol{F}_m \quad \forall t \in T \tag{7c}$$

where $\widehat{k}_{P,P}$ [-] is the phosphorus-yield-dependent fraction of phosphorus lost due to settling at typical conditions; $\widehat{k}_{P,\frac{1}{Q}}$ [mg P/L] is the aggregate water-yield-dependent (i.e. flow-dependent) fraction of phosphorus lost due to algal uptake (j = 1), algal decomposition (j = 2), and benthic uptake (j = 3) at typical conditions; $\boldsymbol{Q}_{inv} \in \boldsymbol{R}^N$ are vectors analogous to cumulative residence time at each subwatershed outlet during month t; $\boldsymbol{q}_{inv} \in \boldsymbol{R}^N$ are vectors analogous to residence time in each subwatershed in month t; and $(\boldsymbol{Y}_{m,t})_{\frac{1}{Q}} \in \boldsymbol{R}^N$ are response vectors for the inverse of water yield in each subwatershed for conservation practice *m* during month t. While this formulation for $\boldsymbol{q}_{inv}$ does not precisely represent the mean residence time (i.e. $\boldsymbol{q}_{inv}$ does not equal $\boldsymbol{q}^{-1}$), it preserves a linear relationship between the decision variables $\boldsymbol{F}_m$ and the estimated output $\boldsymbol{P}$.

For water bodies, SWAT2012 simulates settling only, with no nutrient transformations. As with in-stream settling, settling in water bodies is formulated as a first-order reaction with respect to the body's nutrient concentration and depends on the residence time in the reservoir. Distinctly, the settling rate in water bodies scales linearly with the water body area, and the residence time can exceed the model time step – therefore settling depends geometrically upon residence time

$$\Delta \boldsymbol{P} = \left(1 - \left(1 - k_{P,R}\right)^{\frac{\Delta T}{\Delta t}}\right) \cdot A \cdot \boldsymbol{P} \tag{8}$$

where $k_{P,R}$ [1/hr/m^2] is the reservoir trapping rate per unit area; Δt [hr] is the model time step.

Therefore, with regard to reservoir trapping, the effectiveness of a modified RM method depends upon the accuracy of the linear approximation:

$$\Delta \boldsymbol{P} \approx \widehat{k}_{P,R} \cdot \boldsymbol{P} \tag{9}$$

where $k_{P,R}$ [-] is the fraction of phosphorus trapped by the reservoir at typical conditions. Notably, because the residence time in water bodies is much greater in magnitude and variance than in streams, the simulated reservoir trapping may also be much larger in magnitude and span. Therefore, it is not immediately clear whether this approximation at typical conditions will hold well.

Equations (6)–(10) are generalized forms provided and discussed as the background and justifications for the modifications we make in the following section. In Section 3.2.2, we evaluate the validity of the phosphorus in-stream approximation (Equation (7)) and phosphorus reservoir trapping approximation (formulation provided in section 3.2.1) in concert. For watersheds where nitrogen is not adequately estimated by the traditional RM approach, the same formulation and validation process may be applied as that for phosphorus.

3.2.2. *Incorporation of effects of point sources, reservoirs, and in-stream processes in phosphorus estimation*

Here we present phosphorus response matrix modifications that adjust for 1) point-source discharges, 2) reservoir trapping, and 3) in-stream processes. The modifications are embedded within the framework of Equations (1) and (2) by adding additional sources contributing to subwatershed yield ($\boldsymbol{p}$), scaling the phosphorus-yield response vectors ($(\boldsymbol{Y}_{m,t})_P$), adding dependence on the inverse water-yield response vectors ($(\boldsymbol{Y}_{m,t})_{\frac{1}{Q}}$), and selectively scaling elements of the connectivity matrix ($\boldsymbol{W}$). Fig. 4 illustrates a simple conceptual watershed with the features considered by the modified response matrix approach for phosphorus. Greater attention is given to reservoir trapping and in-stream processes since adding point-source impacts is straightforward: point-source loads are simply added to the subwatershed yield (Equation (1)) where the point sources are located:

$$\boldsymbol{p}_t = \sum_{m \in M}\left(diag\left(\boldsymbol{Y}_{m,t}\right)_Q \cdot diag(\boldsymbol{A}) \cdot \boldsymbol{F}_m\right) + \boldsymbol{p}_{t,\,ps} = \boldsymbol{p}_{t,nps} + \boldsymbol{p}_{t,\,ps} \quad \forall t \in T \tag{10}$$

where $\boldsymbol{p}_{t,\,nps} \in \boldsymbol{R}^N$ is the total non-point source yield from each subwatershed in month t and $\boldsymbol{p}_{t,\,ps} \in \boldsymbol{R}^N$ is the total point source yield from each subwatershed in month t. The inclusion of point-sources is trivial methodologically to the point that we include it in Fig. 2 when comparing traditional RM to SWAT simulations.

We evaluate the linear approximation for reservoir trapping (Equation (9)) in the USRW baseline scenario by conducting a linear regression for phosphorus effluent according to phosphorus influent. Including an intercept term to allow for some minimum trapped load, the regression performs very well (R-squared > 0.95, see Fig. 5):

$$P^{out}_{t,\,res} = \left(1 - \widehat{k}_{P,R,1}\right) \times P^{in}_{t,res} - \widehat{k}_{P,R,2} \quad \forall t \in T \tag{11}$$

where $P^{in}_{t,res}$ and $P^{out}_{t,res}$ are the simulated phosphorus loads into and out of the reservoir at time *t* and $\widehat{k}_{P,R,1}$ and $\widehat{k}_{P,R,2}$ are linear regression coefficients. The accuracy of the linear regression indicates that a linear filter, compatible with the response matrix approach, may reasonably approximate the trapping effect. Because the trapping efficiency (i.e. $\widehat{k}_{P,R,1}$) is time-invariant, it may be incorporated directly within the (also time-invariant) stream connectivity matrix. For all phosphorus stream connectivity matrix elements $(w_{i,j})_P$ such that $i \in \boldsymbol{D}$ and $j \in \boldsymbol{U}$, where $\boldsymbol{D}$ is the set of stream reaches downstream from the reservoir and $\boldsymbol{U}$ is the set of reaches upstream of the reservoir, we set $w_{i,j}$ equal to one minus the

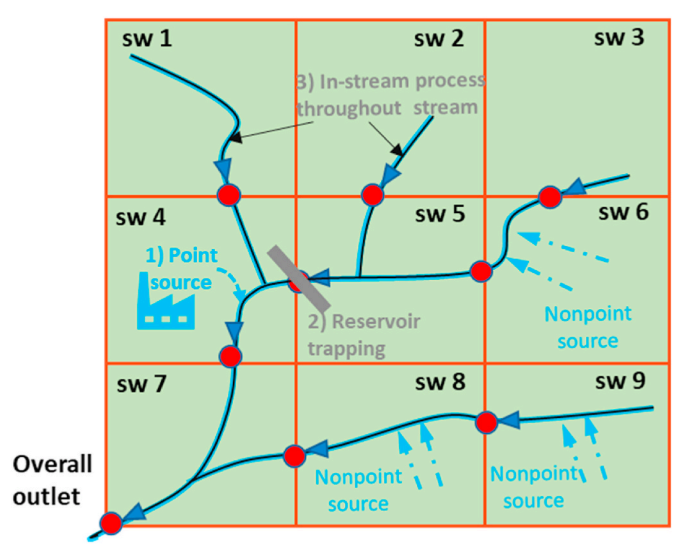

Fig. 4. A graphical example of a simplified watershed with 9 subwatersheds considering: 1) point source discharge, 2) reservoir trapping, and 3) in-stream process. sw: subwatershed.

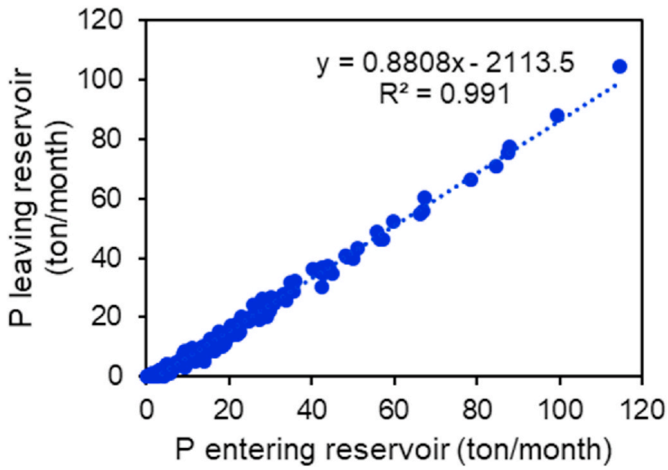

Fig. 5. Linear filtering of reservoir on monthly phosphorus (P) load.

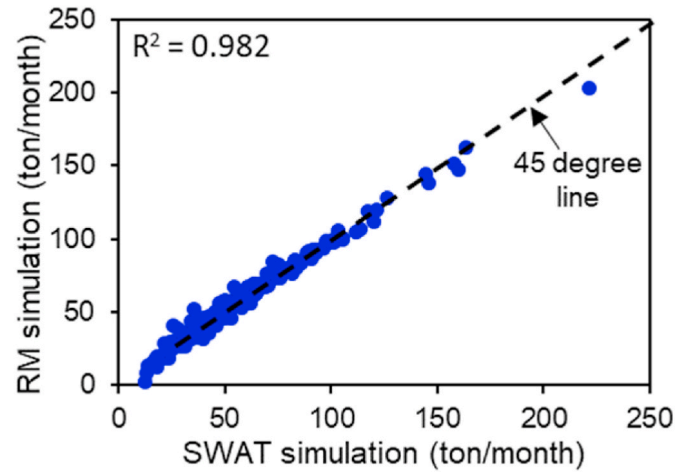

Fig. 6. Performance comparison of in-stream phosphorus (P) loading at outlet between SWAT simulation and RM simulation using linear regression (Eq. [14]).

trapping efficiency (rather than 1, as had been before):

$$(\boldsymbol{w}_{i,j})_P = \begin{cases} \left(1-\widehat{k}_{P,R,1}\right) & \text{if } i \in D \text{ and } j \in U \\ 1 & \text{otherwise, if } j \text{ is upstream of } i \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

Meanwhile, the minimum trapped load (i.e. $\widehat{k}_{P,R,1}$) is subtracted directly from the yield of the downstream subwatershed, as if a "point sink":

$$\boldsymbol{p}_t = \boldsymbol{p}_{t,nps} + \boldsymbol{p}_{t,\,ps} - \boldsymbol{p}_{t,\,res} \quad \forall t \in T \tag{13}$$

where the minimum trapped load, $\boldsymbol{p}_{t,\,res} \in \boldsymbol{R}^N$, is equal to $\widehat{k}_{P,R,2}$ for the subwatershed immediately downstream of the reservoir and zero otherwise.

We evaluate the linear approximation for in-stream processing (Equation (7)) in the USRW baseline scenario by conducting linear regressions (at each subwatershed outlet) for phosphorus effluent according to upstream phosphorus yields and "inverse-flow" yields:

$$\boldsymbol{P}_t = \left(I_N - \widehat{k}_{P,P}\right)\cdot \boldsymbol{W}\cdot \boldsymbol{p}_t + \left(I_N - \widehat{k}_{P,\frac{1}{Q}}\right)\cdot \boldsymbol{W}\cdot \boldsymbol{q}_{invt} \; \forall t \in T \tag{14}$$

where $\widehat{k}_{P,P} \in \boldsymbol{R}^{NxN}$ and $\widehat{k}_{P,\frac{1}{Q}} \in \boldsymbol{R}^{NxN}$ are diagonal matrices whose elements are the regression coefficients estimated using SWAT simulation data (i.e., P yield and streamflow) and indicating the fraction of phosphorus lost in a stream; and $I_N \in \boldsymbol{R}^{NxN}$ is an identity matrix. Recall that $\boldsymbol{W}$ is the connectivity matrix accounting for upstream-downstream relationships. The right-hand side of Equation (14) is then the P export at each reach broken into a term dependent on upstream landscape P loading and a term dependent on upstream cumulative residence time. Larger values of $\widehat{k}_{P,P}$ and $\widehat{k}_{P,\frac{1}{Q}}$ indicate that the stream acts as more of a phosphorus sink. As with the traditional RM method, we apply equations (10)–(14) at the monthly scale. We find again that the regression performs very well (R-squared > 0.95, see Fig. 6), suggesting that Equations 10 and 12-14 constitute an effective modified response matrix approach for phosphorus.

Below, we compare SWAT simulation results for monthly in-stream phosphorus loads at the USRW outlet with estimates from both the modified RM formulation and a "traditional + point source" estimate. We choose to show the "traditional + point source" estimate rather than the traditional RM estimate because (1) the point source addition method is trivial and (2) the point source load in the USRW is so large that its omission obscures the value of the other modifications. First, for SWAT simulations, we randomly select one of four conservation practice combinations from Table 1 (filter strips only, cover crops only, fertilizer reduction only, or all practices together) for every agricultural HRU in the watershed. Then, for the RM estimates, we apply the resultant subwatershed land allocation fractions. As described in section 2.2, the RM estimates assume that this fraction of land allocated to a conservation practice is *evenly distributed among all HRUs* within the subwatershed. Some discrepancy between the RM estimates and SWAT simulations may be attributed to this allocation distinction and differences in how conservation practices impact yields in different HRUs (i.e. on different soils and slopes). We measure the surrogate accuracy of the RM approaches by Nash-Sutcliffe efficiency (NSE) and percent bias (P-bias) between the respective RM estimated time-series and the simulated time-series (Moriasi et al., 2007).

We find that the modified RM formulation provides more accurate and less-biased estimates (NSE = 0.98 and P-bias = 0.9%) than the "traditional + point source" formulation (NSE = 0.92 and P-bias = 12.8%) (Fig. 7). Furthermore, the modified RM formulation provides *drastically* more representative estimates during periods of low flow. The difference between the modified RM and "traditional + point source" estimates is most pronounced during the 2011–2012 drought. From July

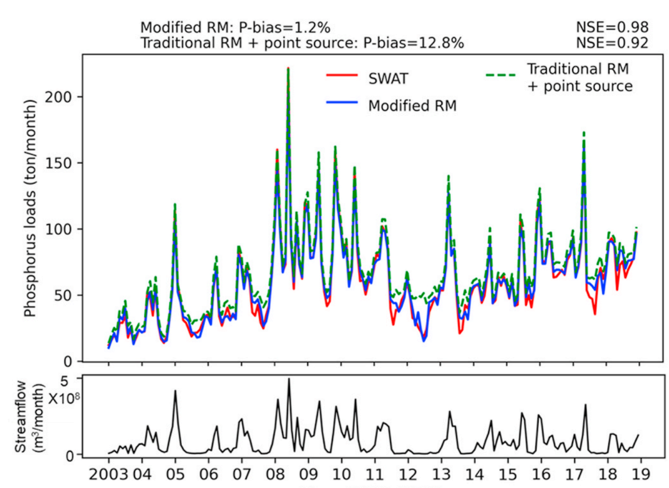

Fig. 7. Performance of modified RM and traditional RM on simulating in-stream phosphorus load at the outlet. The monthly streamflow figure is provided at the bottom to demonstrate the impact of low flow on P discrepancy between modified RM and traditional RM.

2011 to September 2012, SWAT simulates 541,950 kg total phosphorus export from the watershed. The modified RM estimate is 561,562 kg export (3.6% over-estimate) while the "traditional + point source" estimate is 772,064 kg export (42.5% over-estimate). This focused improvement from the modified RM approach during low flows is to be expected – the effects of in-stream processes and reservoir trapping are greatest when residence time is largest. Moreover, this focused improvement is relevant and important since streams and water bodies receiving significant point source discharge can be most vulnerable to harmful algal blooms during low-flow periods (Harrison et al., 2019; Jarvie et al., 2006). The cumulative impact of improved prediction at low flows is a relatively unbiased estimate for total export over the 16 years, rather than the 12.8% over-prediction by the "traditional + point source" method.

We repeat the above evaluation process 100 times for different realizations of randomized conservation practice allocations using modified RM method, and compare the results with the same randomized allocations in SWAT to evaluate the robustness of RM performance under randomized allocations. These realizations provide a more comprehensive picture of the modified RM performance and illustrate the impacts of the conservation practice allocation methods discussed above. Fig. 8 presents the mean and range for the performance metrics (NSE and P-bias) across the 100 realizations and every subwatershed outlet. Overall, the modified RM method has satisfactory performance (NSEs: 0.96 to nearly 1; P-bias: 15%–18% across all subwatersheds except for subwatershed 8) for approximating SWAT in-stream phosphorus loads. Notably, the subwatersheds which perform worst are headwater subwatersheds and consist of relatively few HRUs (for example, subwatershed 8 has only three HRUs). The lesser performance in these subwatersheds is likely due to the divergence between the RM assumption and SWAT-applied method for allocating conservation practice combinations: When there are many upstream HRUs, the upstream land allocated to each conservation practice combination will, in the aggregate, consist of similar soils and slope classes despite different precise allocations. When there are few upstream HRUs, the RM estimate may not capture SWAT sensitivity to which soil type or slope class (i.e. which HRU) a practice is implemented on. Therefore, when applying the modified (or traditional) RM method at the subwatershed level, it is important to acknowledge that the method allows for targeting specific *subwatersheds* with conservation practices but not targeting specific HRUs within a subwatershed.

To evaluate the benefits of computational time saved by modified RM approach, we recorded the computation time of simulations between original SWAT and our modified RM approach with the same conservation practice allocations. Specifically, the computation for 100 randomized simulations in SWAT takes 2 h 20 min, while the modified RM method takes about 4.2 min with a single processor in an Intel Core i7-8700k 64 bit and 32 GB memory Windows PC. Thus, the computational time can be reduced by about 3500% with the modified RM approach.

3.3. Modification 3: Dealing with impacts of in-stream and reservoir processes for sediment

3.3.1. SWAT in-stream and reservoir processes for sediment

Sediment transport in SWAT depends upon flow conditions and sediment supply. Here sediment "supply" refers to sediment which has been lost from the landscape during the current model time step as well as all sediment that was previously deposited and still remains within the stream network. Note also, sediment entrainment (but not deposition) is not controlled by flow conditions and sediment supply *simultaneously,* but rather, one or the other is limiting for a given reach at a specific time, according to the following logic: (1) sediment entrainment may only occur if the stream has sufficient transport capacity (as determined by flow conditions) for increased suspended sediment concentration; (2) any previously deposited sediments which have remained within the reach are entrained first; (3) if all previously deposited sediments are exhausted, the sediment transport capacity may be met by channel bed and bank erosion, but only if the streamflow generates sufficient shear stress upon the streambed/bank.

These interactions between flow and sediment supply and between the event-scale and long-term accumulation present an obstacle for

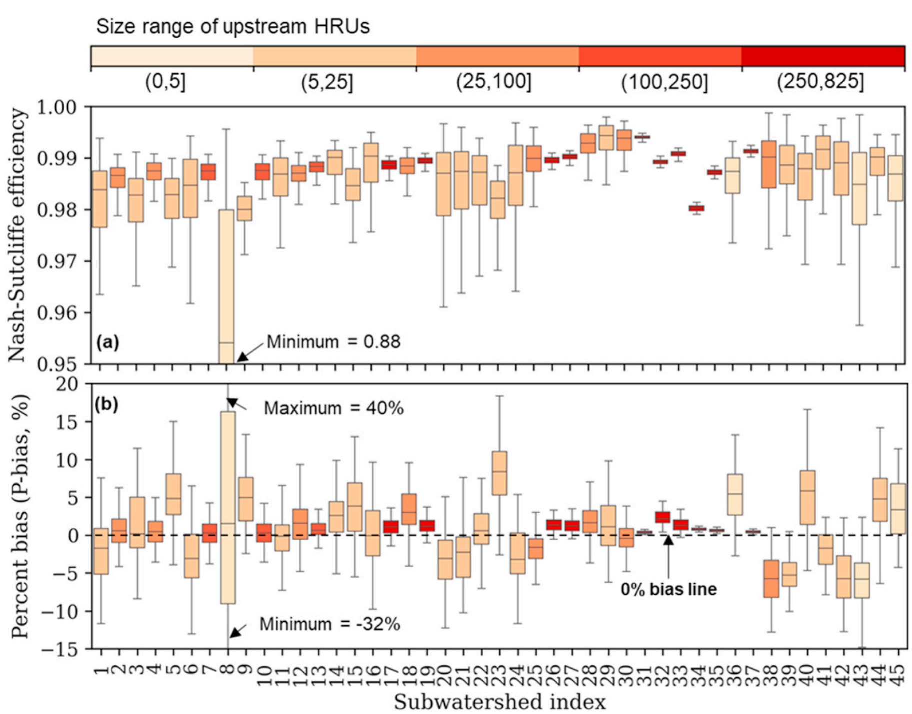

Fig. 8. Performance of modified RM on simulating the in-stream total phosphorus load across all subwatersheds. Each boxplot includes 100 random land allocations of four distinct conservation practices. NSE (a) and percent bias (b) of all subwatersheds in a testbed watershed are presented. Flow directions in the main channel: 4-7-10-13-26-27-32-31-37-35-34 (outlet). HRUs = hydrologic response units.

ascertaining a simple approximation for in-stream sediment processes. For zones of a stream network where transport capacity consistently exceeds sediment supply (we will refer to these as supply constrained), reaches would be controlled by a combination of landscape sediment loss and streambed shear stress (a function of flow). For zones of a stream network where sediment supply consistently exceeds transport capacity (we will refer to these as flow constrained), reaches would be controlled by a combination of the transport capacity (a function of flow, distinct from that for shear stress) and the sediment deposition rate.

3.3.2. Use streamflow to estimate in-stream sediment loads for surrogate model

Here we demonstrate the need to depart from an RM-based approximation for sediment and offer, instead, a simple, nonlinear approximation based on the underlying model processes in SWAT.

For flow-constrained reaches within the stream network, the most sensible linear approximation and RM formulation for sediment export are according to landscape water yield. That is:

$$\boldsymbol{S} = f_4(Q) \approx \widehat{k}_S \cdot \boldsymbol{Q} = \widehat{k}_{S,1} \cdot \sum_{m \in M} \left(diag(\boldsymbol{Y}_{m,t})_Q \cdot diag(\boldsymbol{A}) \cdot \boldsymbol{F}_m \right) + \widehat{k}_{S,2} \quad \forall t \in T \tag{15}$$

where $\widehat{k}_{S,1}$ [mg/L] and $\widehat{k}_{S,2}$ [mg/L] are coefficients representing effects at typical conditions. However, the transport capacity can be a highly nonlinear function of flow; therefore, this approximation may be unlikely to hold. For example, we select the SWAT modeling option to use the Simplified Bagnold model (Neitsch et al., 2011) for transport capacity (one of four options), where:

$$conc_{sed,ch,mx} = C_{sp} v_{ch,pk}{}^{spexp} \tag{16}$$

where $conc_{sed,ch,mx}$ is the maximum sediment concentration (ton/m^3 or kg/L), C_{sp} and *spexp* are parameters defined by SWAT modeler, and $v_{ch,pk}$ is the peak channel velocity (m/s) during the time step – a nonlinear function of inflow (see Eq. (7):2.2.3 from SWAT theory documentation, 2009). In the case where non-linearities must be incorporated, and following the Bagnold equation, a more representative approximation might be:

$$\boldsymbol{S}_n = f_4(Q) \approx \widehat{k}_{S,lin} \cdot \boldsymbol{Q}_n^{\widehat{k}_{S,exp}} \quad \forall n \in N\ \forall t \in T \tag{17a}$$

or perhaps, considering the complex relationship between flow volume ($\boldsymbol{Q}$ here) and peak flow velocity ($v_{ch,pk}$ in the Bagnold equation), even a polynomial approximation is suitable:

$$\boldsymbol{S}_n = f_4(Q) \approx \widehat{k}_{S,1} + \widehat{k}_{S,2} \cdot \boldsymbol{Q}_n + \widehat{k}_{S,3} \cdot \boldsymbol{Q}_n^2 \quad \forall n \in N\ \forall t \in T \tag{17b}$$

where $\widehat{k}_{S,lin}$, $\widehat{k}_{S,exp}$, $\widehat{k}_{S,1}$, $\widehat{k}_{S,2}$, and $\widehat{k}_{S,3}$ are parameters for the sediment-flow relationship at typical conditions. However, this nonlinear approximation is not compatible with a response matrix formulation and would possibly require the modeler to adjust their use of the surrogate – for instance, changing the solution method used for optimization.

On the other hand, for supply-constrained reaches, the most sensible linear approximation and RM formulation likely must account for landscape sediment yield and water yield. That is,

$$\boldsymbol{S} = \boldsymbol{S} + f_5(\boldsymbol{Q}) \approx \boldsymbol{S} + k_{S,Q} \cdot \boldsymbol{Q} = \sum_{m \in M} \left(diag(\boldsymbol{Y}_{m,t})_S \cdot diag(\boldsymbol{A}) \cdot \boldsymbol{F}_m \right) + \widehat{k}_{S,Q} \cdot \sum_{m \in M} \left(diag(\boldsymbol{Y}_{m,t})_Q \cdot diag(\boldsymbol{A}) \cdot \boldsymbol{F}_m \right) \forall t \in T \tag{18}$$

where $\widehat{k}_{S,Q}$ [mg/L] is the streambed sediment contribution per unit of flow at typical conditions. However, the nonlinearity of streambed erosion processes may make this approximation unlikely to hold as well.

Fig. 9(a and b) compares the estimates of best-performing approximations with estimates of the traditional RM approach for selected, illustrative subwatersheds. For flow-constrained subwatersheds such as shown in Figs. 9a and 10b (best approximations from Equations (15) and (17), respectively), the traditional RM approach significantly overestimates the sediment loads, while the proposed modifications provide highly accurate estimates (R-squared > 0.9). For example, in subwatersheds 33 and 34 (the outlet), the NSE for traditional RM sediment estimates are 0.08 and −1.69, respectively, but improve to 0.91 and 0.97 under a flow-based approximation. Evidently, SWAT simulates significant sediment deposition in these zones of the stream network, effectively buffering the upstream landscape sediment loss signal (Jerolmack and Paola, 2010; Romans et al., 2016). In some cases (e.g. as shown in Fig. 9c) the flow-sediment relationship can also be approximately linear, hence the acceptable performance of the response matrix approach

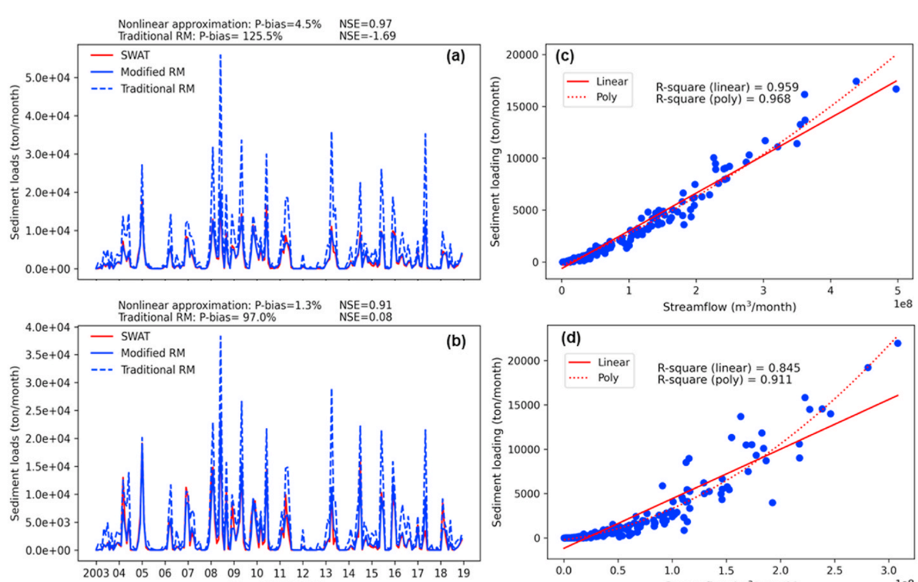

Fig. 9. Performance of nonlinear approximation and traditional RM on simulating the in-stream load of sediment: subwatershed 34 (a) (outlet) and subwatershed 33 (b). Flow-sediment relationships approximated by linear and polynomial regressions for subwatershed 34 (c) and subwatershed 33 (d) are presented for discussion.

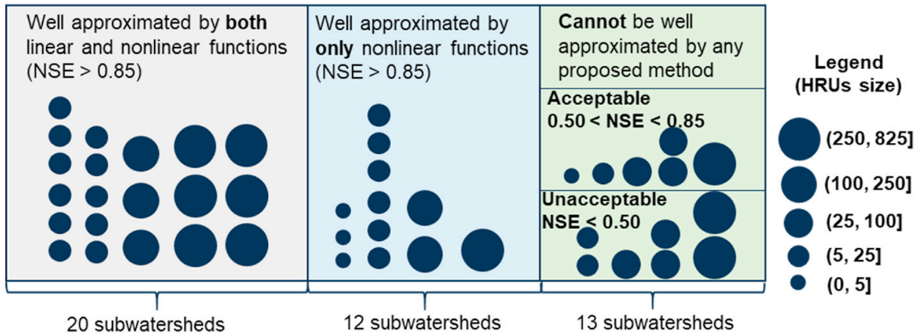

Fig. 10. Performance of linear and nonlinear functions approximating in-stream sediment load. Each subwatershed is represented as a bubble with its size indicating the number of HRUs in its own subwatershed and upstream. HRUs = hydrologic response units; NSE = Nash-Sutcliffe efficiency.

presented in Equation (15). However, in many cases (e.g. as shown in Fig. 9d) the flow-sediment relationship is clearly exponential or polynomial, and a nonlinear approximation is required in place of the RM approach.

We apply the potential RM formulations presented in Equations (15) and (17) (a or b), and 18 and compare their performance in estimating sediment load at the monthly scale at all subwatershed outlets, over 100 realizations for conservation practice allocations. We find that, of the 45 USRW subwatersheds, 20 subwatersheds are approximated well (i.e. average NSE >0.85) by either the linear (Equation (15)) or nonlinear (Equations (17a) and (17b)) approximations aligned with the Bagnold equation (see Fig. 10). For those 20 subwatersheds, the linear approximation method is suggested as it may be incorporated into an RM approach. 12 subwatersheds are only approximated well by the nonlinear estimates (polynomial or power functions). These results align with the suggestions of previous studies that, due to historical management practices, sediment transport in the Upper Mississippi River Basin is generally flow constrained (Neal and Anders, 2015; Trimble, 1999).

Of the other 13 subwatersheds, 6 can be approximated acceptably (i.e. NSE >0.5), but not well, by one of the flow-constrained approximations. The remaining 7 subwatersheds cannot be approximated acceptably by any of the formulations offered here (i.e NSE <0.5). The poor performance in sediment estimation seems to go beyond the unaccounted impacts of soil type and slope class discussed in Section 3.3.2, as there is no clear trend in performance as upstream HRUs increase. It appears that some subwatersheds either frequently switch modes between possible dominant controls on SWAT sediment export, not conforming to a single approximation, or are generally less amenable to approximations for SWAT sediment simulations.

4. Conclusions

In this work, we show that (1) an accurate RM-based approach to SWAT approximation requires that response matrices be generated for all distinct *combinations* of conservation practices, in order to account for the interaction effects between individual practices; (2) a modified RM method, especially accounting for in-stream and reservoir processes, is required to correct estimates for phosphorus export; and (3) a departure from RM-based approximation is required for accurately estimating sediment, instead utilizing a nonlinear flow-based estimate for sediment loads. We hypothesize, based on primary process model understanding, that the modifications presented for phosphorus could also adequately correct estimates for flow and nitrogen when necessary. We also highlight that, when applied for decisions at the subwatershed scale rather than the HRU scale, the proposed approximations perform best at outlets draining large areas. The approximations detailed here provide efficient spatial and dynamic simulations on hydrological responses based on a wide range of spatial applications of agricultural conservation practices. Excluding sediment, the approximations maintain an RM-based approach, facilitating advantages such as the feasibility of linear programming methods. The approximations are especially well-suited to integration within a system of systems modeling framework where modelers wish to consider a mixture of non-point and point source models (Li et al., 2021a). For example, the nutrient effluents simulated by wastewater treatment models can be directly added via the RM method to accurately simulate the in-stream P load. The application here is centered on surrogating the SWAT model, but the process, enlightened by the discussion provided here, could plausibly be generalized to other semi-distributed hydrologic models (e.g., WASP or SWMM). The revised RM enables a more accurate use of a watershed hydrological model for finding optimal watershed management solutions via 1) classic optimization, i.e. linear programming and nonlinear programming (including the suggested nonlinear equation describing the instream and reservoir sediment processes; 2) heuristic optimization such as genetic algorithm which uses the revised RM to replace the original simulation model.

Data and code availability

The data and codes used for constructing the modified RM are available via GitHub (https://github.com/shaobinli/modified_RM).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was funded by the US National Science Foundation (INFEWS/T1 award number 1739788).

References

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

Babbar-Sebens, M., Mukhopadhyay, S., Singh, V.B., Piemonti, A.D., 2015. A web-based software tool for participatory optimization of conservation practices in watersheds. Environ. Model. Software 69, 111–127. https://doi.org/10.1016/j.envsoft.2015.03.011.

Boreux, V., Kushalappa, C.G., Vaast, P., Ghazoul, J., 2013. Interactive effects among ecosystem services and management practices on crop production: pollination in coffee agroforestry systems. Proc. Natl. Acad. Sci. U. S. A. 110, 8387–8392. https://doi.org/10.1073/pnas.1210590110.

Cai, X., Zeng, R., Kang, W.H., Song, J., Valocchi, A.J., 2015. Strategic planning for drought mitigation under climate change. J. Water Resour. Plann. Manag. 141, 1–10. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000510.

Chaubey, I., Chiang, L., Gitau, M.W., Mohamed, S., 2010. Effectiveness of best management practices in improving water quality in a pasture-dominated watershed. J. Soil Water Conserv. 65, 424–437. https://doi.org/10.2489/jswc.65.6.424.

Daniel, E.B., Camp, J.V., Leboeuf, E.J., Penrod, J.R., Dobbins, J.P., Abkowitz, M.D., 2011. Watershed modeling and its applications: a state-of-the-art review. Open Hydrol. J. 5, 26–50.

Femeena, P.V., Sudheer, K.P., Cibin, R., Chaubey, I., 2018. Spatial optimization of cropping pattern for sustainable food and biofuel production with minimal downstream pollution. J. Environ. Manag. 212, 198–209. https://doi.org/10.1016/J.JENVMAN.2018.01.060.

Fitzpatrick, W.P., Bogner, W.C., Bhowmik, N.G., 1987. Sedimentation and Hydrologic Processes in Lake Decatur and its Watershed, Report of Investigation.

Gorelick, D.E., Baskaran, L.M., Jager, H.I., 2019. Visualizing feedstock siting in biomass production: tradeoffs between economic and water quality objectives. Land Use Pol. 88, 104201 https://doi.org/10.1016/J.LANDUSEPOL.2019.104201.

Gorelick, S.M., 1983. A review of distributed parameter groundwater management modeling methods. Water Resour. Res. 19, 305–319. https://doi.org/10.1029/WR019i002p00305.

Gramig, B.M., Reeling, C.J., Cibin, R., Chaubey, I., 2013. Environmental and economic trade-offs in a watershed when using corn stover for bioenergy. Environ. Sci. Technol. 47, 1784–1791. https://doi.org/10.1021/es303459h.

Harrison, S., McAree, C., Mulville, W., Sullivan, T., 2019. The problem of agricultural 'diffuse' pollution: getting to the point. Sci. Total Environ. 677, 700–717. https://doi.org/10.1016/j.scitotenv.2019.04.169.

Housh, M., Cai, X., Ling Ng, T., McIsaac, G.F., Ouyang, Y., Khanna, M., Sivapalan, M., Jain, A.K., Eckhoff, S., Gasteyer, S., Al-Qadi, I., Bai, Y., Yaeger, M.A., Ma, S., Song, Y., 2014. System of systems model for analysis of biofuel development. J. Infrastruct. Syst. 21, 04014050 https://doi.org/10.1061/(ASCE)IS.1943-555X.0000238.

Housh, M., Yaeger, M.A., Cai, X., Mcisaac, G.F., Khanna, M., Sivapalan, M., Ouyang, Y., Al-Qadi, I., Jain, A.K., 2015. Managing multiple mandates: a system of systems model to analyze strategies for producing cellulosic ethanol and reducing Riverine nitrate loads in the upper Mississippi River Basin. Environ. Sci. Technol. 49, 11932–11940. https://doi.org/10.1021/acs.est.5b02712.

Illinois Environmental Protection Agency, Illinois Department of Agriculture, University of Illinois Extension, 2015. Illinois Nutrient Loss Reduction Strategy.

Jarvie, H.P., Johnson, L.T., Sharpley, A.N., Smith, D.R., Baker, D.B., Bruulsema, T.W., Confesor, R., 2017. Increased soluble phosphorus loads to lake Erie: unintended consequences of conservation practices? J. Environ. Qual. 46, 123–132. https://doi.org/10.2134/jeq2016.07.0248.

Jarvie, H.P., Neal, C., Withers, P.J.A., 2006. Sewage-effluent phosphorus: a greater risk to river eutrophication than agricultural phosphorus? Sci. Total Environ. 360, 246–253. https://doi.org/10.1016/j.scitotenv.2005.08.038.

Jerolmack, D.J., Paola, C., 2010. Shredding of environmental signals by sediment transport. Geophys. Res. Lett. 37, 1–5. https://doi.org/10.1029/2010GL044638.

Li, S., Cai, X., Emamineja, S.A., Juneja, A., Niroula, S., Oh, S., Wallington, K., Cusick, R. D., Gramig, B.M., John, S., McIsaacd, G.F., Singh, V., 2021a. Developing an integrated technology-environment-economics model to simulate food-energy-water systems in corn belt watersheds. Environ. Model. Software 143, 105083. https://doi.org/10.1016/j.envsoft.2021.105083.

Li, S., Emaminejad, S.A., Aguiar, S.E., Furneaux, A.G., Cai, X., Cusick, R., 2021b. Evaluating long-term treatment performance and cost of nutrient removal at water resource recovery facilities under stochastic influent characteristics using artificial neural networks as surrogates for plant-wide modeling. ACS ES&T Eng. 1 (11), 1517–1529. https://doi.org/10.1021/acsestengg.1c00179.

Maddock, T., 1972. Algebraic technological function from a simulation model. Water Resour. Res. 8, 129–134. https://doi.org/10.1029/WR008i001p00129.

Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. ASABE (Am. Soc. Agric. Biol. Eng.) 50, 885–900. https://doi.org/10.13031/2013.23153.

Neal, C.W.M., Anders, A.M., 2015. Suspended sediment supply dominated by bank erosion in a low-gradient agricultural watershed, Wildcat Slough, Fisher, Illinois, United States. J. Soil Water Conserv. 70, 145–155. https://doi.org/10.2489/jswc.70.3.145.

Neitsch, S.L., Arnold, J.G., Kiniry, J.R., Williams, J.R., 2011. Soil & Water Assessment Tool Theoretical Documentation Version 2009 (College Station, Texas).

Peterson, S.M., Flynn, A.T., Traylor, J.P., 2016. Groundwater-Flow Model of the Northern High Plains Aquifer in Colorado, Kansas, Nebraska, South Dakota, and Wyoming: U.S. Geological Survey Scientific Investigations Report, pp. 2016–5153. https://doi.org/10.3133/sir20165153.

Razavi, S., Tolson, B.A., Burn, D.H., 2012. Review of surrogate modeling in water resources. Water Resour. Res. 48 https://doi.org/10.1029/2011WR011527.

Romans, B.W., Castelltort, S., Covault, J.A., Fildani, A., Walsh, J.P., 2016. Environmental signal propagation in sedimentary systems across timescales. Earth Sci. Rev. 153, 7–29. https://doi.org/10.1016/j.earscirev.2015.07.012.

Seo, S.B., Mahinthakumar, G., Sankarasubramanian, A., Kumar, M., 2018. Conjunctive management of surface water and groundwater resources under drought conditions using a fully coupled hydrological model. J. Water Resour. Plann. Manag. 144, 04018060 https://doi.org/10.1061/(asce)wr.1943-5452.0000978.

Shafiee-Jood, M., Housh, M., Cai, X., 2018. Hierarchical decision-modeling framework to meet environmental objectives in biofuel development. J. Water Resour. Plann. Manag. 144, 04018030 https://doi.org/10.1061/(ASCE)WR.1943-5452.0000950.

Sinshaw, T., Yuan, L., Forshay, K., 2019. A Review of Watershed and Water Quality Tools for Nutrient Fate and Transport. Washington, DC, EPA/600/R-19/232.

Trimble, S.W., 1999. Decreased rates of alluvial sediment storage in the Coon Creek Basin, Wisconsin, 1975-93. Science (80-.) 285, 1244–1246. https://doi.org/10.1126/science.285.5431.1244.

USDA NASS, 2019. 2017 Census of Agriculture Illinois State and County Data.

Yoon, J., Klassert, C., Selby, P., Lachaut, T., Knox, S., Avisse, N., Harou, J., Tilmant, A., Klauer, B., Mustafa, D., Sigel, K., Talozi, S., Gawel, E., Medellín-Azuara, J., Bataineh, B., Zhang, H., Gorelick, S.M., 2021. A coupled human-natural system analysis of freshwater security under climate and population change. Proc. Natl. Acad. Sci. U. S. A. 118 https://doi.org/10.1073/pnas.2020431118.

Zhang, X., Srinivasan, R., Van Liew, M., 2009. Approximating SWAT model using artificial neural network and support vector machine. JAWRA J. Am. Water Resour. Assoc. 45, 460–474. https://doi.org/10.1111/j.1752-1688.2009.00302.x.