Environmental Modelling and Software 143 (2021) 105083

Contents lists available at ScienceDirect

Environmental Modelling and Software

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

Developing an integrated technology-environment-economics model to simulate food-energy-water systems in Corn Belt watersheds

Shaobin Li [a], Ximing Cai [a,*], Seyed Aryan Emaminejad [a], Ankita Juneja [b], Sundar Niroula [a], Seojeong Oh [c], Kevin Wallington [a], Roland D. Cusick [a], Benjamin M. Gramig [c], Stephen John [d], Gregory F. McIsaac [d,e], Vijay Singh [b]

[a] Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA
[b] Department of Agricultural and Biological Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA
[c] Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, Urbana, IL, USA
[d] Agricultural Watershed Institute, Decatur, IL, USA
[e] Department of Natural Resources and Environmental Sciences, University of Illinois at Urbana-Champaign, Urbana, IL, USA

ARTICLE INFO

Keywords:
Integrated modeling
Surrogate modeling
Machine learning
Nutrient recycling
Phosphorus recovery
Environmental benefits

ABSTRACT

To facilitate understanding and decision making in the food-energy-water (FEW) nexus context, we develop an integrated technology-environment-economics model (ITEEM) at a watershed scale. ITEEM is built as an integration of various models, including models for grain processing, drinking water treatment, and wastewater treatment (technology); a watershed model for hydrology, water quality, crop production, and nutrient cycling (environment); an economics model assessing total benefit, including non-market valuation of environmental benefits. Different data techniques are applied to develop suitable surrogates for computer-based models, including a response matrix method, artificial neural networks, and lookup tables. Empirical equations are applied to develop models of economics and drinking water treatment. The input-output relationships between the models are formulated in a unified computational framework. ITEEM, a spatially semi-distributed dynamic simulation model, can be used to quantify the environmental and socioeconomic impacts of various management practices, technologies, and policy interventions on FEW systems in the Corn Belt.

1. Introduction

Food-energy-water (FEW) systems in the US Corn Belt are highly interconnected and sensitive to stresses and threats. Grain production and subsequent utilization for animal feed, human food, and ethanol production have pervasive effects on water quantity and quality in downstream environments both locally (e.g., lakes and rivers with elevated nitrogen and phosphorus) and nationally (e.g., Hypoxic zone in the Gulf of Mexico) (US EPA, 2017). Water stress associated with increased climatic variability is anticipated to increase (Muttiah and Wurbs, 2002), especially in many mid-sized cities in the Corn Belt that interact with neighboring agricultural lands, major industrial needs (Li et al., 2018), and their shared watersheds. Energy demand and overall costs for wastewater and drinking water treatment have increased, and this trend is expected to be exacerbated by continued expansion of food and bioethanol production (Simpson et al., 2008; Twomey et al., 2010). To deal with these threats to and risks within FEW systems, long-term efforts have been made to resolve the conflicts between agriculture, food industry, water supply, and environmental protection. For example, wastewater treatment and corn ethanol refinery facilities have begun extracting nutrients from "waste" and process byproducts, which results in both the reuse of extracted materials as inorganic mineral fertilizers (e.g. struvite and calcium phytate) and the reduction of point-source discharge of nutrients to the environment. For example, recovering phosphorus (P) can conserve a finite resource (e.g. phosphate rock) (Cordell et al., 2009; Juneja et al., 2019; Margenot et al., 2019); cost-effective water treatment technologies are adopted to conserve energy use (Bhatnagar and Sillanpää, 2011); agricultural best management practices (BMPs) reduce nutrient and soil loss from farmland in upstream watersheds (Lemke et al., 2011; Rao et al., 2009). Researchers have called for holistic integrated modeling development and assessment for FEW systems at various scales to avoid fragmented status quo

* Corresponding author. Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, 3216 Digital Computer Lab MC 250, Urbana, IL, 61801, USA.

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

https://doi.org/10.1016/j.envsoft.2021.105083
Accepted 11 May 2021
Available online 23 May 2021
1364-8152/© 2021 Elsevier Ltd. All rights reserved.

decision making (Leck et al., 2015; Little et al., 2019). This paper presents an integrated technology-environment-economics model (ITEEM) which unites a set of surrogates and empirical models derived from the various primary models simulating key processes at a watershed scale. The developed ITEEM is capable of analyzing complex systems and specific solutions to interconnected problems in FEW systems in Corn Belt watersheds.

There are several major challenges when integrating models from different disciplines. First, most physical models are developed using discipline-specific computer programs or software packages (e.g., SWAT for hydrologic processes, GPS-X for wastewater treatment), which causes a barrier for automatic information transfer. Recently, some interfaces have been developed for simple automated data exchange between two models (Anderson et al., 2018; Xiang et al., 2020). For a large interdisciplinary integrated model involving agricultural, hydrologic, and engineering components developed in various computer programs (including commercial software), as the case of our study, the level of complexity can be overwhelming to modelers, and it usually turns out to be infeasible to directly integrate different models due to incompatibilities among discipline-specific computer programs (Little et al., 2019). Second, some engineering design models (e.g., GPS-X for wastewater treatment, SuperPro Designer for Grain processing) are proprietary which may impose costs and legal constraints on direct coupling. Third, inputs and outputs from separate models are likely to have different temporal and spatial scales with distinct data formats, which need to be harmonized at the points of interaction between models (Cai, 2008). Appropriately building the interactions between various models is a key step to enable information transfer endogenously within a consistent model. Fourth, complex physical models can be highly computationally expensive; an affordable computational burden is especially important if the research of interest will address stochastic problems (Little et al., 2019). Thus directly integrating many computationally heavy models is often computationally infeasible.

Researchers have developed various integrated models (Cai, 2008; Carmichael et al., 2004; Gaddis et al., 2010; Housh et al., 2014). Cai (2008) shared reflective comments on the advantages and challenges of holistic modeling (tight coupling of different components in one consistent model) versus compound modeling approaches ("loose" coupling of different components via external data exchanges). Holistic models embed different components into a single consistent optimization model, such as hydrologic-economic models (Cai, 2008; Cai et al., 2003; Harou et al., 2009), hydro-biogeochemical model (Wu et al., 2016), and "system of systems" models, e.g. a biofuel (biomass and refinery)-infrastructure-environment model (Housh et al., 2014). Holistic optimization models are usually composed of mathematical equations including the objective function(s) and constraint function(s). Other system modeling approaches applied in FEW systems include agent-based models (Ng et al., 2011), life cycle assessment (Li et al., 2020), system dynamics (Feng et al., 2016; Gaddis and Voinov, 2010), etc. However, these system modeling methods have less focus on integrating detailed physical process modeling, but more focus on other perspectives. For example, agent-based models focus on simulating the behavior and decision-making of multiple stakeholders, life cycle assessment focuses on quantifying environmental impacts from cradle to grave, and system dynamics focuses on modeling the feedbacks among stock variables and drivers. The degree of process details at which those system modeling approaches have may not lend themselves to coupling multiple complex process models in a system of systems.

Little et al. (2019) proposed a generic tiered system of systems (GTSoS) to upscale physical models from the process level to the system level via integration while keeping computational tractability and minimizing the loss of fidelity (Little et al., 2019). Models that are developed at the process level in various computer programs (or software packages) with domain-specific knowledge and data can be replaced by surrogates (also termed reduced-order models, meta-models, or emulators), if process models cannot be integrated directly due to complexity and incompatibilities among discipline-specific computer programs. Various data techniques can be applied for emulating a process model, such as polynomial response surfaces, artificial neural networks, and supporting vector machine using numerical samples of inputs and outputs of the primary model under a systematic sampling strategy (Leperi et al., 2019; Lu and Ricciuto, 2019). Those surrogates typically build statistical relationships between inputs and outputs of a system modeled by a primary model. Another type of surrogate is based on hybrid theory and data (also termed as lower-fidelity physically-based surrogates) (Razavi et al., 2012). By replacing complex process models with appropriate surrogates, one can integrate them into a consistent model, maintaining reasonable fidelity of the primary process models without causing a serious computational burden, as most surrogates do not have a rich internal structure (Carmichael et al., 2004).

Although the GTSoS framework provides a promising direction on model integration for analyzing a system of systems, the development of such a framework is challenging. Specific challenges include the selection of an appropriate mathematical form of a surrogate for a particular process model, and the integration of the surrogates across multiple spatial and temporal scales (Cai, 2008). In addition, examples of real-world problems are needed to demonstrate the effectiveness and applicability of GTSoS to the various complex system modeling cases. Here, we explain the methodology used to overcome these challenges in the construction and execution of ITEEM. Disciplinary-specific process models are replaced by surrogates, and these surrogates are integrated within a unified computational software framework to form a holistic model. ITEEM is demonstrated in a watershed in the Corn Belt to analyze inter-connected problems of crop production, grain processing, water and wastewater treatment, and nutrient management, with consideration of technologies, management practices, and policies for multiple sectors.

2. Research problem and FEW systems characterization in Corn Belt watersheds

2.1. Research problem

FEW systems are usually highly interconnected crossing multiple sectors.in many regions. For the Corn Belt watersheds, FEW systems are sensitive to stresses and threats with respect to food production, and increasing biomass production, and energy supply and demand, which pose impacts on water quality, water supply, energy demand and cost, resources conservation, and economic growth and financial stability. These interconnected components of FEW systems are depicted in Fig. 1. Managing phosphorus (P) within these systems has proven especially challenging over the last 40 years due to the so-called "phosphorus paradox". On the one hand, phosphorus is an essential nutrient for plant growth, and correspondingly, copious applications of phosphorus fertilizer have been critical to meeting demand for food, livestock feed, and biofuel (Jarvie et al., 2015). However, on the other hand, phosphorus fertilizer applied in agricultural fields is at risk of being transported into water bodies where, in excess, it contributes to water quality degradation, namely, toxic algae blooms (Bennett et al., 2001; Carpenter, 2008). Efforts to navigate these conflicting objectives have been undermined by long-lasting stores of P in fields and streams (i.e. P legacy) which create time-lags between changed agricultural practices and their impact on water quality or crop yields (Jarvie et al., 2017; Powers et al., 2016; Sharpley et al., 2013). Developing technologies for P removal and recovery from waste streams with feasible costs are important given that these interventions play unique and under-emphasized roles within the FEW nexus regarding water pollution, resource recovery, and agriculture production, as shown in Fig. 1.

Traditionally, various management options are evaluated and implemented within individual systems (F, E, or W) or processes at local scales. Nutrient pollution in a Corn Belt watershed typically comes from

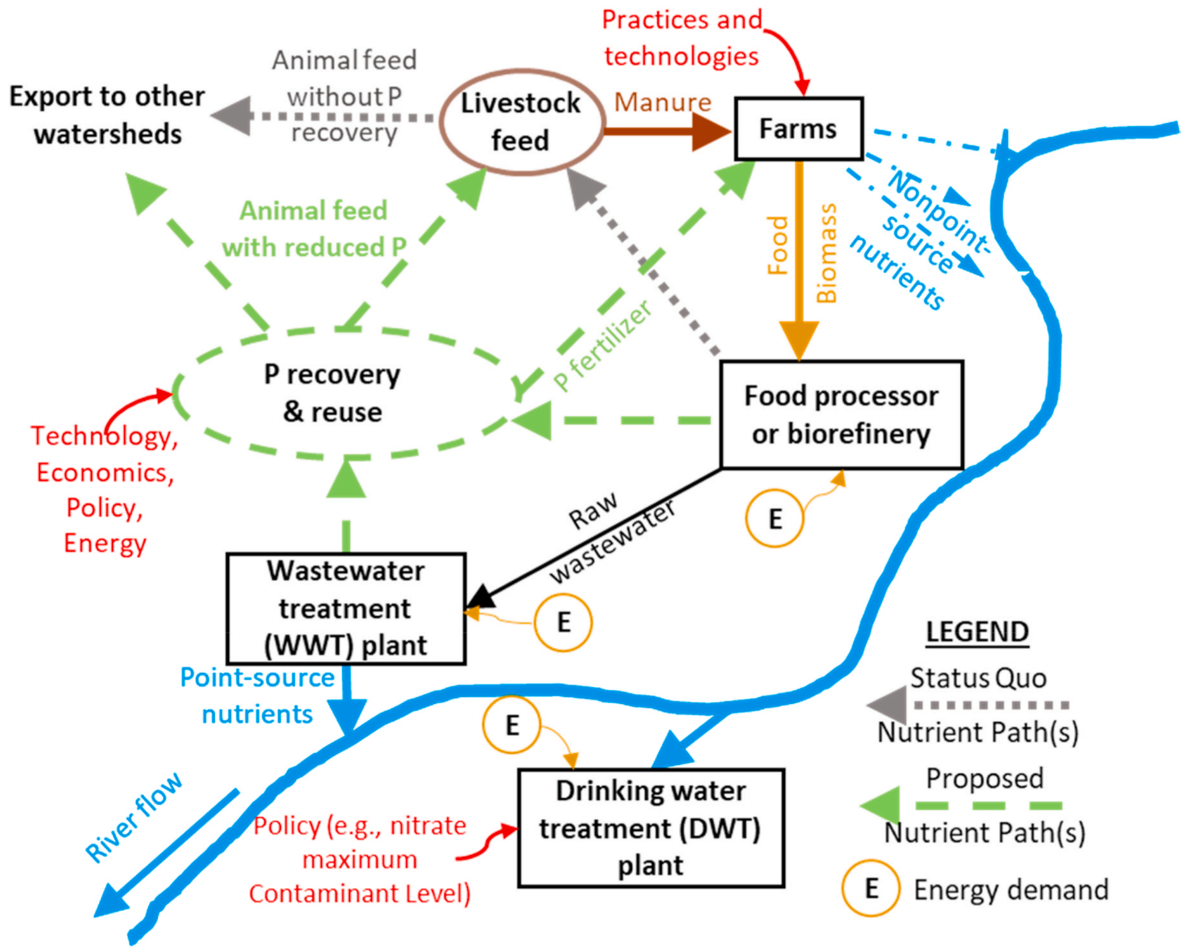


Fig. 1. FEW systems in a Corn-Belt watershed with phosphorus (P) recovery as a key technology.

a combination of identifiable pollution discharges (i.e. point sources, such as wastewater treatment plants) and diffuse pollution (i.e. non-point sources, such as agricultural runoff). Grain processing also indirectly contributes to nonpoint source P pollution by concentrating P in coproducts (corn gluten feed, CGF, and dried distillers grains with solubles, DDGS) to the extent that they exceed livestock dietary requirements, leading to pollution by P-enriched manure in livestock feedlots (Nahm, 2002). Studies have shown the P concentration in CGF and DDGS (two commonly used ingredients for cattle and poultry diets) can be reduced by recovering P from light steepwater (wet milling plant) and thin stillage (dry grind plant), for potential use as a fertilizer (Juneja et al, 2019, 2020). Drinking water treatment is considered as a local-scale process that takes raw water from lakes that could have upstream point and nonpoint sources of nitrate and sediment.

Traditional approaches usually use separate disciplinary-specific models and ignore or do not fully consider the impact of the FEW nexus relations that exist at certain spatial scales. Such approaches cannot capture the interconnected influence of measures taken across the interdependent systems. To address this general deficit in the Corn Belt and other regions, a seamless integrated technology-environment-economics model (ITEEM) is developed to assess the tradeoffs and synergies within FEW systems in the Corn Belt.

2.2. Primary models for different components of FEW systems

Components of FEW systems shown in Fig. 1 are modeled by various computer-based programs and empirical relationships on data and knowledge in individual disciplinary domains. Specifically, the Soil and Water Assessment Tool (SWAT), a semi-distributed and physically-based watershed management model (Jayakrishnan et al., 2005), is used to simulate water quality, quantity, and crop yield based on different land uses and BMPs in a Corn Belt watershed. The wastewater treatment (WWT) is modeled in GPS-X (Hydromantis Environmental Software Solutions, Inc.) with advanced mathematical modeling, optimization, and management of wastewater treatment processes. Grain processing (GP) is modeled in Superpro Designer (Intelligen Inc.) to evaluate the potential of P recovery from corn coproducts with existing physical-chemical and enzymatic technology. SWAT, WWT, and GP models involve detailed processes characterized by biological, chemical, and physical principles. The drinking water treatment (DWT) model is empirical and driven by historical plant data for energy requirement and cost according to influent nitrate and sediment concentrations. In addition to the physical component of FEW systems, we also develop an economic model that represents the human dimension of the FEW systems. The economic model is semi-theoretical and empirical, based on choice experiments evaluating the relationships between water quality improvements and farmers' or the public willingness to pay that is assessed using survey data. Details on the development of each primary model are provided in the Supplementary Information (SI), Section 1.

3. Development of ITEEM

Via multi-disciplinary teamwork, three process models (SWAT, WWT, and GP) and empirical (DWT) or theoretical-empirical models (Economics) are first established at the process level (the lower part of Fig. 2). Then the components of ITEEM are developed in the form of surrogates or empirical relationships, which are coupled by integrating input and output relationships crossing temporal and spatial scales at

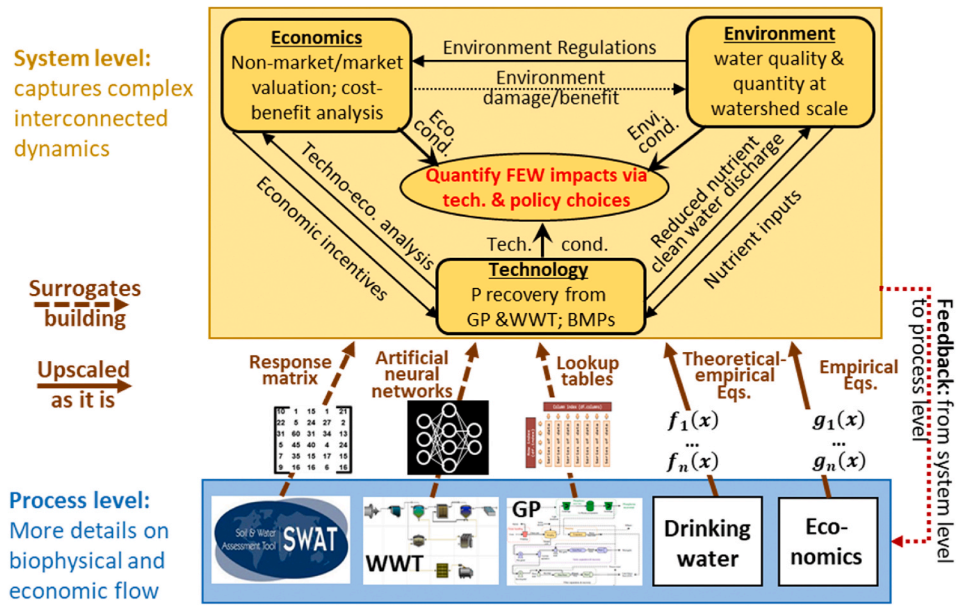


Fig. 2. A tiered modeling framework for ITEEM. The modeling framework starts from establishing disciplinary-specific primary models at the process level and upscales the primary models to the system level, which captures complex interactions between technological, environmental, and economics components. Note: SWAT = Soil and Water Assessment Tool; GP = Grain processing; WWT = Wastewater treatment.

the interaction points between the components at the system level (the upper part of Fig. 2). Such a hierarchical structure allows modelers to drill down to the process level and access details for better interpreting results simulated at the system level. All components of the ITEEM are coded in the same programming platform, Python.

The interaction between the technology (T), environment (En), and economics (Ec) at the system level of ITEEM are shown in the upper part of Fig. 2. The technology is composed of BMPs simulated in SWAT and engineering technologies simulated in WWT and GP components. The relationships between T-Ec include capital, operation, and maintenance costs for P extraction and the cost savings from a) changing farm management practices to use P recovered from biorefineries as crop fertilizer; b) introducing cover crops with no-till, etc. The relationships between En–Ec include a) non-market benefit as a measure of the value to the population of an improvement in water quality and a measure of people's preferences for alternative ways of achieving lower P pollution levels; b) water treatment and water supply cost due to extra nutrient discharge. The relationships between T-En include a) P removal from grain coproducts and hence "downstream" reductions of P in manure and feedlot runoff; b) nonpoint source change of P, nitrate, and sediment loads in rivers and to lakes; c) point-source nutrient discharge reduction; d) mined P offset with P recovered from biorefineries and wastewater treatment plant.

3.1. Selection of model forms for components of ITEEM

As discussed before, it is challenging to select appropriate model forms for different model components (Razavi et al., 2012). As there are no well-defined standards for selecting model forms, we select an appropriate model form for each component based on the model availability (i.e., if a process model is available for a specific component), complexity and attributes of existing models, and data availability. A decision tree used for selecting model forms is presented in Fig. 3. In general, we start by examining if a component (either physical or economic) can be represented by a set of empirical equations for the purpose of our study, especially for those components for which primary computer-based models are not available or do not have external sources to develop a computer-based model. In the current study, the economics and drinking water treatment (DWT) components fall within this category. Empirical equations are directly modularized in Python to represent such components.

For the components for which computer-based models (i.e., SWAT, WWT, and GP) exist, distinct surrogate forms are chosen according to the particularities of each case. First, if there are sufficient simulation data determined by the number of inputs and outputs of a computer-based model, a machine learning model (e.g., artificial neural network, ANN) can be used to surrogate a complex and nonlinear process model, such as the case for the WWT component of ITEEM. Note that the sample size of the simulation data varies case by case and is dependent on model complexity (Davis et al., 2018); sufficient simulation data support the development of a surrogate model with desirable and stable performance. Second, if a computer-based model has a large number of inputs and simulate copious spatial and temporal outputs, it becomes more challenging to obtain sufficient simulation data for training traditional surrogate models (e.g., ANNs, SVM). The challenge arises from two perspectives: 1) the process of generating sufficient SWAT simulations for such a spatially-distributed dynamic model is computationally expensive itself; 2) the process of training ANN and SVM with such high dimension inputs and outputs takes up computer memory and numerous calculations, thus are prone to crash. Although machine learning has also been applied to approximate complex hydrological models (Cai et al., 2015; Zhang et al., 2009), the number of inputs and outputs in their SWAT are usually relatively low (only several or dozens of inputs and outputs at most), thus requiring fewer simulation data (only thousands of simulations at most). For our case, we aim at developing a surrogate model that can reasonably replicate SWAT simulations, including temporal and spatial heterogeneity, while varying a large number of inputs (e.g., BMP applications at each of the subwatersheds). Since neither ANN nor SVM with limited model simulations (e.g., less than 100 model runs) could be effective in generating such a spatially-distributed and dynamic surrogate for copious inputs and outputs, we choose a response matrix (RM) method as a surrogate model to produce spatially-distributed dynamic outputs of SWAT, which we show is appropriate in this study. The RM method estimates water, sediment, and nutrient yields from landscapes with partial adoption of management practices by interpolating between simulation results when those practices are applied to all or none of the landscape. The loading of nitrate, total P, sediment, and streamflow in each channel

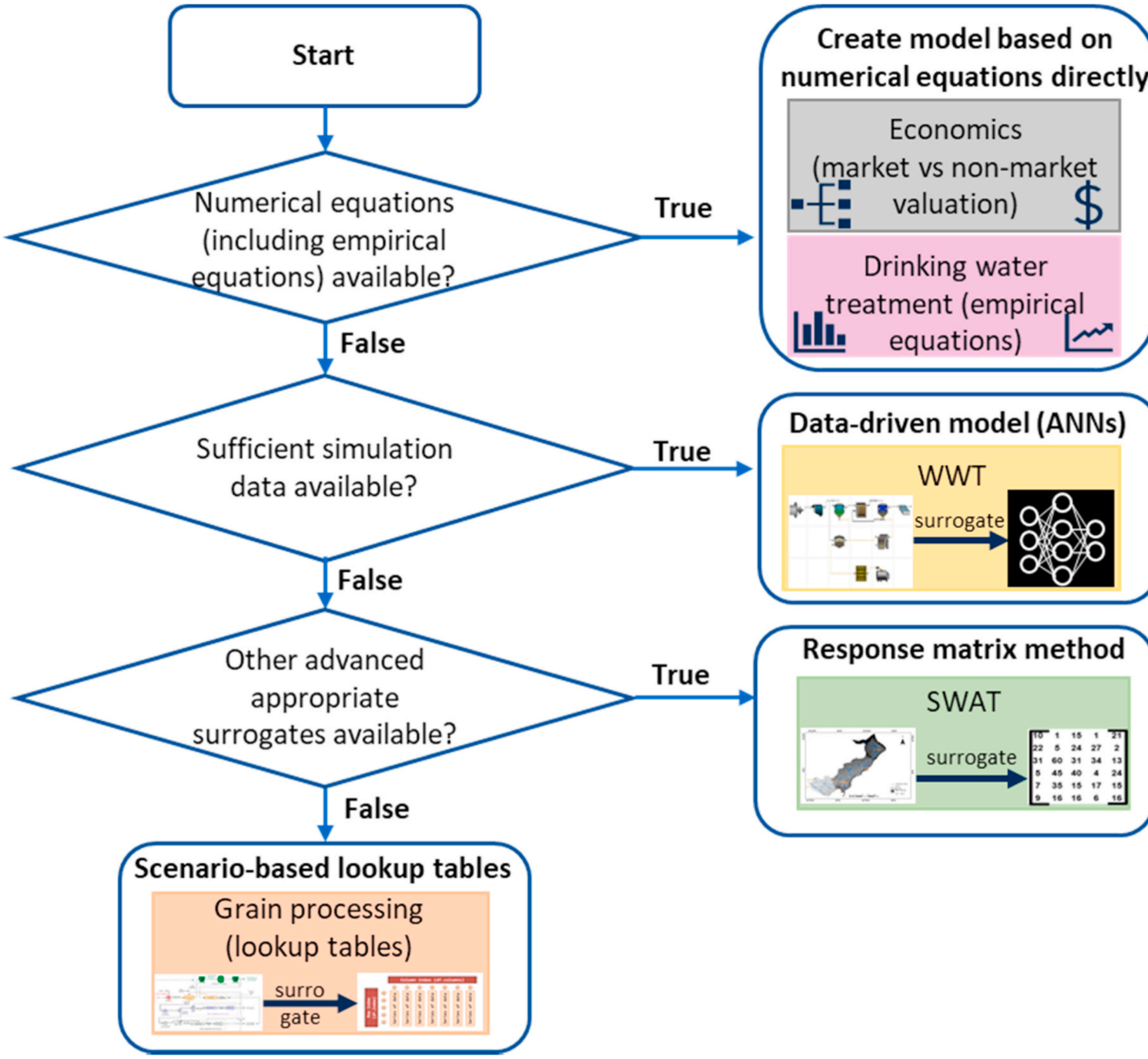


Fig. 3. Decision tree for selecting appropriate forms of surrogates or empirical models in the ITEEM. Primary modes are decided by each group. Note: WWT = Wastewater treatment; ANNs = artificial neural networks; SWAT = Soil & Water Assessment Tool.

reach is then the sum of all upstream landscape loads. Due to its simplicity and effectiveness, the RM method has been widely used in different areas, such as groundwater management models and watershed models (Gorelick, 1983; Housh et al., 2014). The details of the mathematical definitions adopted to implement the RM method are provided in Section 3.1.1.

Third, if advanced surrogates (e.g., ANNs, SVM, RM method) are not appropriate due to data limitation and the characteristics of a computer-based model, as the case of the GP component, we look for another surrogate form too. The data limitation arises that although Superpro Designer can generate sufficient samples via its built-in Monte Carlo simulator, the simulator does not provide outputs with the degree of detail needed for calculating the amount of P flow recovered from process streams. The characteristics of the GP model refers to the fact that, unlike the SWAT that simulates spatial and temporal variables, the GP model generates steady-state operation outputs which are only determined by the plant capacity. For this case, lookup tables, a most basic form of surrogates, are created to determine GP simulation outputs under a set of plant capacities reflecting a common range of commercial grain processing plants. All data in the lookup tables are directly derived from the high-fidelity computer-based model.

3.2. Interactions between components of the ITEEM

After selecting a model form for each of the ITEEM components, the next challenge is to couple the surrogates whose inputs and outputs have distinct spatial and temporal scales. This is a common challenge for integrating multiple components within a consistent model, while each case may have its unique complexity when coupling spatially and temporally varied data and processes across components. Cohesive spatial and temporal scales are chosen for ITEEM to couple point source nutrient loadings from WWT with nonpoint source loadings simulated with SWAT. The temporal scale of SWAT outputs can be daily, monthly, or annual, and the spatial scale can be by hydrological response units (HRUs), subwatersheds, or the entire watershed. In contrast, the steady-state WWT model operates at the weekly, bi-weekly, or monthly scale and discharges to a specific point in a watershed. Given such inconsistent spatial and temporal scales, we couple SWAT and WWT models at the monthly scale; to match the spatial scale, we couple SWAT and WWT models at the subwatershed scale (12-digit Hydrologic Unit Code). This reduces hundreds of HRUs to dozens of subwatersheds. Point source nutrient loadings are added as inflows to the channel reach in the subwatershed where WWT plants are located. More details about coupling over temporal and spatial scales are provided in mathematical formulations (Section 3.3.1.1.).

Fig. 4 illustrates the multiple interacting feedback loops among the components in ITEEM. SWAT has the most interactions with other components. The inputs to SWAT include BMP (e.g., tillage, P fertilizer rates, grassed waterways/riparian buffers, etc.) land allocations at the subwatershed level. The simulated nitrate and sediment from SWAT are inputs to the DWT model to estimate the required treatment cost and energy consumption to purify drinking water. SWAT also simulates corn and soybean yields, which are inputs to the GP model that simulates the amount of recovered P (rP). The resulting amount of phytin-based fertilizer produced from rP is then simulated as a substitute that displaces mined P fertilizer. Besides, the P recovery from corn grain byproducts (CGF, DDGS) causes a reduction of P content in livestock diets, leading to reduced manure P content and ultimately reduced P runoff from feedlots. The selection of a wastewater treatment technology also has implications for P recovery, nutrient discharge, and cost. The four WWT alternatives are:1) activated sludge, 2) activated sludge with chemical precipitation, 3) a modified 5-stage Bardenpho process with enhanced biological phosphorus removal (EBPR), and 4) a modified 5-stage Bardenpho EBPR process with struvite (a form of P) recovery (EBPR_StR)).

Replacing mined P with an rP product (i.e. struvite from WWT or phytin from GP) is an additional agricultural BMP within ITEEM. Via techno-economic analysis, the costs (e.g., capital, operation, labor, and maintenance costs) for the WWT, GP, and DWT technologies and practices are first calculated as total present value and then converted into equivalent annualized cost (EAC), expressed as "$ per year", and their associated energy requirements are also calculated. The other two economic components of ITEEM include: 1) linking water quality levels to public willingness-to-pay (WTP) and farmers' willingness-to-accept (WTA) payment to adopt new conservation practices, and 2) calculating the total net benefit accounting for engineering technologies, farm management practices, and non-market environmental benefit. Beyond the interactions between these components, ITEEM as a whole is driven by the climate, market price of crops and rP fertilizer products, policy regulations on maximum contaminant level (MCL) for nitrate in drinking water and wastewater nutrient effluent limits, and technology options proposed for WWT and GP components.

3.3. *Overview of components in the ITEEM*

The basic overview of each surrogate and empirical model in ITEEM are provided in this section. Detailed mathematical formulations of each modeling component are provided in SI Section 2.

3.3.1. *Response matrix (RM) for SWAT*

SWAT simulates water quality (i.e., nitrate, total phosphorus, and sediment yield), water quantity (i.e., streamflow), and crop yield (i.e., corn, soybean, corn silage, perennial grass) for each hydrologic response unit (the smallest spatially homogeneous unit in SWAT) and aggregates to the subwatershed scale. The computational time for running SWAT can be expensive, from minutes to hours, depending on temporal and spatial scales, and the number of simulations. The RM method has been used previously to approximate the impacts of different crop allocations on simulated water and nutrient yield from landscapes (Housh et al., 2014). To apply the RM method, SWAT-simulated water, sediment, nutrient and crop yields under various scenarios of complete BMP adoption are stored in a set of response matrices. This initial simulation may require large computational efforts. However, the resulting RM can efficiently handle a large set of decision variables (i.e., the land area of BMP adoption in each subwatershed) involved in watershed management. It is worth noting that the traditional RM method only estimates the landscape loss (e.g., nonpoint source nutrient and sediment contributing to rivers) (See detailed calculations in SI Section 2.1.1). In

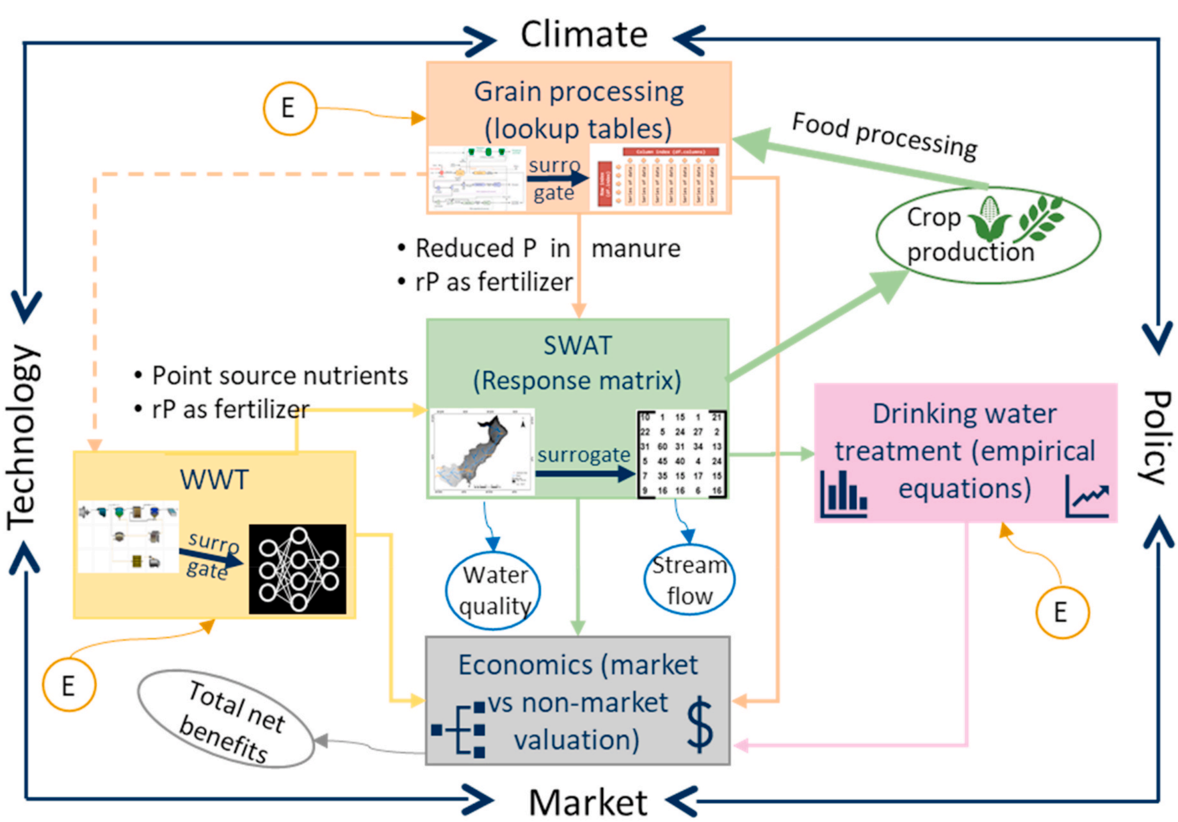


Fig. 4. Interaction of components in the ITEEM. A total of five components are represented as rectangles while the inputs and outputs are shown in ellipses. The interactions between different components are evaluated in ITEEM and denoted by solid arrows. The dashed arrow from grain processing to WWT implies an intrinsic connection between the two components, but the interaction is not explicitly evaluated in ITEEM due to data unavailability. Note: SWAT = Soil & Water Assessment Tool; rP = recovered phosphorus; E = energy demand; WWT = Wastewater treatment.

this work, we extend the traditional RM to further account for point-source pollution, reservoir trapping, and in-stream processes so that in-stream loading can be accurately estimated.

First, the point-source loading of nitrate and TP simulated from the WWT component (see Section 3.3.2) is added to the subwatershed where WWT plants are located. Second, modifications of the traditional RM method are also required to account for the trapping of sediment and nutrients in reservoirs. Third, in-stream processes such as nutrient cycling (e.g. settling and microbial uptake/respiration) and sediment deposition must be considered in order to estimate the final in-stream loading. For the special case of sediment, SWAT-simulated loads are strongly controlled by in-stream deposition and degradation. These in-stream sediment processes cannot be effectively accommodated for by an affine function land area allocation. Therefore, we instead assume that all streams carry their full, flow-limited sediment capacity, calculated similarly to the simplified version of the Bagnold sediment stream equation which is an option within SWAT (though another, better-performing option is applied in our SWAT simulations); where the simplified Bagnold equation of SWAT determines sediment-carrying capacity according to flow velocity, we estimate the capacity according to volumetric flow rate (Bagnold, 1977). That is, whenever incoming sediment loads exceed the flow-determined capacity for sediment, sediment is deposited. When capacity exceeds incoming load, sediment is eroded from the streambed. Detailed calculations of those modifications can be found in SI Section 2.1.2. Unlike streamflow, sediment, and nutrients that are dependent on its spatial reach network and in-stream processes, the total watershed crop production is simply the sum of that in each subwatershed (SI Section 2.1.3).

3.3.2. Artificial neural networks for WWT

WWT plants in the Corn Belt can contribute considerable point source nutrient loading (nitrate-N, TP) due to combined sewerage from stormwater, domestic, and high-strength industrial (especially from biorefineries) wastewater. The WWT component of ITEEM includes four wastewater treatment plant design alternatives to treat the combined influent. The four alternatives include: 1) activated sludge (AS), 2) activated sludge with chemical precipitation (ASCP) to reduce effluent P concentrations from the WWT plant, 3) a modified 5-stage Bardenpho process with enhanced biological phosphorus removal (EBPR), 4) and a modified 5-stage Bardenpho EBPR process with struvite (magnesium ammonium phosphate) recovery (EBPR_StR). We include the impact of stormwater that causes highly variable treatment performance during the process development using GPS-X software. Detailed descriptions of process development for each treatment alternative are provided in SI Section 1.2.

As the wastewater treatment involves complex and nonlinear physical and biological processes, advanced data-driven techniques can be applied to predict treatment performance under fluctuations of influents. Artificial neural networks (ANNs) have been widely applied in various fields to capture nonlinear, complex relationships between inputs and outputs. For a generic ANN, a vector of input data ($\boldsymbol{x}$) can be mapped to a vector of output data ($\boldsymbol{y}$), i.e., $\boldsymbol{y} = f_{ANN}(\boldsymbol{x})$, where $f_{ANN}(\boldsymbol{x})$ represents a function of neural networks. In this study, four feed-forward back-propagation ANNs are applied to surrogate the four WWT alternatives (i.e., AS, ASCP, EBPR, EBPR_StR). Once plant layouts of the four WWT alternatives are designed and optimized, we simulate stochastic influent conditions and run process simulations for each WWT design alternative, to account for WWT performance variability. Each treatment alternative is simulated 10,000 times in the original high-fidelity model using the GPS-X software and the dataset is split 60%, 20%, 20% into training, validation, and test datasets. Details of ANNs training are provided in SI Section 2.2.

After successfully training ANNs for the four treatment alternatives, the next step is to predict the effluents under stochastic conditions of influent using ANNs. To be consistent with the temporal scale of the SWAT component and reduce the computational time, we simulated monthly loading from the WWT, assuming each month is run as a steady state. For each month, the total inflow (domestic and industrial wastewater + rainwater) is determined using historical data while the chemical oxygen demand (COD), total Kjeldahl nitrogen (TKN), total phosphorus (TP)are randomly sampled 1000 times from their fitted historical distributions. Since SWAT simulations provide deterministic monthly values, the monthly mean values of effluent loadings from the 1000 simulations are calculated and added into the subwatershed where the WWT is located. This is a key step to integrate point source and nonpoint source pollutant loadings from the different components. The techno-economic analysis of the four treatment alternatives is conducted using a combination of modeling and calculations. Specifically, the capital costs (e.g., construction) and fixed operational costs (e.g., labor, maintenance) are calculated in CapdetWorks, a proprietary software compatible with GPS-X for estimating fixed costs (capital, labor, maintenance cost) for WWT models. Operational costs that vary with influent characteristics are calculated with process design and cost estimate equations from the US Environmental Protection Agency (Harris et al., 1982).

3.3.3. Lookup tables for grain processing

Two grain processing (GP) models (i.e., corn wet milling and corn dry grind) are developed in SuperPro Designer (Intelligen, Inc.), which contains rigorous reactor modules for mechanical and chemical engineering of corn grain processing (Juneja et al, 2019, 2020) and details of process development are provided in SI Section 1.3. Since SuperPro Designer is commercially programmed and cannot be directly connected with the other ITEEM components, we develop lookup tables that store results simulated from SuperPro Designer. The lookup tables contain two plant layouts for each plant capacity. The capital cost, operational cost, energy and water use, and P content of CGF and DDGS are simulated for each plant capacity. In both wet milling and dry grind corn processing models, two plant layouts are simulated: 1) status quo grain processing without P recovery; and 2) alternative technology that processes grain and recovers P as P complex, which can be further purified as phytin (a calcium magnesium salt of phytic acid). Calculations of cost and energy use are provided in SI Section 2.3.

3.3.4. Empirical equations for DWT

The DWT model is developed based on operational data from a drinking water treatment plant located in the Corn Belt. The cost data include fixed and variable costs for the nitrate removal facility (NRF). The fixed cost includes management overhead, labor cost for operation and maintenance, depreciation cost, and NRF energy cost. Note that the daily energy consumption in NRF is assumed constant as detailed data are unavailable. The variable cost includes the use of sodium chloride as the regenerant chemical for ion-exchange resins and alum and polymer for turbidity treatment. The consumption of sodium chloride is dependent on the nitrate level in the untreated water entering the DWT plant. The consumption of alum and polymer is dependent on the sediment concentration in the raw water entering the DWT plant. The costs do not include the total cost in the main treatment facility as the purpose of the DWT component is to estimate costs and energy requirements associated with excess nitrate and sediment treatment only.

The nitrate-N (NO_3–N) and sediment loadings and streamflow estimated from SWAT in the subwatershed where the DWT plant is located are inputs to the DWT plant component. The decision to operate the NRF is based on daily NO_3–N concentration in the untreated water entering the DWT plant. The NRF will operate on any day where the influent NO_3–N concentration exceeds the threshold of 8.0 NO_3–N mg/L, based on the current maximum contaminant level (MCL) of 10 mg/L; that is, we assume the operation threshold nitrate concentration to operate the NRF is 80% of the MCL. Calculations of cost and energy use for nitrate and sediment treatment are provided in SI 2.4.

3.3.5. Theoretical-empirical equations for economics

3.3.5.1. Non-market valuation for water quality improvement. Choice Experiments (CEs) are conducted to elicit the general public's WTP for water quality improvements and farmers' WTA payment to change management practices. CEs are a widely used non-market valuation method in which respondents are asked to select the most preferred alternative in a hypothetical decision-making situation while varying the levels of different attributes of interest (Louviere et al., 2000). We utilize fractional factorial designs to allocate attribute levels (i.e. cost, recreational options, attainment of nutrient loss reduction goals) to each alternative — a single combination of attribute levels — and choice sets (a combination of different alternatives to choose from). The responses to the CEs are analyzed using statistical methods based on Random Utility Theory (RUT). The utility of an individual survey respondent n from alterative a ($U_{n,a}$) includes systematic ($V_{n,a}$) and stochastic ($\varepsilon_{n,a}$) components with $U_{n,a} = V_{n,a} + \varepsilon_{n,a} = \boldsymbol{\beta}^T \boldsymbol{X}_{n,a} + \varepsilon_{n,a}$. Vector $\boldsymbol{X}_{n,a}$ contains attribute levels faced by individual n in alternative a and $\boldsymbol{\beta}$ is a vector of coefficients estimated corresponding to the attributes. Because the non-market valuation study is not yet complete, this study uses a published WTP estimate from a related study (Parthum and Ando, 2020) to demonstrate how the choice experiment results will be used in the ITEEM.

3.3.5.2. Evaluation of ITEEM total costs and benefits. After estimating the costs of engineered technologies and agricultural management practices, the total economic net benefits of the entire system are calculated for each option. Each option includes a combination of WWT and bioprocessing technology together with a spatially explicit configuration of agricultural practices. Total benefits include the sum of revenue from product sales and non-market benefits associated with water quality changes, and the total cost is the sum of those incurred for wastewater and drinking water treatment, grain processing, and agricultural management practice implementation. The total net benefits are calculated as:

$$\Delta B = B_{WTP} + R_{rP} + R_{crop} + R_{GP} - C_{WWT} - C_{DWT} - C_{GP} - C_{ag} \tag{1}$$

where ΔB (\$/yr) is the economic total net benefits for a given option. B_{WTP} (\$/yr) is the monetary measure of public WTP for water quality improvements. R_{rP} (\$/yr) is the revenue generated by selling recovered P product; R_{crop} (\$/yr) is crop revenue. R_{GP} (\$/yr) is the revenue from grain processing products sold. C_{WWT} (\$/yr) is the cost of wastewater treatment. C_{DWT} (\$/yr) is the cost of drinking water treatment. C_{GP} (\$/yr) is the cost of grain processing plants. C_{ag} (\$/yr) is the total cost of all agricultural management practices applied a given scenario. All the terms in Eq. (1) are annualized cash flow that takes factors of the time value of money and inflation into account. Detailed calculations of terms in Eq. (1) are provided in SI Section 2.

3.4. Sensitivity analysis of ITEEM

As each of the components contributes uncertainty to ITEEM via BMPs, environmental engineering technologies and policies, it is important to investigate how the uncertainties from different component models propagate and affect the overall outputs of ITEEM at the system level. The multiple sources of uncertainties, which can be correlated, complicate the sensitivity analysis. For a demonstration purpose, we conduct a simple one-at-a-time (OAT) sensitivity analysis of key parameters, and leave more complete global sensitivity analysis for future work, which indeed can be standard-alone study. We use a sensitivity indicator calculated as below:

$$Sensitivity_{i,j} = \frac{1}{k} \times \sum_{1}^{k} \left| \frac{\Delta output_{j,k} / output_{j,k,baseline}}{\Delta parameter_{i,k} / parameter_{i,k,baseline}} \right| \tag{2}$$

where $Sensitivity_{i,j}$ represents the averaged relative change of parameter i on output j across different scenarios (k). $\Delta output_{j,k}$ is the change of output j of scenario k from baseline output ($output_{j,k,baseline}$). $\Delta parameter_{i,k}$ is the change of parameter i of scenario k from baseline ($parameter_{j,k,baseline}$). The multi-dimension outputs of ITEEM are aggregated into four categories: 1) water quality and quantity; 2) energy consumption of engineering systems; 3) costs and benefits; 4) production of crop and recovered P. The key parameters investigated in ITEEM includes six parameters from SWAT (e.g., runoff curve, water capacity in the soil, soil evaporation, etc), two parameters (i.e., influent nutrient strength and inflow) in WWT model, and other parameters (i.e., crop price, chemical price, utility cost, willingness to pay per household, and interest rate) related to costs and revenues across various models. Detailed descriptions and the range of each of these parameters are provided in SI Table S4.

4. Computational implementation of ITEEM in object-oriented programming platform

A coherent computational framework is developed to link and execute models from individual knowledge domains in an orderly manner. Standards of integrated modeling have been promoted by researchers to produce a useable and low-friction simulation environment, such as the Community Surface Dynamics Modeling System (CSDMS) project by Peckham et al., in 2013. The design criteria include but not limited to support of multiple operating system, use of open-source tools rather than proprietary software, ease of reusability and maintenance, etc. We develop ITEEM using the object-oriented programming in Python, which fits the several standards promoted by CSDMS. An object-oriented framework connects models as inherited objects where some models are parent objects for others. There are several advantages for using an object-oriented framework. First, by inheriting attributes (variables of an object) and methods (functions of an object), new child objects can be easily built, which meets CSDMS' design criterion of code reusability. Second, creating various methods within the same object allows distinguishing separate functionalities (e.g., technology cost versus treatment performance) thus exhibiting a clear structure for the ease of maintenance. The ITEEM developed in the object-oriented language can also be easily converted to a different language using language interoperability tool (e.g., Babel) (Peckham et al., 2013). There are some other features that need to be improved in the future, such as the support of serial and parallel computation.

The five component models of ITEEM are modularized as five independent objects. Note that the five component models of ITEEM are represented either in not original primary model at the process level; they are rather their surrogates derived from primary process models or simplified or empirical models, which are all integrated at the system level. For example, Class "SWAT()" is simulated by its surrogate model, the response matrix method. Variables are stored as attributes; functions are partitioned into various methods within each object. Fig. 5 shows the implementation and integration of ITEEM described in the unified modeling language (UML). Routines of data exchange are specified either in the attributes or inputs of methods in objects. Specific outputs of interest are obtained by calling specific methods. For example, a method called "get_loading("nitrate")" is defined within the "SWAT" object to obtain nitrate loading specific to a particular scenario, instead of calling SWAT to produce all outputs simultaneously. Based on the five objects that represent the five components, respectively, an overall object that integrates the five components is created as an "ITEEM" object that incorporates attributes (variables) and methods (functions) from all components into a single entity. Such an entity provides a computationally efficient model based on a large set of interconnected technology, environment, and economic relationships. The detailed descriptions of attributes and methods are provided in SI Table S3.

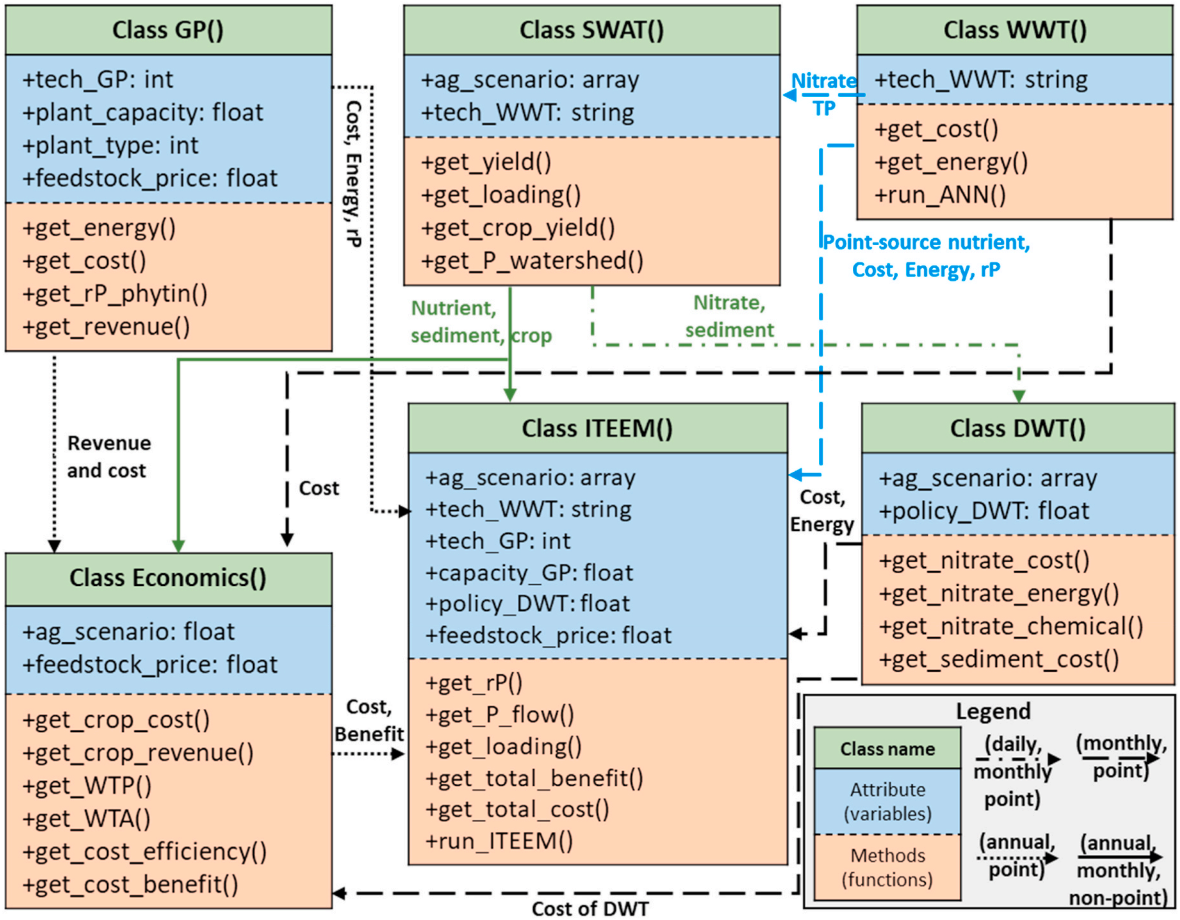


Fig. 5. Unified modeling language description of ITEEM in an object-oriented programming platform. For brevity, only attributes and selected methods are shown in each class and arguments in methods are not included. Dashed lines between objects represent monthly and point scale data flow; dotted lines represent annual and point scale data flow; The only solid line from "SWAT" object to "ITEEM" represents a mixture of monthly (e.g., nutrient loadings), annual (e.g., crop yield), and nonpoint source scale that covers all subwatersheds.

5. Demonstration of ITEEM in the Upper Sangamon River Watershed, Illinois

We demonstrate ITEEM via a testbed watershed, the Upper Sangamon River Watershed (USRW). Different scenarios are tested to explore a portfolio of alternative engineering technologies, policies, and BMPs; the results of the scenarios are compared to a baseline scenario in terms of multiple FEW systems indicators.

5.1. Study area for a testbed watershed

The USRW, located in central Illinois, USA, is selected as an illustrative testbed for its FEW nexus issues, data availability, and existing modeling studies for this watershed (Fig. 6). Water quality in the USRW is threatened by both agricultural runoff and municipal and industrial nutrient discharges. The relatively flat prairie soils in the watershed are highly productive, extensively underlain by subsurface drainage systems, and cultivated for maize and soybean production. Within the watershed, Lake Decatur, created by a dam on the Sangamon River, is the source of municipal water supply for the City of Decatur and the Village of Mount Zion (combined population of 79,000) and industrial water supply for grain processors. The lake has been classified as impaired because of high nitrate and P concentrations and low dissolved oxygen. Periodic dredging of sediment has been necessary to maintain the lake's storage capacity. The cost of nitrate, P, and sediment delivered to the lake from agricultural runoff has been born by the water and wastewater ratepayers in Decatur and Mount Zion. The Sanitary District of Decatur (SDD) treats stormwater, industrial wastewater, and domestic wastewater. SDD discharges treated effluent to the Sangamon River downstream of the Lake Decatur dam. The total discharge of SDD is approximately 600 Mg NO_3–N/yr and 582 Mg P/yr, the largest of any facility in the state of IL, at concentrations typically ranging from 6 to 10 mg NO_3–N/L and 5–30 mg P/L. SDD is faced with the challenge of complying with an impending effluent standard of 1 mg P/L. This is a major challenge because influent concentrations from the biorefineries, responsible for approximately 90% of SDD discharge, are more than twice the typical high range for effective biological P removal. Other WWT plants discharge in the watershed but their contribution to nutrient pollution is relatively small (<5% of total point-source TN and TP).

The three corn grain processing facilities located in the USRW have combined processing capacity of 8.1 million tonnes of corn per year. In addition, one major dairy feedlot produces an estimated 9400 metric tonnes of manure per year from around 3100 milking cows. The status quo P content of manure is assumed to be 9.5 g/kg dry fecal matter. If P can be recovered from grain processing plant waste streams, the P concentrations in manure can be reduced to 6.3 g P/kg dry manure (see Section 3.3.3).

5.2. Scenario analysis

ITEEM enables us to quantify the impacts of various nutrient management strategies, technologies, and policies that could enhance the beneficial synergies of FEW systems in the Corn Belt. To illustrate the use

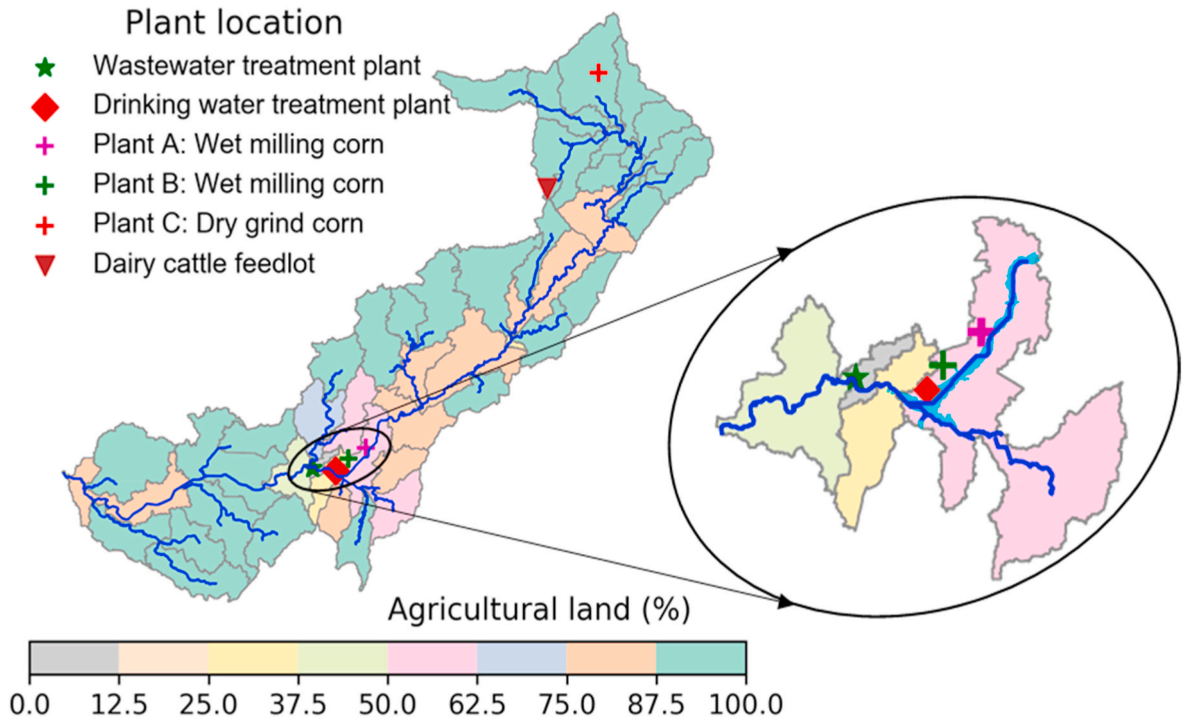


Fig. 6. Map of testbed Upper Sangamon River Watershed in Illinois.

of ITEEM, we simulate and compare four scenarios of agricultural management practices, engineering technologies, and drinking water standards. The baseline scenario consists of the status quo agricultural management practices, WWT technology WWT (i.e., activated sludge without rP), and GP technology (wet-milling and dry-grind corn processing without rP), and current MCL policy governing nitrate in drinking water (i.e., 10 NO_3–N mg/L). The scenario definitions are provided in Table 1.

5.3. Results and discussion

5.3.1. Surrogate model performance

Reducing complex process models into simpler surrogates almost inevitably introduces new uncertainty. Therefore, it is imperative to ensure acceptable performance for the various surrogates used in ITEEM, compared to the original high-fidelity models. Here we evaluate the performance of three surrogate models, i.e., RM for SWAT, ANNs for WWT, and lookup tables for GP. For DWT and Economics components, they are already developed as empirical or theoretical-empirical equations at the process level and there are no additional surrogates applied to upscale them to the system level. The primary component models (e.g., SWAT, WWT, GP) have been developed in different software packages (Juneja et al, 2019, 2020), which however is not the focus of this study. In fact, some software packages, such as WWT, have consensus models that have been validated and used in engineering design.

Traditionally, the RM method is used for landscape loss estimates (Housh et al., 2014). In this work, we extended the RM to account for reservoir trapping, point-source loading, and some in-stream processes (authors who are interested the details of the modified RM, please contact the corresponding author). We chose two widely-used goodness-of-fit measures to assess the performance of the RM method: percent bias (P-bias) and Nash-Sutcliffe efficiency (NSE) (Moriasi et al., 2007). The ideal value of P-bias is zero, indicating no long-term overestimation or underestimation, with positive values indicating overestimation and negative values underestimation. NSE varies from negative infinity to one, one indicating a perfect match between the RM results and SWAT results and with values less than zero indicating that model prediction is less accurate than using the mean of observed data.

To test the performance of the RM method, we select a combination of five BMPs with randomly assigned agricultural land area, compare the results using the RM method versus SWAT, and present detailed results of one realization in Fig. 7. Details regarding the five selected BMPs can

Table 1
Descriptions of alternative scenarios for ITEEM demonstration.

Scenario number	Agricultural management practices[a]	WWT[b]	Grain Processing (GP)	Regulation on drinking water (NO_3–N)
S0 (baseline)	noCC_CT_0red_36%FS	AS	no P recovery	10 mg/L
S1	noCC_RTF_15red_36%FS	ASCP	P recovery	10 mg/L
S2	noCC_RTF_15red_36%FS	EBPR	P recovery	5 mg/L
S3	CC_RTF_30red_50%FS	EBPR_StR	P recovery	10 mg/L

Note:
[a] Each agricultural management practice has four components: 1) cover crop practice, 2) tillage practice; 3) fertilizer rate; 4) waterways/buffers.For cover crop practice: noCC = no cover crop applied; CC = a winter cover crop after corn.For tillage practices: CT = conventional tillage in Fall, reduced tillage in Spring (baseline); RTF = reduced tillage in Fall and Spring.For fertilizer rate: 0red = no reduction from baseline 207 kg diammonium phosphate (DAP); 15red = 15% reduction from baseline; 30red = 30% fertilizer reduction from baseline. For waterways/buffers: 36% FS = 36% of agricultural land are installed with filter strips (baseline); 50%FS = 50% of agricultural land are installed with filter strips.
[b] Wastewater treatment (WWT) has four alternative technologies: 1) activated sludge (AS); 2) activated sludge with chemical precipitation (ASCP); 3) modified Bardenpho enhanced biological phosphorus removal (EBPR); 4) Modified Bardenpho Enhanced Biological Phosphorus Removal with struvite recovery (EBPR_StR).

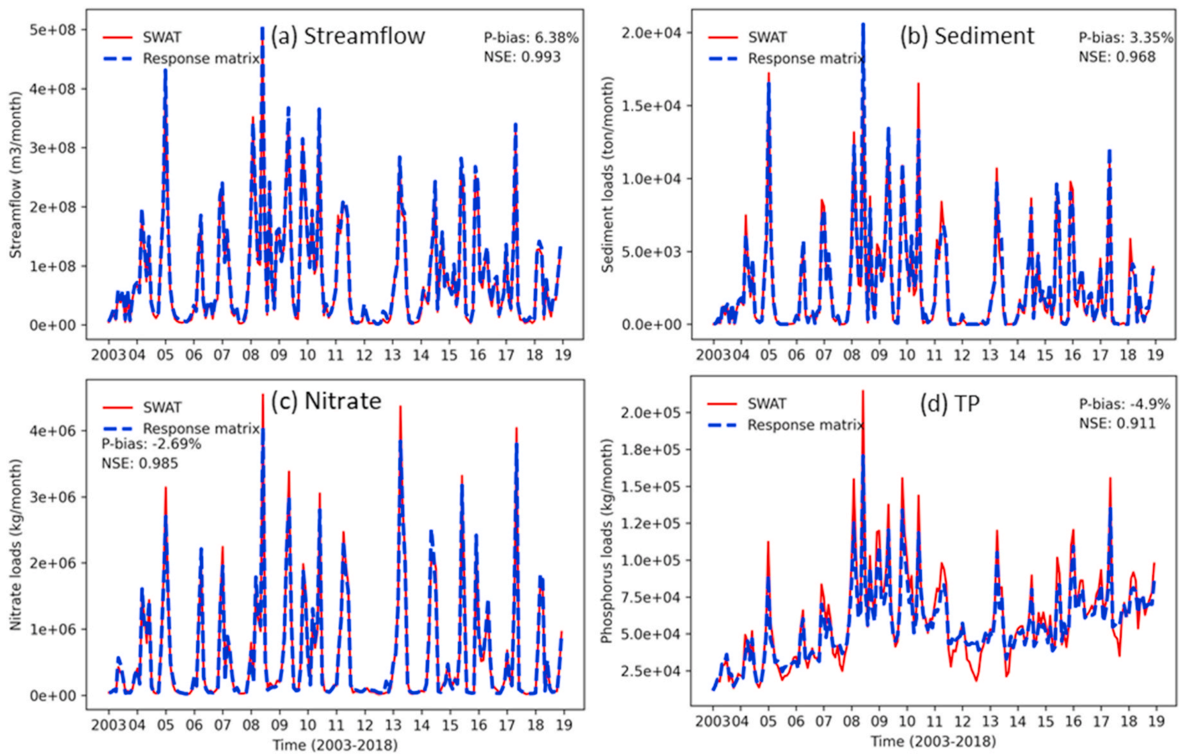


Fig. 7. Simulated results from response matrix and SWAT for the baseline: (a) streamflow, (b) sediment, (c) nitrate, (d) total phosphorus (TP).

be found in SI Table S5, and the spatial agricultural land area allocations for the realization demonstrated in Fig. 7 can be found in SI Table S6. The RM method almost perfectly predicts streamflow and nitrate loads, as evidenced by both P-bias (nearly 0%) and NSE (nearly 1) in Fig. 7a and c. This is because nitrate loading and streamflow are minimally influenced by in-stream processes in the study watershed. Streamflow and nitrate load at the watershed outlet are very similar to the sum of the loads from individual subbasins minus a constant percentage trapped by the reservoir. The RM method also predicts the total phosphorus (TP) and sediment loading with high accuracy in general (P-bias = −4.9% and NSE = 0.911 for TP, P-bias = 3.35% and NSE = 0.968 for sediment). For TP, major discrepancies between the RM method and SWAT simulation are observed during the low streamflow periods (e.g., 2012–2013). This discrepancy arises because the RM method accounts for in-stream P settling and biological P uptake/respiration by applying a constant percent reduction to TP loss from the landscape (11 percent). The constant percent reduction corresponds to some representative travel time for water through the channel network. When water spends more time within the local channel network, the impact of the in-stream processes grows, and accordingly, the SWAT simulation shows the stream acting as a stronger P sink than does the RM method.

As discussed above, sediment is predicted by applying a simplified Bagnold equation to the RM estimates for streamflow (Fig. 7b). We apply streamflow constrained equations to estimate the in-stream sediment loading. The flow-constrained method performs better than simply summing up the landscape losses of sediment. This is because, in this tested watershed, sediment transport in streams is controlled by streamflow, not landscape sediment loss. To test the robustness of the performance of the RM method, we run 10 realizations. The performance of the RM method remains stable, as shown by the detailed results of 10 realizations in SI Table S7. Overall, the RM method in ITEEM performs satisfactorily as a surrogate for SWAT across an explicitly spatial and temporal scale. However, future efforts should be devoted to incorporating a more realistic depiction of sediment deposition and degradation processes.

The ANNs exhibit satisfactory performance in surrogating the complex process-based wastewater treatment model (WWT). Using the simulation data from each of the four WWT alternatives, the ANNs shows high prediction accuracy for all outputs (MSE < 0.001 and R-squared >0.95 as shown in SI Fig. S7). Detailed data of MSE and R-squared for each output in each WWT alternative are provided in SI Table S8. Note that among the outputs predicted by the ANNs, the total nitrate and phosphorus loading from WWT is coupled with the RM method to account for total nutrient loading for the watershed. For the GP model surrogated by lookup tables, all data stored in lookup tables are directly from the simulation results in the high-fidelity model. We assume that each GP plant is operated at steady-state and at its plant capacity, and the simulation data are extracted directly from lookup tables. Therefore, there is no additional uncertainty introduced to this surrogate model.

5.3.2. Suitability of surrogate modeling and applicability of surrogate-based model coupling

As introduced in Section 3.1, we develop surrogate models for complex process models first and then couple the surrogates (along empirical models for water supply and economic analysis) to formulate ITEEM. The surrogate-based model coupling method is suitable only if the following conditions exist. First, if primary process models cannot be integrated directly with compatibilities among discipline-specific computer programs, as discussed by Little et al. (2019). Second, the coupled primary process models are not computationally tractable, especially it is difficult if not possible to use the coupled models for decision analysis, e.g., being coupled with an optimization algorithm (e.g., genetic algorithm) to find optimal solutions.

While surrogate modeling has also been applied to acting as emulators and approximate uncertainty quantification in many domains (Alemazkoor and Meidani, 2020; Razavi et al., 2012; Wu et al., 2014), the difference between surrogate modeling and surrogate-based model coupling should be noted. Surrogate modeling assesses a surrogate to one single high-fidelity simulation model and uses it for certain modeling purposes; surrogate-based model coupling assesses the joint application of multiple surrogate models derived from multiple process simulation models in a consistent modeling framework. Various other model coupling methods have been applied to couple complex process models together, such as wrapper model, scripting, and model translation. Those methods couple high-fidelity models directly and thus have higher accuracy. For instance, Anderson et al. (2018) integrated DSSAT (a crop simulation model) and GREET (a simulation model for energy use and emissions for various vehicles) models via their application programming interfaces (APIs) to simulate the inter-relationships between crop production and environmental impacts of biofuel production (Anderson et al., 2018). However, not all software developers provide APIs, especially for commercial software packages. Xiang et al. (2020) integrated DSSAT with MODFLOW (a groundwater simulation model) by writing scripts for external controls on both models in a batch mode in Python. Model translation constructs different individual models from scratch in a common platform. However, this method can only be feasible for simple simulation models as it requires rewriting all equations included in the primary simulation models (Malard et al., 2017). There is not a single model coupling method that is deemed to be better than others under all cases. Users should consider their strengths and weaknesses for particular applications. For FEW systems analysis present in this study, with appropriate simplifications as described above, our surrogate-based model coupling approach can deliver a computational tractable integrated model at the system level.

5.3.3. Tradeoffs and limitations of surrogate-based integration design choices

Various spatial and temporal scales exist in different component models. For example, the temporal scales can be daily, monthly, and annual; the spatial scales can be a point, watersheds (small and large) and river basin for different component models. Our integration design choices on selecting targeted spatial and temporal scales and their interaction points are driven by decision-making requirements for the model, as well as technical considerations. One of our study purposes is to evaluate solutions based on BMPs and environmental engineering technologies for combined point and non-point source nutrient management. We chose a monthly temporal scale for both point source and non-point source simulation, which is usually sufficient to maintain nutrient mass balance. For water quality related decision-making, the daily raw nitrate level is important and thus the DWT model uses a daily time scaler. Correspondingly, we use the results of SWAT at a daily scale to estimate the daily raw nitrate at the sub-watershed where the DWT locates.

Technically there is a typical tradeoff between the modeling accuracy and computational requirement in the choices of temporal and spatial scales and the aggregation level for the integrated model. In particular, the DWT is modeled at a point scale, where a treatment plant takes raw water from a storage or a river segment within the study watershed. It would be is ideal to have the nitrate concentration at a finer spatial scale because of the spatial variance of the nitrate concentration. However, SWAT only simulates the in-stream loads of nutrients at a sub-watershed level, and the average nitrate concentration in the sub-watershed is taken for the simulation. For WWT, the effluent of point-source nutrients is dynamic and impacted by domestic influent fluctuations, as well as wet weather. However, dynamic modeling of WWT over a long-term period is challenging due to the lack of detailed knowledge of influent wastewater characteristics and of the rainfall translation to plant influent and operation changes. Eventually, we use a steady-state approach to simulate monthly point source nutrients. Such a temporal aggregation from nearly real-time to monthly scale limits the capability of the WWT model in simulating peak stormwater demand caused by an extreme rainfall event. To rigorously quantify the tradeoffs between accuracy and computational time due to different choices of spatial and temporal scales would be an interesting investigation. We do not explicitly quantify such tradeoffs in this paper as we focus more on developing an integrated model that tightly couples process-based and empirical models at the same platform that can be used to test hypotheses and generate insights for watershed nutrient management.

5.3.4. Tradeoffs among food, energy, water, and economics

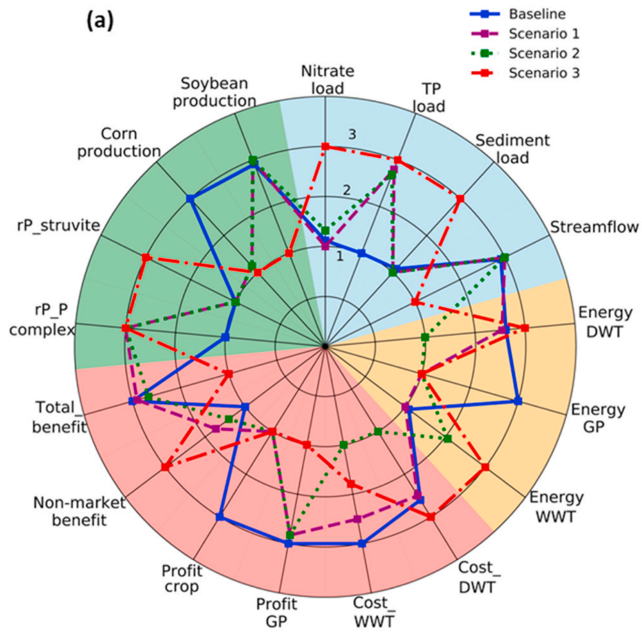
To demonstrate how ITEEM can be applied to explore tradeoffs among multiple metrics of FEW systems, outputs from the three alternatives and the baseline are simulated using the method "run_ITEEM()". To facilitate tradeoff evaluation, we normalize the performance of each metric indicator from 1 to 3, with 1 indicating the worst and 3 indicating the best among all scenarios, as shown in Fig. 8a. The minimum and maximum indicator values are provided in Fig. 8b. Compared to Scenarios 1–3, the Baseline Scenario has the lowest overall performance for water quality and quantity indicators, but the best overall performance on energy consumption, cost of technologies (GP and WWT), and crop production. The result arises because all three alternatives introduce best agricultural management practices, upgrade the existing technology to recover P, and advance point-source P removal.

The non-market benefits represent the estimated willingness-to-pay of general public living in upstream of Decatur reservoir for increases in likelihood of achieving the nutrient reduction target (45% by 2045) in Illinois based on the study of Parthum and Ando (2020). Their study estimates that each household located upstream of the reservoir would be willing to pay \$0.95 per year to increase the likelihood of meeting the nutrient target by one percentage point. Using the estimate, we simplify the non-market benefits under each scenario as: \$0.95/household/year multiplied by 113,700 (approximate number of households in the upstream of the reservoir), and then multiplied by the extent to which a scenario attains the 45% target for Nitrate-N or TP (e.g., if a scenario reduces TP by 45%, then this value is 100). The values in Fig. 8 are the sum of benefits derived from reduced nitrate and TP loads. According to this formulation, the three alternatives generate environmental benefits ranging from 3.0 to 8.1 million dollars/yr. However, those environmental benefits are not enough to offset the associated costs to upgrade treatment technologies and adopt new agricultural BMPs. As evidenced in Fig. 8a, the baseline scenario provides greater total benefits than do the three alternative scenarios. Note that the non-market benefits will be updated based on our choice experiment results and the results might change considerably.

Among the three alternative scenarios, Scenario 3 has the best nitrate reduction and second-best TP reduction due to the adoption of cover crops and the choice of Modified Bardenpho Enhanced Biological Phosphorus Removal with struvite recovery (EBPR_StR) for WWT. The nitrate reduction in Scenario 3 also results in decreased energy consumption and cost associated with nitrate removal at the DWT. Scenarios 1–3 reduce the DAP fertilizer application rate by 15%, 15%, and 30%, respectively, but the impact on corn yield is relatively negligible. This could be because either the baseline is currently over-applying DAP fertilizer or there is enough P accumulated from prior years to make up the gap in crop demand for P. Overall, the three alternatives illustrate the tradeoffs between reduced nutrient loading, energy demand, and cost for alternative technologies.

5.3.5. Sensitivity analysis of ITEEM

We conduct the sensitivity analysis for key parameters, including six parameters from SWAT, two parameters from the WWT, and the other eight parameters of benefits and costs (see Fig. 9 for the list of the parameters). Note that the WW nutrient parameter in the WWT model varies the influent COD, TKN, and TP altogether (Details provided in SI Table S4). The heatmap (Fig. 9) shows the sensitivity results with parameters listed horizontally and multi-dimensional outputs listed vertically under the four scenarios (baseline + three alternatives in Table 1) investigated in this study. We consider a change of ±20% as upper and lower bound for most key parameters, except for the interest rate and willingness to pay per household. As mentioned earlier, the sensitivity results of the key parameters are scenario dependent. Values in the heatmap represent the change to an output responding to one unit change of a parameter. For example, the sensitivity of parameter



(b)

Metrics	Worst (score=1)	Best (score=3)	Unit
Nitrate load	8.07E+06	5.60E+06	kg/yr
TP load	8.28E+05	1.78E+05	kg/yr
Sediment load	3.16E+04	2.85E+04	ton/yr
Streamflow	9.84E+08	1.07E+09	m^3/yr
Energy_DWT	8.42E+00	2.61E+00	TJ/yr
Energy_GP	3.51E+03	3.50E+03	TJ/yr
Energy_WWT	5.22E+01	3.86E+01	TJ/yr
Cost_DWT	4.56E+05	3.13E+05	$/yr
Cost_WWT	2.23E+07	1.91E+07	$/yr
Profit_crop	2.13E+08	2.51E+08	$/yr
Profit_GP	2.696E+08	2.70E+08	$/yr
Benefit_WTP	0.00E+00	8.12E+06	$/yr
Total_benefit	4.70E+08	5.02E+08	$/yr
rP_P_complex	0.00E+00	4.71E+07	$/yr
rP_struvite	0.00E+00	1.28E+06	$/yr
Corn production	1.69E+09	1.71E+09	kg/yr
Soybean production	4.99E+08	5.11E+08	kg/yr

Fig. 8. Tradeoffs between the multiple-dimension indicators of FEW systems. Each colored line represents one scenario described in Table 1. The multiple-dimension indicators are aggregated into four groups: 1) water quality and quantity; 2) energy consumption of engineering systems; 3) cost and benefits; 4) crop and rP production. Note: DWT = Drinking water treatment; GP = Grain processing; WWT = Wastewater treatment; WTP = Willingness to pay; rP = recovered phosphorus.

(SWAT_runoff curve) on the output (nitrate load) is 2.1, as shown in the top left corner of the heatmap, meaning that a 1% change of the runoff curve value from the baseline can cause a 2.1% change of nitrate load at the outlet.

It is found that the six SWAT parameters have the most significant impacts on water quality (i.e., nutrient and sediment load) and quantity (i.e., streamflow) and crop production. The uncertainty from those parameters further propagates into crop revenue, the willingness to pay (WTP), and the total net benefit. The total net benefit is also significantly impacted by the market prices of products (e.g., starch and ethanol) from grain processing plants and the cost of feedstock (e.g., corn sold for grain processing), as the profit from GP has a large contribution to the total net benefit. In contrast, the prices of chemicals, rP, and utility (e.g., electricity and natural gas) have negligible impacts across the outputs.

The two parameters from the WWT model are evaluated with the four treatment alternatives (AS for baseline, ASCP for S1, EBPR_acetate for S2, and EBPR_StR for S3), corresponding to the four scenarios provided in Table 1. The influent nutrient strength (WW nutrient) and inflow of wastewater (WW inflow) have a noticeable impact on the energy use of WWT, the amount of recovered struvite, and the WTP, which demonstrates that the uncertainty of influent characteristics from the WWT model can have a significant impact at the system level outputs. However, we also have two interesting observations. 1) parameters from the WWT model do not have a significant impact on TP load at the outlet (Fig. 9), despite the fact the point source P is the leading contributor to the total TP load for the testbed watershed. 2) parameters are sensitive on the outputs (point source only) of individual WWT treatment alternatives. For example, the sensitivity of WW influent for the four treatment alternatives is 1.38 for point source nitrate and 0.87 for point source TP (details provided in SI Table S9). However, the sensitivity of WW nutrient is decreased when integrated at the system level (point + nonpoint sources) with sensitivity being 0.03 for nitrate and 0.25 for TP (details provided in SI Table S9). Both observations can be attributed to the fact that our local sensitivity analysis is scenario-dependent; for the cases of ASCP, EBPR, and EBPR_StR, the point source nutrients are significantly reduced and are not the leading TP contributor anymore. Therefore, for the scenarios with ASCP, EBPR, and EBPR_StR, the parameter changes on WWT model will not have sensitive impact on total TP load, which ultimately decreases its sensitivity on TP load at the system level.

6. Conclusions and future research

Addressing large-scale environmental sustainability challenges requires integrated analysis of complex inter-relationships within FEW systems. This paper presents the development of an integrated technology-environment-economics model (ITEEM) for typical watersheds in the Corn Belt. We use various data techniques to convert complex models simulating physical & engineering processes and socioeconomic relationships into computationally tractable surrogates and link these surrogates via input-output relationships within a consistent computer-based modeling platform. The procedures for developing ITEEM for a case study watershed (Upper Sangamon River Watershed, USRW) can be applied to other watersheds in the Corn Belt, with required data and model preparation as shown for the USRW.

Based on our experience developing ITEEM with a team including researchers from hydrology and water resources system analysis, environmental engineering, environmental economics, and sociology, we reflect on steps for selecting surrogate models, i.e., which type of data-driven surrogates is most suitable for a particular physical and process model based on data and model availability, as well as the purpose of the integrated model. In this study, we applied the response matrix method and artificial neural networks, respectively, to create surrogates for SWAT and WWT. A detailed process model for DWT is not available for our project, therefore we adopted empirical equations for the DWT component. The lookup table method used for the GP component is simple but sufficient since GP can be assumed to operate at steady state and is only impacted by the size of plant capacity. The economics component is formulated with equations that include non-market valuation estimates and overall economic net benefits. In our one-at-a-time sensitivity analysis, we show to what extent the uncertainty of selected key parameters in the component models can impact the outputs of ITEEM; we identify some critical parameters that are worthy of further investigation. Future work should adopt a global sensitivity analysis considering the correlation of the uncertainties from different component models, as well as the uncertainty due to future climate

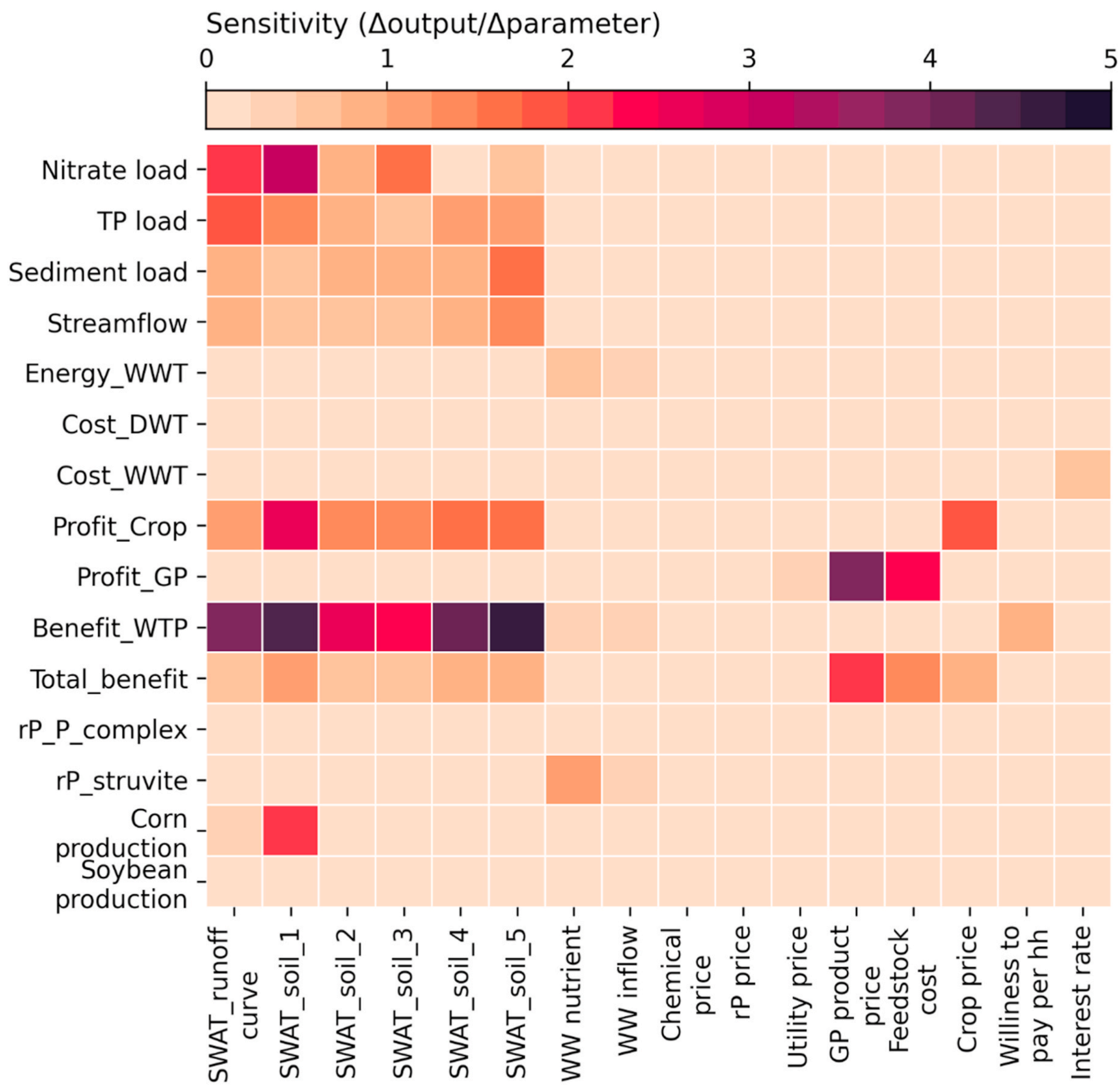


Fig. 9. Sensitivity analysis results of ITEEM outputs (vertical axis) with key parameters (horizontal axis) from component models. All outputs are annual average with the unit provided in Fig. 8b. Note: WWT = Wastewater treatment; DWT = Drinking water treatment; GP = Grain processing; WTP = Willingness to pay; rP = recovered phosphorus.

change.

ITEEM enables testing hypotheses for FEW systems analysis and exploring solutions to resolve inter-connected FEW problems. For example, one hypothesis to test is that the most economically efficient way to improve water quality in Corn Belt watersheds should be to jointly employ a combination of agricultural land management practices and P recovery from co-products generated by grain biorefinery facilities or wastewater treatment. It is noted that ITEEM is designed for evaluating long-term strategic planning for FEW systems in the Corn Belt but not for evaluating short-term events, such as extreme rainfall events that can cause peak stormwater flow that affects both point and non-point pollution. Future work will be conducted to evaluate FEW system resilience under a set of stress and disturbance scenarios. Last but not the least, ITEEM will be coupled with a multi-objective optimization algorithm to search for optimal technologies and policies.

Data and code availability

The data and codes are available from the corresponding author upon request.

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 supported by the US National Science Foundation (INFEWS/T1 award number 1739788). We are grateful to industrial, governmental, and agricultural stakeholders in Decatur, IL for providing valuable data for developing components of the ITEEM. We appreciate valuable comments and constructive suggestions from the editors Daniel P. Ames and Tatiana Filatova, and three anonymous reviewers, which have considerably improved the quality of this work.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envsoft.2021.105083.

References

Alemazkoor, N., Meidani, H., 2020. Fast probabilistic voltage control for distribution networks with distributed generation using polynomial surrogates. IEEE Access 8. https://doi.org/10.1109/ACCESS.2020.2987787.

Anderson, R., Keshwani, D., Guru, A., Yang, H., Irmak, S., Subbiah, J., 2018. An integrated modeling framework for crop and biofuel systems using the DSSAT and GREET models. Environ. Model. Software 108, 40–50. https://doi.org/10.1016/J.ENVSOFT.2018.07.004.

Bagnold, R.A., 1977. Bed load transport by natural rivers. Water Resour. Res. 13, 303–312. https://doi.org/10.1029/WR013i002p00303.

Bennett, E.M., Carpenter, S.R., Caraco, N.F., 2001. Human impact on erodable phosphorus and eutrophication: a global perspective. Bioscience 51, 227–234. https://doi.org/10.1641/0006-3568(2001)051[0227:HIOEPA]2.0.CO;2.

Bhatnagar, A., Sillanpää, M., 2011. A review of emerging adsorbents for nitrate removal from water. Chem. Eng. J. https://doi.org/10.1016/j.cej.2011.01.103.

Cai, X., 2008. Implementation of holistic water resources-economic optimization models for river basin management - reflective experiences. Environ. Model. Software 23, 2–18. https://doi.org/10.1016/j.envsoft.2007.03.005.

Cai, X., Asce, M., Mckinney, D.C., Asce, A.M., Lasdon, L.S., 2003. Integrated hydrologic-agronomic-economic model for river basin management. J. Water Resour. Plann. Manag. 129, 4–17. https://doi.org/10.1061/ASCE0733-94962003129:14.

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.

Carmichael, J., Tansey, J., Robinson, J., 2004. An integrated assessment modeling tool. Global Environ. Change 14, 171–183. https://doi.org/10.1016/j.gloenvcha.2003.12.002.

Carpenter, S.R., 2008. Phosphorus control is critical to mitigating eutrophication. Proc. Natl. Acad. Sci. U.S.A. https://doi.org/10.1073/pnas.0806112105.

Cordell, D., Drangert, J.O., White, S., 2009. The story of phosphorus: global food security and food for thought. Global Environ. Change 19, 292–305. https://doi.org/10.1016/j.gloenvcha.2008.10.009.

Davis, S.E., Cremaschi, S., Eden, M.R., 2018. Efficient surrogate model development: impact of sample size and underlying model dimensions. In: Computer Aided Chemical Engineering. Elsevier B.V., pp. 979–984. https://doi.org/10.1016/B978-0-444-64241-7.50158-0.

Feng, M., Liu, P., Li, Z., Zhang, J., Liu, D., Xiong, L., 2016. Modeling the nexus across water supply, power generation and environment systems using the system dynamics approach: Hehuang Region, China. J. Hydrol. 543, 344–359. https://doi.org/10.1016/j.jhydrol.2016.10.011.

Gaddis, E.J.B., Falk, H.H., Ginger, C., Voinov, A., 2010. Effectiveness of a participatory modeling effort to identify and advance community water resource goals in St. Albans, Vermont. Environ. Model. Software 25, 1428–1438. https://doi.org/10.1016/j.envsoft.2009.06.004.

Gaddis, E.J.B., Voinov, A., 2010. Spatially explicit modeling of land use specific phosphorus transport pathways to improve TMDL load estimates and implementation planning. Water Resour. Manag. 24, 1621–1644. https://doi.org/10.1007/s11269-009-9517-z.

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.

Harou, J.J., Pulido-Velazquez, M., Rosenberg, D.E., Medellín-Azuara, J., Lund, J.R., Howitt, R.E., 2009. Hydro-economic models: concepts, design, applications, and future prospects. J. Hydrol. https://doi.org/10.1016/j.jhydrol.2009.06.037.

Harris, R.W., Cullinane, M.J., Sun, P.T., 1982. Process Design and Cost Estimating Algorithms for the Computer Assisted Procedure for Design and Evaluation of Wastewater Treatment Systems (CAPDET). Environmental Engineering Division, Environmental Laboratory, U.S. Army Engineer Waterways Experiment Station.

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.

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., Sharpley, A.N., Flaten, D., Kleinman, P.J.A., Jenkins, A., Simmons, T., 2015. The pivotal role of phosphorus in a resilient water-energy-food security nexus. J. Environ. Qual. 44, 1049–1062. https://doi.org/10.2134/jeq2015.01.0030.

Jayakrishnan, R., Srinivasan, R., Santhi, C., Arnold, J.G., 2005. Advances in the application of the SWAT model for water resources management. Hydrol. Process. 19, 749–762. https://doi.org/10.1002/hyp.5624.

Juneja, A., Cusick, R., Singh, V., 2020. Recovering phosphorus as a coproduct from corn dry grind plants: a techno-economic evaluation. Cereal Chem. 10260 https://doi.org/10.1002/cche.10260.

Juneja, A., Sharma, N., Cusick, R., Singh, V., 2019. Techno-economic feasibility of phosphorus recovery as a coproduct from corn wet milling plants. Cereal Chem. 96, 380–390. https://doi.org/10.1002/cche.10139.

Leck, H., Conway, D., Bradshaw, M., Rees, J., 2015. Tracing the water-energy-food nexus: description, theory and practice. Geogr. Compass 9, 445–460. https://doi.org/10.1111/gec3.12222.

Lemke, A.M., Kirkham, K.G., Lindenbaum, T.T., Herbert, M.E., Tear, T.H., Perry, W.L., Herkert, J.R., 2011. Evaluating agricultural best management practices in tile-drained subwatersheds of the Mackinaw River, Illinois. J. Environ. Qual. 40, 1215–1228. https://doi.org/10.2134/jeq2010.0119.

Leperi, K.T., Yancy-caballero, D., Snurr, R.Q., You, F., 2019. 110th Anniversary : surrogate models based on artificial neural networks to simulate and optimize pressure swing adsorption cycles for CO2 capture. Ind. Eng. Chem. Res. https://doi.org/10.1021/acs.iecr.9b02383.

Li, S., Qin, Y., Subbiah, J., Dvorak, B., 2020. Life cycle assessment of the U.S. beef processing through integrated hybrid approach. J. Clean. Prod. 121813 https://doi.org/10.1016/j.jclepro.2020.121813.

Li, S., Ziara, R.M.M., Dvorak, B., Subbiah, J., 2018. Assessment of water and energy use at process level in the U.S. beef packing industry: case study in a typical U.S. large-size plant. J. Food Process. Eng. 41, e12919 https://doi.org/10.1111/jfpe.12919.

Little, J.C., Hester, E.T., Elsawah, S., Filz, G.M., Sandu, A., Carey, C.C., Iwanaga, T., Jakeman, A.J., 2019. A tiered, system-of-systems modeling framework for resolving complex socio-environmental policy issues. Environ. Model. Software 112, 82–94. https://doi.org/10.1016/j.envsoft.2018.11.011.

Louviere, J.J., Hensher, David A., Swait, J.D., 2000. Stated Choice Methods: Analysis and Applications. Cambridge University Press.

Lu, D., Ricciuto, D., 2019. Efficient surrogate modeling methods for large-scale Earth system models based on machine-learning techniques. Geosci. Model Dev. (GMD) 12, 1791–1807. https://doi.org/10.5194/gmd-12-1791-2019.

Malard, J.J., Inam, A., Hassanzadeh, E., Adamowski, J., Tuy, H.A., Melgar-Quiñonez, H., 2017. Development of a software tool for rapid, reproducible, and stakeholder-friendly dynamic coupling of system dynamics and physically-based models. Environ. Model. Software 96, 410–420. https://doi.org/10.1016/j.envsoft.2017.06.053.

Margenot, A., Kitt, D., Gramig, B.M., Berkshire, T., Chatterjee, N., Hertzberger, A., Aguiar, S., Furneaux, A., Sharma, N., Cusick, R.D., 2019. Toward a regional phosphorus (re)cycle in the U.S. Midwest. J. Environ. Qual. 1413, 1397–1413. https://doi.org/10.2134/jeq2019.02.0068.

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.

Muttiah, R.S., Wurbs, R.A., 2002. Scale-dependent soil and climate variability effects on watershed water balance of the SWAT model. J. Hydrol. 256, 264–285. https://doi.org/10.1016/S0022-1694(01)00554-6.

Nahm, K.H., 2002. Efficient feed nutrient utilization to reduce pollutants in poultry and swine manure. Crit. Rev. Environ. Sci. Technol. https://doi.org/10.1080/10643380290813435.

Ng, T.L., Eheart, J.W., Cai, X., Braden, J.B., 2011. An agent-based model of farmer decision-making and water quality impacts at the watershed scale under markets for carbon allowances and a second-generation biofuel crop. Water Resour. Res. 47, 1–17. https://doi.org/10.1029/2011WR010399.

Parthum, B., Ando, A.W., 2020. Overlooked benefits of nutrient reductions in the Mississippi river basin. Land Econ. https://doi.org/10.5281/zenodo.3692738.

Peckham, S.D., Hutton, E.W.H., Norris, B., 2013. A component-based approach to integrated modeling in the geosciences: the design of CSDMS. Comput. Geosci. 53, 3–12. https://doi.org/10.1016/j.cageo.2012.04.002.

Powers, S.M., Bruulsema, T.W., Burt, T.P., Chan, N.I., Elser, J.J., Haygarth, P.M., Howden, N.J.K., Jarvie, H.P., Lyu, Y., Peterson, H.M., Sharpley, A.N., Shen, J., Worrall, F., Zhang, F., 2016. Long-term accumulation and transport of anthropogenic phosphorus in three river basins. Nat. Geosci. 9, 353–356. https://doi.org/10.1038/ngeo2693.

Rao, N.S., Easton, Z.M., Schneiderman, E.M., Zion, M.S., Lee, D.R., Steenhuis, T.S., 2009. Modeling watershed-scale effectiveness of agricultural best management practices to reduce phosphorus loading. J. Environ. Manag. 90, 1385–1395. https://doi.org/10.1016/j.jenvman.2008.08.011.

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.

Sharpley, A., Jarvie, H.P., Buda, A., May, L., Spears, B., Kleinman, P., 2013. Phosphorus legacy: overcoming the effects of past management practices to mitigate future water quality impairment. J. Environ. Qual. 42, 1308–1326. https://doi.org/10.2134/jeq2013.03.0098.

Simpson, T.W., Sharpley, A.N., Howarth, R.W., Paerl, H.W., Mankin, K.R., 2008. The new gold rush: fueling ethanol production while protecting water quality. J. Environ. Qual. 37, 318–324. https://doi.org/10.2134/jeq2007.0599.

Twomey, K.M., Stillwell, A.S., Webber, M.E., 2010. The unintended energy impacts of increased nitrate contamination from biofuels production. J. Environ. Monit. 12, 218–224. https://doi.org/10.1039/b913137j.

US EPA, 2017. Mississippi River/Gulf of Mexico Watershed Nutrient Task Force: 2017 Report to Congress. Washington, DC.

Wu, B., Zheng, Y., Tian, Y., Wu, X., Yao, Y., Han, F., Liu, J., Zheng, C., 2014. Systematic assessment of the uncertainty in integrated surface water-groundwater modeling based on the probabilistic collocation method. Water Resour. Res. 50, 5848–5865. https://doi.org/10.1002/2014WR015366.

Wu, Y., Liu, S., Qiu, L., Sun, Y., 2016. SWAT-DayCent coupler: an integration tool for simultaneous hydro-biogeochemical modeling using SWAT and DayCent. Environ. Model. Software 86, 81–90. https://doi.org/10.1016/j.envsoft.2016.09.015.

Xiang, Z., Bailey, R.T., Nozari, S., Husain, Z., Kisekka, I., Sharda, V., Gowda, P., 2020. DSSAT-MODFLOW: a new modeling framework for exploring groundwater conservation strategies in irrigated areas. Agric. Water Manag. 232, 106033. https://doi.org/10.1016/j.agwat.2020.106033.

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.