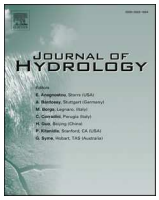
Journal of Hydrology 624 (2023) 129959

Contents lists available at ScienceDirect

Journal of Hydrology

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



Research papers

# Pay-for-practice or Pay-for-performance? A coupled agent-based evaluation tool for assessing sediment management incentive policies

Chung-Yi Lin [a,1], Y.C. Ethan Yang [a,*], Anil Kumar Chaudhary [b]

[a] Department of Civil and Environmental Engineering, Lehigh University, Bethlehem 18015, PA, USA
[b] Department of Agricultural Economics, Sociology, and Education, Pennsylvania State University, University Park 16802, PA, USA

ARTICLE INFO

This manuscript was handled by Nandita Basu, Editor-in-Chief, with the assistance of Jan Adamowski, Associate Editor

*Keywords:*
Water quality management
Chesapeake Bay Watershed
Sediment modeling
Cost-share program
Behavioral sensitivity analysis
Best management practice

ABSTRACT

Cost-shared programs have been applied to incentivize the adoption of agricultural best management practices (BMPs) to address the long-standing water quality issue in the Chesapeake Bay watershed, US. However, the business-as-usual (BAU) incentive program (i.e., pay-for-practice, paying cost share for implementing BMPs) is likely to miss the Total Maximum Daily Load target to reduce 20% of the total suspended sediment (TSS) in 2010 by 2025. Some field experiments indicate that pay-by-performance (PFP; paying lower cost share but with additional bonus payment per unit sediment reduction) can better motivate community involvement leading to greater water quality control outcomes. However, the effectiveness of different incentive policies is still unclear at a basin scale. We propose a coupled agent-based modeling tool to quantify the performance of different incentive policies. The tool considers farmers' (i.e., agents') BMP adoption dynamics affected by the social norm and the potential bonus payment. Specifically, we compare individual-based PFP ($PFP_i$) and group-based PFP ($PFP_g$) with BAU. Results of our proposed model applied to the selected study area, the Susquehanna River Basin, Chesapeake Bay's largest tributary watershed, suggest that PFP can achieve higher TSS reduction with less cost. $PFP_g$ shows the best basin-wide TSS reduction associated with the least uncertainty among all tested policies. Also, the performance of $PFP_g$ is less impacted by the change in the bonus payment compared to $PFP_i$ attributed to farmers' collaboration efforts. Potentially, the proposed policy evaluation tool can better inform an achievable target with policy suggestions in assistance with social studies (e.g., surveys and behavioral experiments).

## 1. Introduction

Suspended sediment transport plays an important role in carrying nutrients (e.g., nitrogen, N, and phosphorus, P) from soil storage pools to oceans or lakes through river systems (Li et al., 2022). With increasing nonpoint inputs of nutrients from commercial fertilizers, surplus nutrients impair waterbodies and induce eutrophication at the coastal wetlands or deltas around the world (Carpenter et al., 1998; Mee, 2006). Such eutrophication-induced hypoxia leads to more than half of marine dead zones, including Chesapeake Bay in the US, Kattegat in northern Europe, the Black Sea in southeastern Europe, the Gulf of Mexico in North America, and East China Sea (Diaz and Rosenberg, 2008).

Compared to point sources, regulating nonpoint sources is highly challenging since the responsible entities for nonpoint pollutants are often not identifiable (Hardy and Koontz, 2008). Efforts are required to mitigate the environmental damages. Authorized by Section 303(d) of the Clean Water Act, the US Environmental Protection Agency (EPA) assists states and regional authorities in developing the Total Maximum Daily Load (TMDL). A TMDL establishes the maximum amount of a pollutant (e.g., N, P, or sediment) allowed in a water body and often serves as a target guiding environmental management policy (Borah et al., 2019).

Agriculture is the major source of nonpoint inputs. Although best management practices (BMPs) like conservation tillage, buffer strips, and cover crops are indicated to be an effective way to mitigate N, P, and sediment emissions to the river systems (Liu et al., 2017), the adoption of BMPs is often voluntary and not mandated by law as a requirement. Therefore, incentive programs become a typical way to motivate voluntary compliance to reduce agricultural nonpoint inputs (Feather and Cooper, 1995). For example, cost-share programs are carried out to complement the additional cost and perceived loss in crop production for implementing BMPs (Commender et al., 2020).

* Corresponding author.
*E-mail addresses:* chl319@lehigh.edu, chungyi@vt.edu (C.-Y. Lin), yey217@lehigh.edu (Y.C.E. Yang), auk259@psu.edu (A. Kumar Chaudhary).
[1] Present affiliation and address: Department of Civil and Environmental Engineering, Virginia Tech, Blacksburg 24061, VA, USA.

https://doi.org/10.1016/j.jhydrol.2023.129959
Received 22 February 2023; Received in revised form 1 May 2023; Accepted 11 July 2023
Available online 20 July 2023

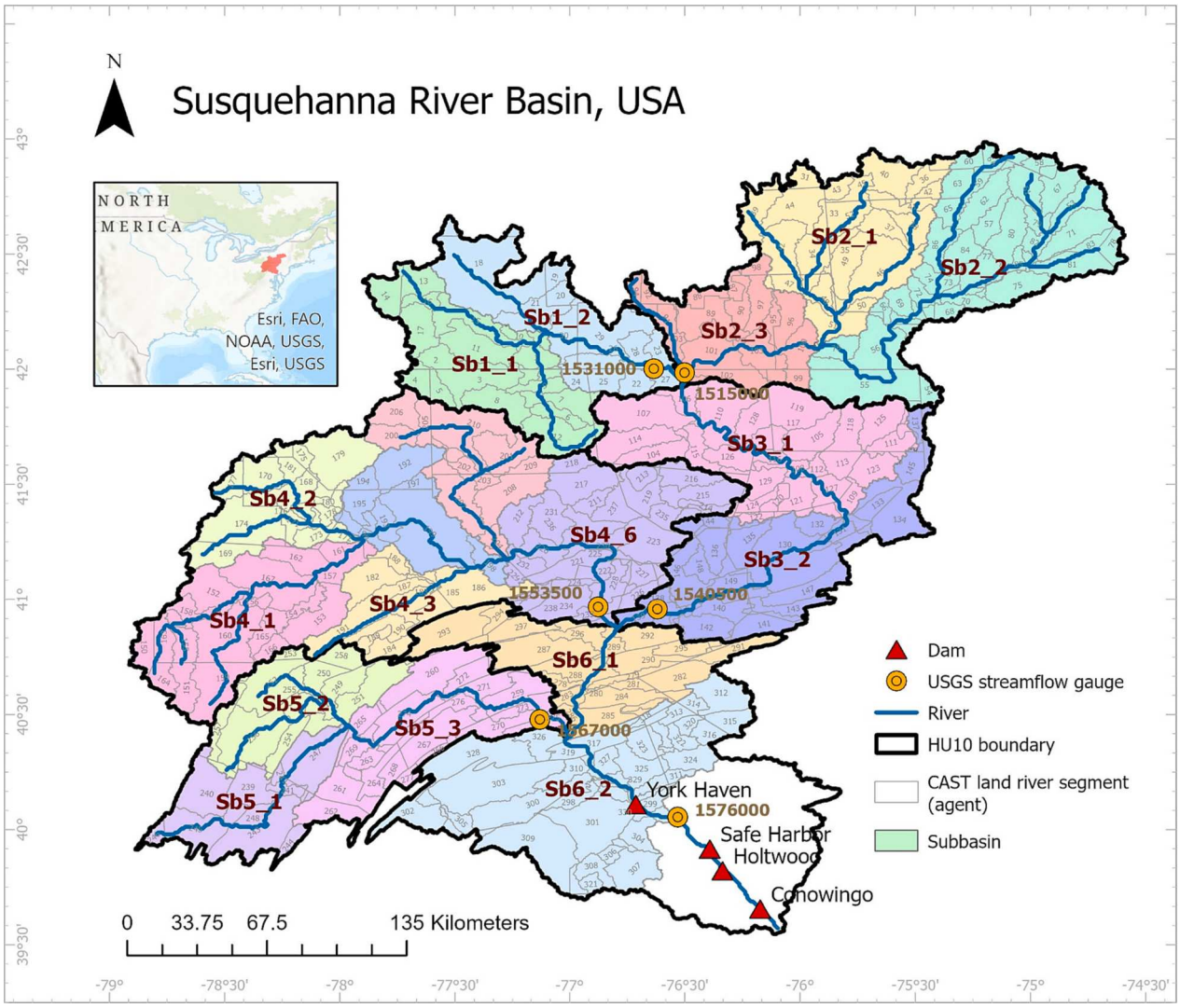



**Fig. 1.** The Susquehanna River Basin, US. Grey boundaries are CAST land and river segments defining the spatial unit of agents. Color patches are subbasins. Points of USGS streamflow gauges are presented with yellow circle symbols. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

"Pay-for-practice" is the most prevalent incentive policy adopted in the US (Campbell et al., 2021; da Costa et al., 2012; Evans and Skaggs, 2004; Radcliffe, 2001), where the funding agencies cost share partial or full BMP implementation cost based on ranking formula. The level of support depends on the rank. Each subsidy payment is generally associated with a pre-approval BMP contract listing eligible BMPs, commitment implementation length, and cost share amount (e.g., *Virginia BMP Cost-Share Program | Fauquier County,* VA, 2020). However, studies have identified some causes that limit the pay-for-practice policy's efficiency and efficacy. Those causes include but are not limited to functional and geographical mismatch to the targeted pollutants (Talberth et al., 2015), inconsistent maintenance over time (Reimer et al., 2012), and low motivations of community involvement (Collins and Maille, 2011; Graversgaard et al., 2021; Maille and Collins, 2012).

"Pay-for-performance" (PFP) provides an alternative subsidy allocation platform that tends to be more cost-efficient (Claassen and Weinberg, 2006; Fales et al., 2016). Instead of paying for implementing BMPs, PFP pays for the actual reduction of targeted pollutants. For example, Talberth et al. (2015) used an optimization model to prove that PFP can achieve the same nutrient reduction as the conventional pay-for-practice with less than half of the policy cost. However, measuring the actual pollutant reduction of each cropland is costly and challenging (Engel, 2016). As a result, field-scale baseline models are usually established to determine the subsidy payment for the individual-based PFP ($PFP_i$; Muenich, 2017). As an alternative, Maille and Collins (2012) conducted a field experiment to test the effectiveness of group-based PFP ($PFP_g$), where the pollutant reduction is evaluated based on the gauge measurement at a watershed outlet. During the experiment, they observed the importance of teamwork among farmers to maximize the subsidy gain.

However, field experiments are generally small in scale due to high costs and intensive labor requirements. It is also unclear the validity of projecting field-scale results to a basin-scale that is often associated with greater heterogeneities in land use/cover, hydrological responses, cultures, and society. Studies have shown that cropland characteristics, farmers' demographics, and social interactions like the social norm (one's decision correlates to some degree with neighbor decisions; Bicchieri and Muldoon, 2011) could influence individuals' BMP adoption decisions in addition to the financial concerns (Liu et al., 2018; Prokopy et al., 2008; Rodriguez et al., 2009). For PFP, dynamic environmental feedback plays another critical driving factor since the pollutants' reduction performance may directly affect farmers' BMP adoption decisions. These create a need for a policy evaluation tool that can capture the system's nonlinearity and heterogeneity, as well as the feedback

dynamics between natural and human systems, to quantify and compare the effectiveness of different incentive policies systematically.

Agent-based modeling (ABM), a bottom-up modeling approach, has been widely adopted to capture heterogeneous human behaviors, including the interactions among agents (e.g., farmers) and the environment in many water resources management problems (Berglund, 2015; Lin et al., 2022a,b). However, few studies have applied ABM as an incentive policy evaluation tool for water quality management. Liu and Ruebeck (2020) demonstrated the effectiveness of generic performance-based environmental payment programs across different communities' compositions through a grid-based ABM model. However, the environmental feedback is represented by a simple embedded equation. We would like to further elaborate the work to couple an ABM (human system) with a process-based sediment simulation model (natural system).

Overall, this study aims to develop an incentive policy evaluation tool to systematically compare the effectiveness of pay-for-practice and PFP at a basin scale, considering dynamic farmers' BMP adoption decisions through a coupled ABM. We focus on two key driving factors affecting farmers' adoption decisions, the social norm and the potential bonus gain. Their sensitivity to the policy effectiveness and corresponding model output uncertainties will be analyzed. Without losing the generality, we demonstrate the proposed tool on the suspended sediment management issue in the Susquehanna River Basin, US, the largest tributary watershed to the Chesapeake Bay. This work will also pave the way for more comprehensive research involving social studies to inform real-world policy better.

The structure of the rest of the paper begins with the introduction of the study area and incentive policies for comparison in Section 2. The proposed policy evaluation tool is shown in Section 3. Then, the results will be presented in Section 4, followed by a Discussion in Section 5. Finally, the Conclusions are shown in Section 6. A lookup table for acronyms and notations (Table A.1) is provided in the Appendix.

## 2. Materials

### 2.1. Susquehanna River Basin

Susquehanna River Basin (SRB) is the largest tributary watershed associated with the Chesapeake Bay. SRB has a drainage area of 71,000 $km^2$ and contributes to more than half of the inflow of the Bay (Schubel and Pritchard, 1986). Given substantial increasing anthropogenic activities and the use of commercial fertilizers, eutrophication in the Chesapeake Bay was observed around the mid-1930s. The eutrophication-induced hypoxia zone impairs the ecosystem (Diaz and Rosenberg, 2008; Kemp et al., 2005). Seventy-three percent of the total P load and 18% of the total N load transported to the Bay are typically attached to the sediment particles (Noe et al., 2020). In 2010, US EPA reached a mutual agreement with seven states and other local authorities to set a TMDL target. The goal is to reduce 20% of the total suspended sediment (TSS) load in 2010 by 2025 (EPA, 2010). Although millions of dollars are invested in the BMP cost-share program each year, we only achieved a 12% TSS reduction in 2021 (12 years after the establishment of the TMDL target), indicating that we are likely to miss the 20% reduction target (EPA, 2022). Such results might be partially attributed to the loss of sediment trapping capability (i.e., the loss of the Conowingo effect; Palinkas et al., 2019) from the Conowingo reservoir. However, inadequate BMP adoptions and lower BMP effectiveness of the most prevalent incentive policy, pay-for-practice, are argued to be another reason for missing the TMDL target. Focusing on evaluating the effectiveness of different incentive policies on a basin scale, we select US Geological Survey (USGS) gauge ID 1576000 as the basin outlet (Fig. 1). Namely, we only model a subarea of SRB to avoid the Conowingo effect or other influences from the reservoir operations in this study. York Heaven Dam is small in capacity and has a limited impact on TSS transportation.

**Table 1**
Summary of BAU, $PFP_i$, and $PFP_g$.

| Incentive policy | Pay-for-practice | Pay-for-performance | |
|---|---|---|---|
| | | Individual-based | Group-based |
| Notation | BAU | $PFP_i$ | $PFP_g$ |
| Cost-share % | 75% | 50% | 50% |
| Social norm | The adoption rate of adjacent neighbors | The adoption rate of adjacent neighbors | The adoption rate of neighbors in the same agent group |
| Bonus | × | ✓ | ✓ |

To facilitate the watershed implementation plan of each jurisdiction to achieve the TMDL target, the Chesapeake Assessment Scenario Tool (CAST; Chesapeake Bay Program, 2020), an online nutrient and sediment load estimator, was developed in 2011 (Devereux and Rigelman, 2014). Unlike the Chesapeake Bay Program's Watershed Model–Hydrological Simulation Program–Fortran (Watershed Model; Moyer and Bennett, 2007) used by US EPA to determine the TMDL, CAST is closer to an accounting tool aiming to be used by stakeholders. To pave the way for future integration between CAST and our proposed policy evaluation tool, the same land and river segments used by CAST are adopted to represent the spatial units of farmer agents (grey boundaries in Fig. 1). We called the spatial boundary of such land and river segments "the CAST boundary unit" thereafter. Although the choice of the CAST boundary unit is a subjective decision of the study, the CAST boundary unit captures the administrative and hydrological boundaries as the segments are defined based on county and watershed boundaries (Chesapeake Bay Program, 2020). Another reason for choosing the CAST boundary unit is due to the limitation of finer spatial resolution data, as the reported BMP data are often collected at the county and downscaled to land and river segments. There are 580 land and river segments in the study area and we only select 330 segments (labeled in Fig. 1; Supplementary Information of an Excel worksheet) with an area larger than 5000 ha as our decision-making agents, who make annual BMP adoption decisions, including the number of the installing unit. Eighteen subbasins (color patches in Fig. 1) are defined as agent groups, where incentive program participants in the same group are expected to collaborate under the group-based policy (e.g., $PFP_g$). Although various BMPs are available in real-world implementations (Chesapeake Bay Program, 2018), we only consider conservation tillage with a five-year contract length in this proof-of-concept study. The BMP capacity for each agent is constrained by its cropland area calculated from USGS National Land Cover Database (Dewitz and USGS, 2021). The corresponding annualized implementing cost is set to be $56.16/ha (in 2012 dollars; Talberth et al., 2015). Streamflow and TSS loads of six USGS gauges (Fig. 1) are adopted to calibrate the sediment simulation models (Section 2.1). The historical daily temperature and precipitation data for each subbasin are estimated from Daymet V4 (Thornton et al., 2022). The Daymet dataset covers the period of 1980 to 2021 for the SRB.

### 2.2. Incentive policies

In this study, we compare individual-based and group-based PFP (i.e., $PFP_i$ and $PFP_g$) with the conventional pay-for-practice policy denoted as a business-as-usual (BAU) scenario. With BAU, government and authorities will cost share 75% of the BMP implementation cost (Talberth et al., 2015). Although the actual cost-share percentage varies in the real world, we assume 75% cost-share amount can already complement the adoption hesitation of the potential loss from a rational perspective. Namely, the remaining hesitations are attributed to their perceived bias. We will consider the social norm as an additional driving factor, where adjacent neighbors' BMP adoption rate will alleviate a farmer's adoption concern. In $PFP_i$, farmers will receive less cost share (i.e., 50%) while an additional bonus based on the actual TSS reduction is available.

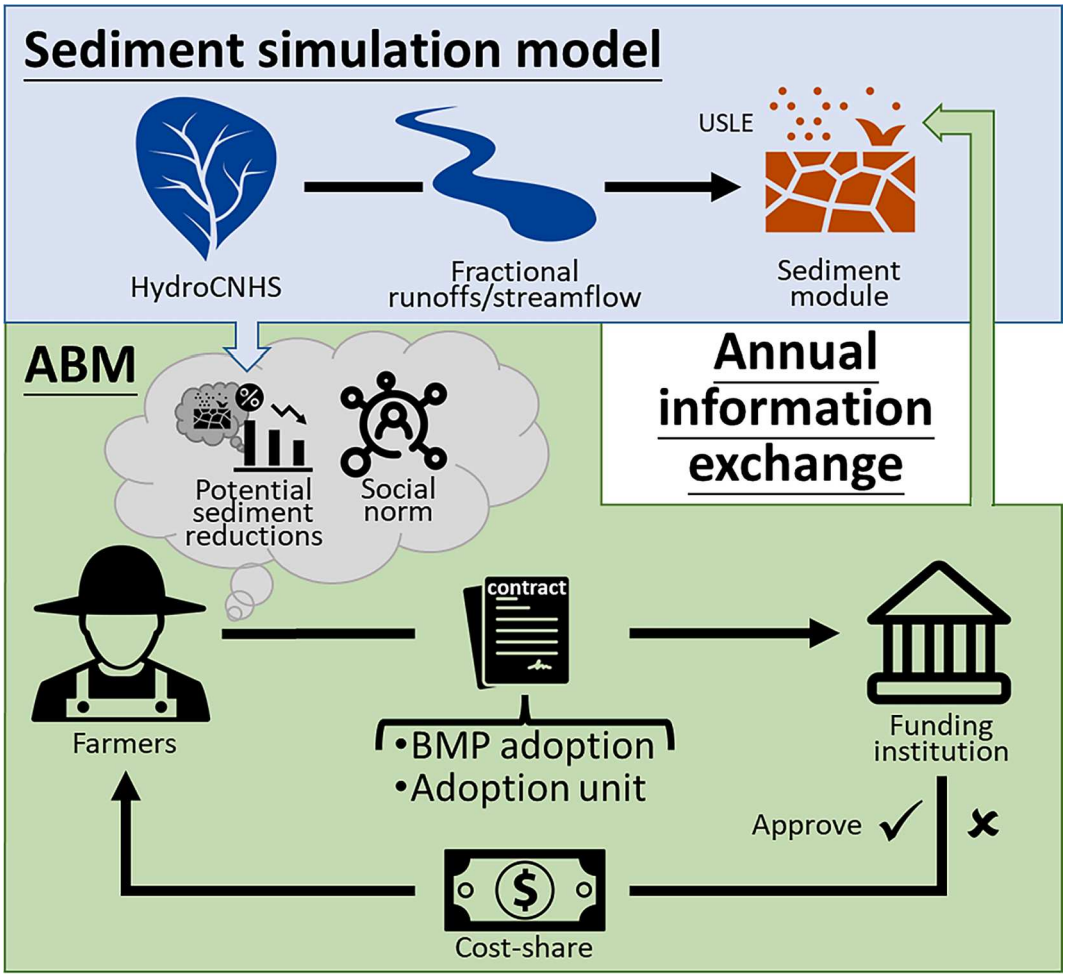



**Fig. 2.** Coupled ABM incentive policy evaluation simulation schema. Farmers make the BMP adoption decision based on the social norm and potential sediment reduction estimated from the sediment simulation model. If the contract is approved by the funding agency, farmers will receive cost-share, and the effectiveness of BMP in total suspended sediment removal will be added to the sediment module for next year's simulation to complete the feedback loop.

Therefore, potential bonus gain plays a role in farmers' decisions in addition to the social norm. $PFP_g$ has a similar setup; however, the social norm is defined as the BMP adoption rate of the agents within the same agent group (i.e., a subbasin) since the bonus gain is shared among the program participants in a group. We summarize the three incentive policies in Table 1.

## 3. Methods

We propose a coupled ABM incentive policy evaluation tool shown in Fig. 2. The two-way coupling technique is used to establish mutual information exchange between natural and human systems. The natural system, simulated by a process-based sediment simulation model, serves as an external environment for ABM. ABM describes agents' annual BMP adoption decisions, including the amount of implementing units. The BMP contracts will be sent to the funding institution for approval and receive a cost share. For simplicity, the subsidy is on first come, first serve basis instead of a ranking system. Collectively, we can observe the TSS reduction compared with a baseline model (i.e., the same model but without any new BMPs installed within the simulation period) at the basin outlet (i.e., USGS gauge ID 1576000) to determine the effectiveness of different incentive policies. We will introduce each component shown in Fig. 2 in the following sections.

### 3.1. Sediment simulation model

The sediment simulation model used in this study is built upon the Hydrological model for Coupled Natural-Human Systems (HydroCNHS; Lin et al., 2022a,b) by adding a sediment module.

#### 3.1.1. Hydrological Model - HydroCNHS

HydroCNHS is an open-source Python package (Lin et al., 2022a,b) supporting the integration of customized human models into a semi-distributed hydrological model through four application programming interfaces (APIs). For example, Dam API, RiverDiv API, Conveying API, and InSitu API can integrate abstracted human decisions of man-made infrastructures such as reservoirs, off-stream diversions, transbasin aqueducts, and drainage systems programmed with the ABM concept. The daily hydrological responses (e.g., streamflow) are computed by routing the runoffs simulated in each subbasin to the routing outlet in HydroCNHS. Each of the HydroCNHS APIs has a unique plug-in structure that respects within-subbasin and inter-subbasin (i.e., river) routing logic under the Lohmann routing schema (Lohmann et al., 1998) to maintain the water balance. Runoffs of each subbasin are calculated by lump models like the General Water Loading Function (GWLF; Haith and Shoemaker, 1987) or the ABCD model (Thomas, 1981). Additionally, HydroCNHS supports model calibration using parallel computing power. HydroCNHS has been applied to characterize modeling uncertainty in coupled natural-human systems (Lin and Yang, 2022).

In this study, runoffs of 18 subbasins are first calculated by the

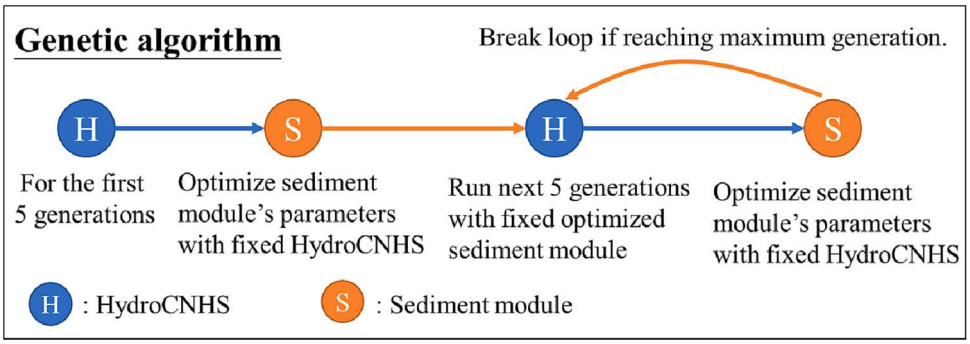



**Fig. 3.** Alternating GA calibration strategy for the sediment simulation model.

GWLF. The nine parameters of GWLF in each subbasin control the process of surface quick flow calculated by the SCS-curve-number method, subsurface flow, and baseflow. After that, runoffs are routed to six routing outlets (i.e., six USGS gauges in Fig. 1). In total, 232 parameters need to be calibrated for the hydrological model built by HydroCNHS.

### 3.1.2. Sediment module

The sediment module describes the sediment generation, delivery, and transportation processes on top of HydroCNHS. We adopt the commonly used Universal Soil Loss Equation (USLE) to compute the amount of sediment generated from the land of an agent (i.e., CAST boundary unit). USLE empirically determines soil erosion by multiplying factors accounting for rainfall erosivity (*RE*), soil erodibility (*K*), surface condition (*CP*), and topographic characteristics (*LS*), as shown in Eq. (1).

$$x_{k,t} = 0.132 \times RE_t \times K_k \times CP_k \times LS_k \times Ar_k \tag{1}$$

where $x_{k,t}$ [Mg/day] is the generated sediment from agent $k$ at day $t$. Factors $K$ and $LS$ are estimated based on the soil data and the digital elevation model data (Moore and Burch, 1986), respectively. Factor $CP$ remains to be calibrated. Term $Ar_k$ [ha] is the area associated with agent $k$. Rainfall erosivity (i.e., $RE$) is calculated by the following equation.

$$RE_t = 64.6 \times s_a^{'} \times R_t^{s_b} \tag{2}$$

where $s_a^{'}$ is $s_a \times A_c$ in cool season (Oct-Mar) and $s_a \times A_w$ in warm season (Apr-Sep). Factors $A_c$ and $A_w$ are equal to 0.12 and 0.3 for this study area (Selker et al., 1990), respectively. Parameters $s_a$ and $s_b$ are required to be calibrated.

After $x$ is computed, we calculate sediment supply to the subbasin outlets by adopting the delivery ratio approximation (Haith et al., 1992), as shown below.

$$sx_{s,m} = DR_s \times \sum_{k \in s} \sum_{t=1}^{d_m} x_{k,t} \tag{3}$$

where $sx_{s,m}$ [Mg/month] is the sediment supply to the outlet of subbasin $s$ in month $m$. Agents within a subbasin share the same delivery ratio ($DR$). Term $d_m$ is the number of days in month $m$.

After that, we route the sediment transportation with the following equation based on the fractional runoffs and streamflow from the upstream outlets that contribute to the streamflow of the downstream outlets (Eq. (4)).

$$sed_{ro,m} = \sum_{u \in U_{ro}} \sum_{i=1}^{12} qr_{u,i}^{m-i-1} \times sx_{u,m-i-1} \tag{4}$$

where $sed_{ro,m}$ [Mg/month] is the TSS load at routing outlet $ro$ in month $m$. The set of $ro$'s upstream outlets (i.e., streamflow contributors) is denoted $U_{ro}$. Term $qr^m$ is an allocation ratio vector, which is dynamically calculated based on the fractional runoffs/streamflow ($fq$ [$m^3$/s]) with 12 months of moving windows, as shown in Eq. (5).

$$qr^m = [fq_{u,m}^{s_q}, fq_{u,m+1}^{s_q}, \cdots, fq_{u,m+10}^{s_q}, fq_{u,m+11}^{s_q}] \tag{5}$$

where $s_q$ is a calibrated parameter. Note that this sediment routing process requires the hydrological model to run 12 months ahead of the sediment simulation module. In total, there are 90 parameters, 5 per subbasin, which need calibration.

### 3.1.3. Model calibration

We sequentially calibrate the sediment simulation model from the upstream routing outlets to the downstream with an alternating genetic algorithm (GA) strategy. Such a way is more efficient in calibrating the large number of parameters based on our modeling experience. In GA, we use roulette wheel selection, uniform crossover (crossover probability = 0.5), elite strategy (one elite), and uniform mutation (mutation probability = 0.15). The population size is 200, with a maximum generation equal to 100. The population is initialized by Latin hypercube sampling. Also, we adopt three random seeds (i.e., 3, 5, and 11). Under the GA framework, we optimize the sediment module's parameters every five generations while calibrating the HydroCNHS model's parameters. Details are shown in Fig. 3.

The objective (*Obj*; Eq. (6)) for the GA is the mean Kling-Gupta efficiency (KGE; Eq. (7)) over the monthly streamflow ($KGE_{Q_M}$) and TSS load ($KGE_{Sed_M}$).

$$Obj = (KGE_{Q_M} + KGE_{Sed_M})/2 \tag{6}$$

$$KGE = 1 - \sqrt{(r-1)^2 + \left(\frac{\sigma_{sim}}{\sigma_{obs}} - 1\right)^2 + \left(\frac{\mu_{sim}}{\mu_{obs}} - 1\right)^2} \tag{7}$$

where $r$ is the Pearson correlation coefficient, and $\mu$ and $\sigma$ denote the mean and standard deviation of flows, respectively. The subscripts *obs* and *sim* refer to observed and simulated streamflow time series, respectively. The streamflow and TSS load observations from the six USGS gauges (Fig. 1) are used for calibration and validation. We apply the 1985–2011 and 2000–2019 observed streamflow to calibration and validation, respectively. However, due to the limited TSS load data (only available from 2000 to 2019), there is no validation for the sediment module. We optimize the sediment-related parameters with the L-BFGS-B algorithm (Byrd et al., 1995; Zhu et al., 1997) given a fixed-parameter HydroCNHS. In addition, we did not explicitly consider the historical BMP adoption dynamics in the calibration process, which means our calibrated model implicitly captures the average BMP effects from 2000 to 2019 through the observed TSS load. As the purpose of the study is to demonstrate the policy evaluation tool and the sediment reduction is calculated against the baseline model, the impact of such simplification can be ignored. Calibration and validation data are summarized in Table S1. Calibration parameters and their bounds are provided in Table S2.

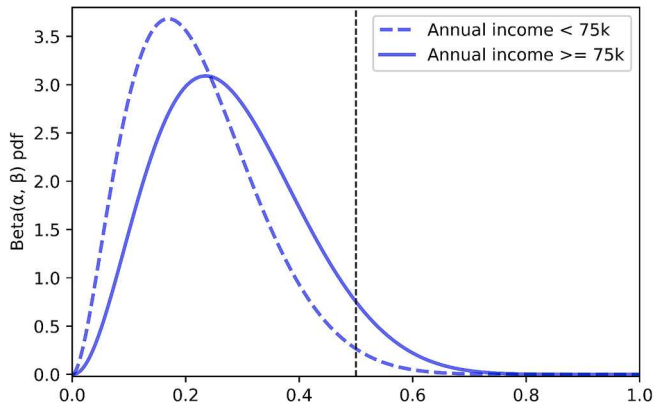



**Fig. 4.** The estimated beta distributions of conservation tillage adoption willingness for agents with an annual income less than $75 k and greater or equal to $75 k.

### 3.2. Agent-based model

An agent-based model is built to capture heterogeneous and nonlinear human behaviors under different incentive policies with environmental feedback. We design two agent types, including farmer agents (i.e., agents) with adaptive decision-making rules, and the funding institution served as a passive respondent to farmers' subsidy requests. The ODD + D description (Müller et al., 2013) for the ABM is shown in Table S3.

#### 3.2.1. Farmer agents

Farmer agents make annual BMP adoption decisions based on an adaptive decision-making rule affected by the social norm and the potential bonus gain (Table 1). An agent's willingness to adopt BMPs is described by a beta distribution with shape ($\alpha$) and scale ($\beta$) parameters. Parameters $\alpha$ and $\beta$ are estimated from 2017 Stewardship Index Survey at Chesapeake Bay (https://www.chesapeakebehaviorchange.org/survey-data) using the moment method, as shown below.

$$\alpha = \left(\frac{\mu(1-\mu)}{\sigma^2} - 1\right) \times \mu \tag{8}$$

$$\beta = \left(\frac{\mu(1-\mu)}{\sigma^2} - 1\right) \times (1-\mu) \tag{9}$$

where $\mu$ and $\sigma^2$ are the mean and variance calculated from the survey responses. We assume higher incomes are associated with larger croplands. To that, we group 330 agents into two equal size clusters according to their cropland areas. Then, we infer two sets of beta distribution's parameters to represent adoption willingness of annual income less than $75 k ($\alpha = 3.254$, $\beta = 8.325$) and greater or equal to $75 k ($\alpha = 2.872$, $\beta = 10.175$), as shown in Fig. 4. This setup aligns with the findings that farmers with larger farms are more likely to adopt BMP (Liu et al., 2018; Prokopy et al., 2008). A detailed parameter estimation process is provided in Text S1.

Given the willingness likelihood depicted by a beta distribution, the actual action of BMP adoption is determined by an adaptive threshold ($C$), as shown below.

$$\begin{cases} rn \geq C, adopt \\ rn < C, don^{'}tadopt \end{cases} \tag{10}$$

where $rn$ is the random number sampled from the assigned beta distribution. The adaptive threshold is updated every year before making the adoption decision (Eqs. (11) and (12)).

$$C_y^{'} = \mathrm{C}_{ini} - (a \times S_y + b \times PB_y) \tag{11}$$

$$C_y = \min\{1, \max\{0, C_y^{'}\}\} \tag{12}$$

where initial $C$ value ($\mathrm{C}_{ini}$) is set to 0.5. Subscript $y$ is the year. The adaptive threshold $C$ is bounded to be in [0, 1] to fit the range of beta distribution. Parameters $a$ and $b$ are the weights of two selected driving factors. $S$ is the social norm, BMP adoption rate of neighbors (Table 1). Potential bonus gain ($PB_y$ [$]) is only available under $\mathrm{PFP_i}$ and $\mathrm{PFP_g}$, and it is calculated by Eq. (13).

$$PB_y = ECDF_y(PR_y) - thr \tag{13}$$

where ECDF stands for empirical cumulative distribution function, which is annually established by the potential sediment reductions of agents that still have BMP capacity. The potential sediment reductions ($PR$ [Mg]) are calculated by the past five-year-averaged TSS loads using the sediment module results (Section 3.1.2). For $\mathrm{PFP_i}$, $PB$ is estimated by the $PR$ of the agent's own land. Given the assumption that agents will collaborate to maximize the bonus gain by implementing BMPs at the agent's cropland with the highest $PB$, $PR$ is estimated by the highest $PR$ for all agents within an agent group under $\mathrm{PFP_g}$. The quantile threshold ($thr$) is computed in Eq. (14).

$$thr = BC \times GP/BP \tag{14}$$

where $BC$ is the annualized unit BMP cost (i.e., the annualized unit cost of conservation tillage). Cost-share percentage gap between PFP and BAU (i.e., 0.25) is denoted as $GP$. The term $BP$ [$] is the bonus payment per one Mg TSS reduction.

Once an agent decides to adopt BMPs, the implementing area is randomly sampled between 1000 ha and 3000 ha from a uniform distribution to form a 5-year BMP contract that will be sent to the funding institution for approval (Section 3.2.2). Note that the area is capped by the remaining cropland without BMPs (i.e., remaining capacity). The land will be released back to capacity after the 5-year BMP contract ends. We assume all croplands are eligible for implementing conservation tillage at the initial time step.

The generated sediment after implementing BMPs ($x_{k,t}^{BMP}$ [Mg/day]) is computed by Eq. (15).

$$x_{k,t}^{adj} = x_{k,t} \times 0.75 \times CR_{k,t} \times \overline{EFF} \tag{15}$$

$$\overline{EFF} = \frac{\sum_{bmp \in BMPs} EFF_{bmp} \bullet Ar_{bmp}}{\sum_{bmp \in BMPs} Ar_{bmp}} \tag{16}$$

where $CR$ is the BMP coverage rate of the cropland. The factor of 0.75 represents that, on average, 75% of the generated sediment is attributed to croplands (Stenfert Kroese et al., 2020). The area-weighted BMP efficiency is calculated by Eq. (16), where $EFF_{bmp}$ and $Ar_{bmp}$ [ha] are the actual efficiency and the implementing area associated with an active BMP contract, respectively. Given the uncertainty of BMP's performance (Liu et al., 2017), the actual efficiency of the conservation tillage varied between 0 and 1 is sampled from a truncated normal distribution with mean and standard deviation under normal distribution equal to 0.4 (Chesapeake Bay Program, 2018) and 0.1 (Liu et al., 2017), respectively.

#### 3.2.2. Funding institution

The funding institution is a passive respondent to approve or deny farmers' BMP subsidy requests based on the funding availability and the adopted incentive policy. The cost-share percentage is defined in Table 1. With $\mathrm{PFP_i}$ and $\mathrm{PFP_g}$, the bonus payment will be paid first each year before approving new cost-share requests. In this study, the bonus payment per one Mg TSS reduction (i.e., $BP$) is set to be $28.18.

### 3.3. Numerical experiments

The experiment design aims to compare the effectiveness of different

**Table 2**
Monthly and annual calibration and validation results for streamflow and TSS at six USGS gauges.

| USGS gauge ID | Monthly | | Annual | |
|---|---|---|---|---|
| | Streamflow | TSS | Streamflow | TSS |
| 1531000 | (0.821, 0.725) | (0.663, –) | (0.880, 0.808) | (0.656, –) |
| 1515000 | (0.861, 0.821) | (0.631, –) | (0.804, 0.911) | (0.602, –) |
| 1540500 | (0.894, 0.864) | (0.691, –) | (0.939, 0.933) | (0.639, –) |
| 1553500 | (0.885, 0.819) | (0.907, –) | (0.958, 0.850) | (0.797, –) |
| 1567000 | (0.898, 0.863) | (0.812, –) | (0.904, 0.899) | (0.752, –) |
| 1576000 | (0.901, 0.876) | (0.918, –) | (0.947, 0.927) | (0.879, –) |
| Mean | (0.877, 0.828) | (0.770, –) | (0.905, 0.888) | (0.721, –) |

Two values in parentheses are KGE for calibration and validation, respectively. Calibration and validation of streamflow are computed with simulation periods 1985–2011 and 2012–2020, respectively.
Calibration of TSS is computed with simulation period 2000–2019.

incentive policies. To consider the climate uncertainty in our analysis, we adopt a stochastic multi-site weather generator, MulGETS (Chen et al., 2014), to synthesize 100 realization sets of 26-year daily weather time series (i.e., temperature and precipitation) based on 2000 to 2019 historical weather data. This will help us explore the policies' effectiveness under different weather time series with similar statistics (e.g., mean and standard deviation). While the sediment simulation model runs for 26 years, ABM only involves 20 years of simulation. The first five years and the last year's sediment simulation are required to provide the necessary information for ABM. Namely, we will only use 20-year simulation results in the comparison analysis. The simulation of each realization set is repeated ten times to address the stochasticity of the model. Also, we set the funding institution to have infinite funding as we would like to see the policies' effectiveness without financial constraints.

More specifically, we will first compare TSS reduction, BMP adoption rate, and cost-efficiency among BAU, $PFP_i$, and $PFP_g$ with parameters $a$ and $b$ set to be 0.25 and 0.25, respectively. The parameters are chosen to approximate a 12% TSS reduction in 12 years under BAU to mimic the historical sediment control patterns starting from 2010. However, we do not claim the validity of our model in capturing historical BMP adoption patterns, as significant simplifications are made in this proof-of-concept study. After that, we conduct a sensitivity analysis (SA) on parameters $a$ and $b$, the weights of the social norm and the potential bonus gain, respectively, to provide a more comprehensive understanding. We simulate all $a$ and $b$ combinations from 0 to 0.4 with an interval of 0.05. Finally, we explore the impact of bonus payments under $PFP_i$ and $PFP_g$, where bonus adjustment ratios ranging from 0.7 to 1.3 with an interval of 0.15 are used to adjust the original bonus payments ($BP$ = \$28.18).

## 4. Results

### 4.1. Calibration and validation

The calibration and validation results of the sediment simulation model are presented in Table 2 for streamflow and TSS loads at six USGS gauges. The streamflow KGEs perform well on both monthly and annual scales, where the calibration and validation mean KGEs are 0.877 and 0.828 for monthly values and 0.905 and 0.888 for annual values. TSS has slightly lower KGEs than the streamflow, where the monthly and annual mean KGEs are equal to 0.770 and 0.721, respectively. However, given the considerable uncertainty in TSS loads due to, e.g., the legacy effect, we consider the calibrated model performance sufficient for our study.

### 4.2. TSS reduction and cost-efficiency comparison among BAU, $PFP_i$, and $PFP_g$

Fig. 5 shows the comparison of TSS reduction at the basin outlet (Fig. 5a) and the BMP implementation area (Fig. 5b) under BAU (blue), $PFP_i$ (orange), and $PFP_g$ (green). The band is plus and minus one standard deviation. The TSS reduction is referenced from the baseline model without any BMP implementations. The results indicate two PFP policies dominate BAU in TSS reduction. $PFP_g$ has the best TSS reduction performance, and it takes a shorter time to achieve a higher TSS reduction equilibrium. Such outcomes are also revealed in the greater BMP adoptions in the long run, as shown in Fig. 5b. Farmers are more willing to participate in the incentive program under $PFP_g$, which can potentially lead to higher bonus gains with collaboration; hence, a larger area of BMP implementation compared to $PFP_i$. Although a distinct difference between blue and orange lines is found in TSS reduction performance (Fig. 5a), they obtain similar BMP contracts (Fig. 5b) patterns. This is because farmers with higher soil erosion have greater motivation to implement BMPs resulting in better basin-wide TSS reduction compared to BAU. Namely, $PFP_i$ implicitly allocates the incentive program funding to the hotspots through the bonus payment mechanism. We further show the spatial distribution of BMP adoption units at the agent level in Fig. S2.

The TSS reduction equilibriums shown in Fig. 5 imply the TSS reduction upper bounds of the incentive policies. Such information may help the government to analyze the achievability of the reduction target

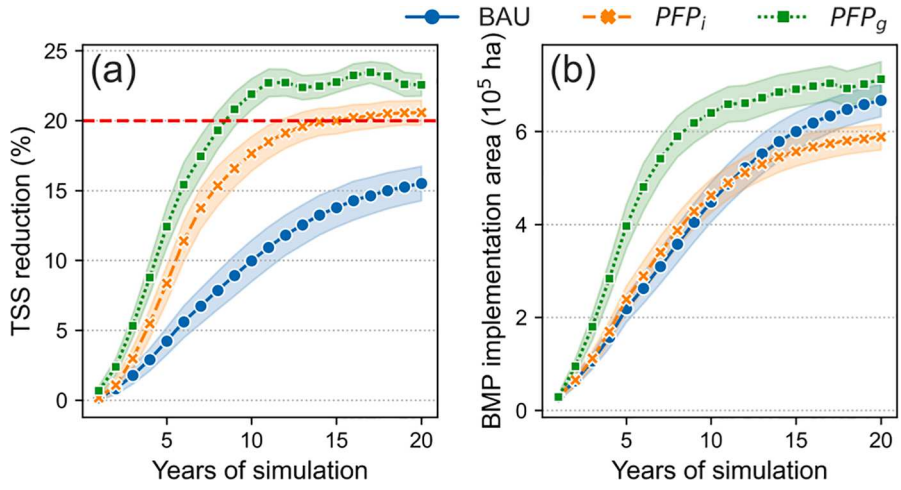



**Fig. 5.** (a) The comparison of TSS reduction at the basin outlet and (b) the BMP implementation area given $a = b = 0.25$ under BAU (blue), $PFP_i$ (orange), and $PFP_g$ (green). The band is plus and minus one standard deviation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

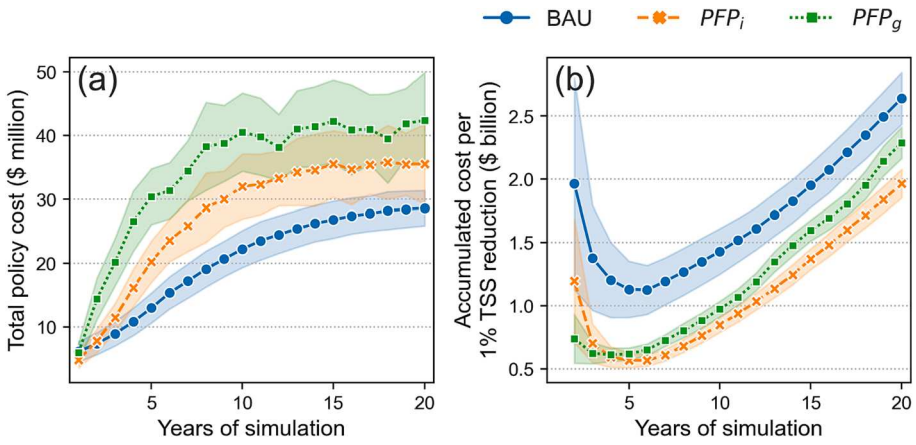



**Fig. 6.** (a) The comparison of the total policy cost in 2012 dollars and (b) the accumulated cost per 1% TSS reduction given $a = b = 0.25$ under BAU (blue), $PFP_i$ (orange), and $PFP_g$ (green). The band is plus and minus one standard deviation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

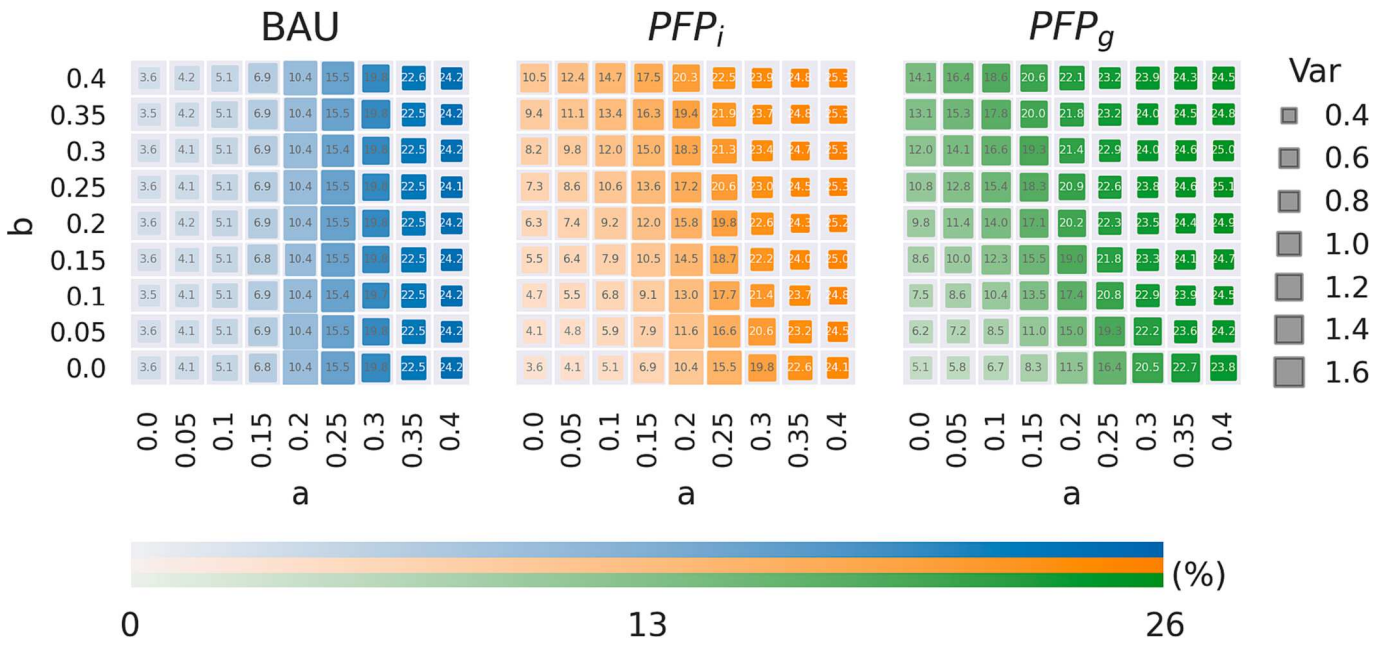



**Fig. 7.** Sensitivity analysis result of the social norm (i.e., *a*) and the potential bonus payment (i.e., *b*) over three incentive policies at year 20. The darker color indicates a higher TSS reduction. The size of the squares represents the uncertainty (i.e., Var) of the results.

with the consideration of human dynamics. For example, the blue line (i.e., BAU) cannot achieve a 20% TSS reduction target (red dashed line) within the 20-year simulation period, while both orange and green lines can reach that target.

From the budget perspective, we seek cost-efficient options to better allocate limited resources. We compare the total policy cost (Fig. 6a) and the accumulated cost per 1% TSS reduction (Fig. 6b). It has the same layout as Fig. 5. In Fig. 6a, $PFP_g$ (green) has the highest policy cost because of more outstanding BMP adoptions (Fig. 5b), followed by $PFP_i$ (orange) and BAU (blue). Although BAU costs the least, it turns out that BAU is not the most cost-inefficient policy, as shown in Fig. 6b. The blue line cost more per 1% TSS reduction compared to the $PFP_i$ and $PFP_g$. Between $PFP_i$ and $PFP_g$, there is a crossover. After the first couple of years, $PFP_g$ becomes more expensive than $PFP_i$ in terms of cost-per-unit TSS reduction. Such a crossover phenomenon is because of the decrease in marginal effect. With more BMPs implemented under $PFP_g$, the newer BMPs will have to be implemented in the less soil erosion cropland (i.e., less effective) as the erosion hotspots have been filled already.

### 4.3. SA of social norm and potential bonus payment

As the quantitative social study is not currently available to determine the weights of social norm and potential bonus payment (i.e., *a* and *b*), we show their SA results of the TSS reduction in Fig. 7 to provide a more comprehensive understanding of the effectiveness of BAU (blue), $PFP_i$ (orange) and $PFP_g$ (green) under different human behavior setups. While the potential bonus payment directly incentivizes farmers' behaviors, social norm plays an important role in determining the diffusion rate of innovations (i.e., BMP adoption). The darker color means greater TSS reduction performance, where the actual reduction percentages are also labeled on the plots. The size of the squares represents the uncertainty across three policies. Results show that PFP outperforms BAU in all cases, and $PFP_g$ generally performs better than $PFP_i$. We can observe that $PFP_g$ has the highest TSS reduction even without the effects of the social norm and the potential bonus payment (i.e., $a = 0$ and $b = 0$). This is because of the collaboration inherent in $PFP_g$.

From the uncertainty perspective, the results with the extreme *a* and

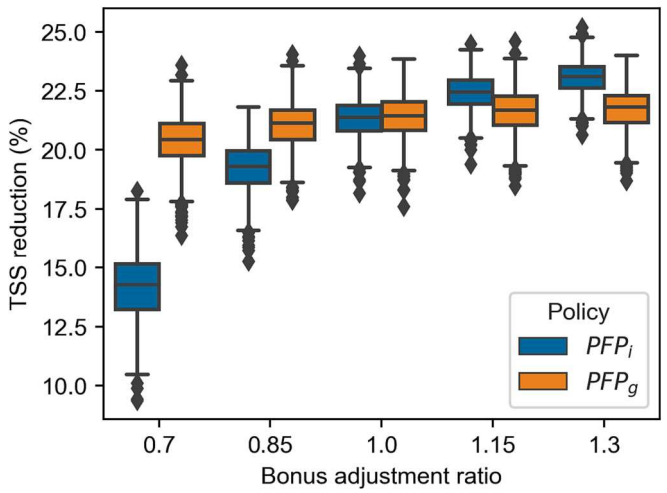



**Fig. 8.** Impacts of the bonus payment on the basin-wide TSS reduction for $PFP_i$ (blue) and $PFP_g$ (orange). Boxplots show the results of year 20 that are simulated with $a = b = 0.25$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

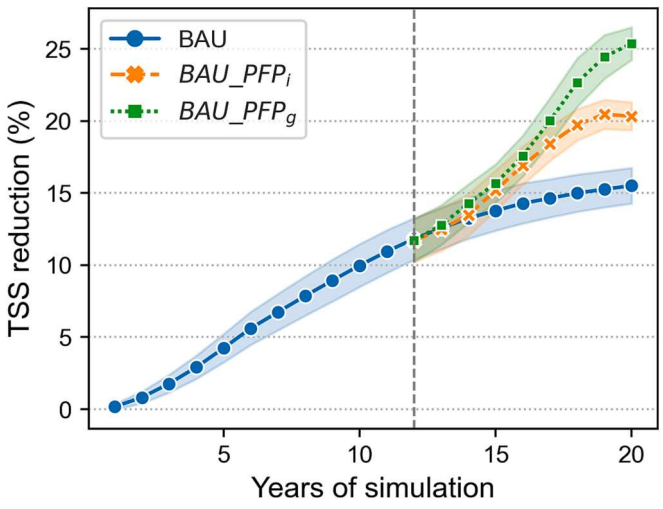



**Fig. 9.** The comparison of TSS reduction at the basin outlet given $a = b = 0.25$ under BAU (blue), switching to $PFP_i$ (orange; $BAU\_PFP_i$) in year 12, and switching to $PFP_g$ (green; $BAU\_PFP_g$) in year 12. The band is plus and minus one standard deviation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

*b* parameter settings (lower left and upper right) tend to be more certain. This is not surprising as we could easily imagine the reduction results if all people refuse or are highly willing to adopt BMPs. However, reality often falls within the parameter range associated with high uncertainty. For example, the chosen parameter setting ($a = b = 0.25$) claimed to approximate 12% TSS reduction at year 12 tends to have relatively high uncertainty. Nonetheless, if we compare the uncertainty across policies, $PFP_g$ has the lowest uncertainty, followed by $PFP_i$ and then BAU. Such results are also indicated in the band in Fig. 5a.

### 4.4. Impact of bonus payments on PFP

From Section 4.2, we show that PFP is generally better in TSS reduction and more cost-efficient given the selected bonus payment per unit TSS reduction (i.e., *BP*). However, different *BP* may lead to a different conclusion. We show the TSS reduction responses with different bonus adjustment ratios for $PFP_i$ (blue) and $PFP_g$ (orange) in Fig. 8. Results support that higher *BP* can lead to better TSS reduction outcomes. While $PFP_i$'s performance dramatically decreases with a lower *BP*, $PFP_g$ reveal a relatively stable performance across different *BP*. When the bonus payment is equal to $0.7BP$, the $PFP_i$'s TSS reduction performance can drop below 15%, which is lower than BAU (15.5%). This implies that if the bonus payment cannot complement the cost-share percentage gap (i.e., *GP*), the $PFP_i$'s performance might be worse than the conventional BAU. In the case of $PFP_g$, the change of *BP* has less impact to the basin-wide TSS reduction outcomes since the BMPs are consistently implemented from the land with the highest soil erosion.

## 5. Discussion

### 5.1. Is it too late to switch incentive policies?

According to the results, we identify the upper bound of BAU which is lower than the 20% TSS reduction target. At the same time, $PFP_i$ and $PFP_g$ indicate the potential to achieve the target within the 20-year simulation period. Such results imply the necessity to switch to PFP if we intend to reach the 20% TSS reduction target. Therefore, the question becomes whether it is too late to make such changes. Fig. 9 shows we might still miss the target in year 15 but achieve the 20% TSS reduction within the last five years of the 20-year simulation, given the incentive policy is switched from BAU to PFP in year 12. Such findings could be analogous to Chesapeake Bay's current situation if we viewed year 1 of our simulation as 2010 when the TMDL target was developed. In year 12 (2021), we achieve a 12% TSS reduction under the pay-for-practice policy (i.e., BAU). The blue line indicates the BAU will likely miss the target in 2025 (year 15), which aligns with the recent-published EPA report (EPA, 2022). In the report, EPA points out that additional efforts and resources must be carried out to meet the target. Some jurisdictions (e.g., Maryland) require its fund flowing to the SRB commission must be used on a PFP basis (Blankenship, 2022). Although such a measure might not lead to the target achievement in the near term (year 15), it will be beneficial from a long-term perspective, as supported by our results (Fig. 9).

### 5.2. Limitations

Although our numerical experiment advocates PFP as a more effective and cost-efficient incentive policy, this study does not address the gap between the modeling result and the actual PFP implementation. For example, how is the bonus payment determined, how is the actual BMP TSS removal quantified, and what spatial scale should be used for $PFP_g$ (Fleming et al., 2022)? In addition, PFP has an inevitable drawback of equity problem (Ribaudo et al., 2011) as the resources implicitly flow to the targeted regions (e.g., sediment generation hotspots). Such an equity issue requires future study to address.

Moreover, we did not consider indirect influences on farmers' BMP adoption decisions from other programs like crop insurance subsidies and other financial supports, which may be worth exploring. Also, we only consider the relatively inexpensive BMP option, conservation tillage, and ignore administrative costs in this study, which might significantly underestimate the total funds needed for conducting those proposed incentive policies. However, information like the total policy cost in Fig. 5 may provide a minimum fund requirement. Also, BMP options like grass/forest buffer strips lasting for multiple years may receive different farmers' perceptions compared to the annual BMPs, such as the conservation tillage. For example, we will need to consider the sustained adoption of practice over time to maintain the conservation gain for the long-lasting BMP options. Multiple beta distributions can be created for each type of BMP to address the perception difference with the assistance of social studies.

Social studies (e.g., surveys and behavioral experiments) may also help parameterize and calibrate/validate the ABM model (e.g., *a* and *b*;

Schrieks et al., 2021). For example, we did not explicitly consider the dynamic of farmers' past experience in the BMP adoption decision due to the lack of empirical social data. To that, the results of social studies can be used to identify other socio-psychological determinants (e.g., farmer identity, different farmer attitudes, past behavior, awareness, information sources, institutional support, barriers to adoption) and theories (e.g., theory of planned behavior, diffusion of innovation, and value-belief-norm) to strengthen the model's representative.

Lastly, without the support of social network and administrative network data, the connections among agents are defined based on the chosen spatial scale of the CAST boundary unit. However, studies have pointed out the spatial scale is a critical factor in water governance (Bitterman et al., 2023; Bitterman and Koliba, 2020). A comprehensive design of a sensitivity analysis on agents' spatial scale could greatly improve our understanding of coupled ABM's modeling property. A future study is encouraged.

## 6. Conclusions

Incentive programs are often adopted to motivate voluntary compliance in implementing BMPs to address surplus nonpoint nutrient and sediment inputs. Although PFP is indicated to be a more cost-efficient alternative than the conventional BAU (i.e., pay-for-practice) through field experiments, such results are usually limited in scale. This study proposes a coupled ABM incentive policy evaluation tool to systematically compare the effectiveness of different incentive policies, including BAU (75% cost share), the individual-based PFP (i.e., $PFP_i$; 50% cost share plus bonus), and the group-based PFP (i.e., $PFP_g$; 50% cost share plus bonus) at a basin scale. The coupled ABM is established by the two-way coupling of a sediment simulation model and an ABM that describes heterogeneous farmers' BMP adoption behaviors and a passive respondent of the funding institution. The Susquehanna River Basin, Chesapeake's largest tributary watershed, is selected as the study area to demonstrate the proposed policy evaluation tool.

Modeling results advocate that PFP can achieve higher TSS reduction equilibrium in a shorter time and is more cost-efficient than the BAU. To that end, $PFP_g$ performs better than $PFP_i$ because of the collaboration among agents in selecting the best BMP-implementing locations. Such collaborative behaviors in maximizing bonus gains also lead to a more stable TSS reduction performance against the change in bonus payment, where $PFP_i$ shows a clear decreasing performance trend as the bonus payment reduces. The SA results on the weights of the social norm (i.e., $a$) and the potential bonus payment (i.e., $b$) provide more comprehensive information about the effectiveness of three incentive policies and indicate a more realistic ABM parameterization is generally associated with larger uncertainty. We analogize our findings to the current water quality management situation in the Chesapeake Bay and suggest that switching from BAU to PFP could be a more promising way to achieve the TMDL target (i.e., 20% TSS reduction) in the long term. Finally, we encourage future studies to elaborate the proposed tool by incorporating social studies (e.g., surveys and behavioral experiments) to improve farmers' decision-making process representation in the model. This includes additional socio-psychological determinants identification, model structural design, the influence of communication at different scales, and the impact of bureaucracy on incentive program participation.

## CRediT authorship contribution statement

**Chung-Yi Lin:** Conceptualization, Methodology, Software, Formal analysis, Writing – original draft. **Y.C. Ethan Yang:** Conceptualization, Writing – review & editing, Supervision, Funding acquisition, Investigation, Project administration. **Anil Kumar Chaudhary:** Conceptualization, Writing – review & editing.

**Table A.1**
Lookup table for acronyms.

| Acronym/notation | Description |
|---|---|
| ABM | Agent-based modeling |
| API | Application programming interface |
| BAU | Business-as-usual |
| BMPs | Best management practices |
| CAST | Chesapeake Assessment Scenario Tool |
| ECDF | An empirical cumulative distribution function |
| EPA | Environmental protection agency |
| GA | Genetic algorithm |
| HydroCNHS | Hydrological model for Coupled Natural-Human Systems |
| KGE | Kling-Gupta efficiency |
| PFP | Pay-by-performance |
| $PFP_i$ | Individual-based PFP |
| $PFP_g$ | Group-based PFP |
| SA | Sensitivity analysis |
| SRB | Susquehanna River Basin |
| TMDL | Total maximum daily load |
| TSS | Total suspended sediment |
| USGS | US Geological Survey |
| USLE | Universal Soil Loss Equation |

## 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.

## Data availability

Data will be made available on request.

## Acknowledgments

The work described in this paper was supported by the US National Science Foundation (NSF): CBET #1941727. We would like to thank the editor, the associate editor, and anonymous reviewers for their comments and suggestions to improve the quality of the manuscript.

## Appendix A

## Appendix B. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhydrol.2023.129959.

## References

Berglund, E.Z., 2015. Using Agent-Based Modeling for Water Resources Planning and Management. J. Water Resour. Plan. Manag. 141 (11) https://doi.org/10.1061/(ASCE)WR.1943-5452.0000544.

Bicchieri C., Muldoon R. 2011. *Social Norms*. https://plato.stanford.edu/archives/spr2014/entries/social-norms/.

Bitterman, P., Koliba, C.J., 2020. Modeling alternative collaborative governance network designs: an agent-based model of water governance in the Lake Champlain Basin, Vermont. J. Public Adm. Res. Theory 30 (4), 636–655. https://doi.org/10.1093/jopart/muaa013.

Bitterman, P., Koliba, C., Singer, A., 2023. A network perspective on multi-scale water governance in the Lake Champlain Basin, Vermont. Ecol. Soc. 28 (1) https://doi.org/10.5751/ES-14036-280144.

Blankenship K. 2022. New plan in place for pollution problems at Conowingo Dam. Bay J. https://www.bayjournal.com/news/pollution/new-plan-in-place-for-pollution-problems-at-conowingo-dam/article_c9d87088-23de-11ed-845c-8f9a80d8f076.html.

Borah, D.K., Ahmadisharaf, E., Padmanabhan, G., Imen, S., Mohamoud, Y.M., 2019. Watershed models for development and implementation of total maximum daily loads. J. Hydrol. Eng. 24 (1), 03118001. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001724.

Byrd, R.H., Lu, P., Nocedal, J., Zhu, C., 1995. A limited memory algorithm for bound constrained optimization. SIAM J. Sci. Comput. 16 (5), 1190–1208. https://doi.org/10.1137/0916069.
Campbell, K.M., Boyer, C.N., Lambert, D.M., Clark, C.D., Smith, S.A., 2021. Risk, cost-share payments, and adoption of cover crops and no-till. J. Soil Water Conserv. 76 (2), 166–174. https://doi.org/10.2489/jswc.2021.00027.
Carpenter, S.R., Caraco, N.F., Correll, D.L., Howarth, R.W., Sharpley, A.N., Smith, V.H., 1998. Nonpoint pollution of surface waters with phosphorus and nitrogen. Ecol. Appl. 8 (3), 559–568. https://doi.org/10.1890/1051-0761(1998)008[0559:NPOSWW]2.0.CO;2.
Chen, J., Brissette, F.P., Zhang, X.J., 2014. A multi-site stochastic weather generator for daily precipitation and temperature. Trans. ASABE 57 (5), 1375–1391. https://doi.org/10.13031/trans.57.10685.
Chesapeake Bay Program, 2018. Chesapeake Bay Program Quick Reference Guide for Best Management Practices (BMPs): Nonpoint Source BMPs to Reduce Nitrogen, Phosphorus and Sediment Loads to the Chesapeake Bay and its Local Waters. Chesapeake Bay Program.
Chesapeake Bay Program, 2020. Chesapeake Assessment and Scenario Tool (CAST) Version 2019. Chesapeake Bay Program Office, Last accessed [April 2023].
Claassen R. L., Weinberg M. 2006. Rewarding Farm Practices versus Environmental Performance (EB-5). USDA Economic Research Service.
Collins, A.R., Maille, P., 2011. Group decision-making theory and behavior under performance-based water quality payments. Ecol. Econ. 70 (4), 806–812. https://doi.org/10.1016/j.ecolecon.2010.11.020.
Commender, K.E., Munsell, J.F., Ares, A., Jay Sullivan, B., Chamberlain, J.L., 2020. The effects of cost-share participant experience on forest buffer retention. Small-Scale For. 19 (3), 253–273. https://doi.org/10.1007/s11842-020-09435-8.
da Costa, P.M.F., Hu, W., Pagoulatos, A., Schieffer, J., 2012. Participation in government cost-share conservation programs in the Kentucky river watershed: A county-level analysis. Environ. Econ. 3 (1).
Devereux, O., Rigelman, J.R., 2014. CAST: An online tool for facilitating local involvement in watershed implementation plans for the Chesapeake Bay total maximum daily load. J. Water Manage. Model. https://doi.org/10.14796/JWMM.C364.
Dewitz J., USGS. 2021. National Land Cover Database (NLCD) 2019 Products ver. 2.0. https://doi.org/doi:10.5066/P9KZCM54.
Diaz, R.J., Rosenberg, R., 2008. Spreading dead zones and consequences for marine ecosystems. Science 321 (5891), 926–929. https://doi.org/10.1126/science.1156401.
Engel, S., 2016. The devil in the detail: A practical guide on designing payments for environmental services. Int. Rev. Environ. Resour. Econ. 9 (1–2), 131–177. https://doi.org/10.1561/101.00000076.
EPA. 2010. Chesapeake Bay total maximum daily load for nitrogen, phosphorus and sediment.
EPA. (2022). Evaluation of Pennsylvania's Final Amended Phase III Watershed Implementation Plan (WIP). https://www.epa.gov/system/files/documents/2022-11/Evaluation_of_Pennsylvania%27s_FINAL_Amended_Phase_III_WIP_11.15.2022%20%28002%29.pdf.
Evans, R.O., Skaggs, R.W., 2004. Development of Controlled Drainage as a BMP in North Carolina. American Society of Agricultural and Biological Engineers. Drainage VIII. https://doi.org/10.13031/2013.15707.
Fales, M., Dell, R., Herbert, M.E., Sowa, S.P., Asher, J., O'Neil, G., Doran, P.J., Wickerham, B., 2016. Making the leap from science to implementation: Strategic agricultural conservation in Michigan's Saginaw Bay watershed. J. Great Lakes Res. 42 (6), 1372–1385. https://doi.org/10.1016/j.jglr.2016.09.010.
Feather P., Cooper J.C. (Eds.). 1995. Voluntary Incentives for Reducing Agricultural Nonpoint Source Water Pollution. https://doi.org/10.22004/ag.econ.33619.
Fleming, P.M., Stephenson, K., Collick, A.S., Easton, Z.M., 2022. Targeting for nonpoint source pollution reduction: A synthesis of lessons learned, remaining challenges, and emerging opportunities. J. Environ. Manage. 308, 114649 https://doi.org/10.1016/j.jenvman.2022.114649.
Graversgaard, M., Jacobsen, B.H., Hoffmann, C.C., Dalgaard, T., Odgaard, M.V., Kjaergaard, C., Powell, N., Strand, J.A., Feuerbach, P., Tonderski, K., 2021. Policies for wetlands implementation in Denmark and Sweden – historical lessons and emerging issues. Land Use Policy 101, 105206. https://doi.org/10.1016/j.landusepol.2020.105206.
Haith D. A., Mandel R., Wu R. S. 1992. GWLF: Generalized Watershed Loading Functions, Version 2.0, User's Manual. Dept. of Agricultural & Biological Engineering, Cornell University, Ithaca, NY.
Haith, D.A., Shoemaker, L.L., 1987. Generalized watershed loading functions for stream flow nutrients. JAWRA J. Am. Water Resour. Assoc. 23 (3), 471–478. https://doi.org/10.1111/j.1752-1688.1987.tb00825.x.
Hardy, S.D., Koontz, T.M., 2008. Reducing nonpoint source pollution through collaboration: policies and programs across the U.S. States. Environ. Manage. 41 (3), 301–310. https://doi.org/10.1007/s00267-007-9038-6.
Kemp, W.M., Boynton, W.R., Adolf, J.E., Boesch, D.F., Boicourt, W.C., Brush, G., Cornwell, J.C., Fisher, T.R., Glibert, P.M., Hagy, J.D., Harding, L.W., Houde, E.D., Kimmel, D.G., Miller, W.D., Newell, R.I.E., Roman, M.R., Smith, E.M., Stevenson, J. C., 2005. Eutrophication of Chesapeake Bay: Historical trends and ecological interactions. Mar. Ecol. Prog. Ser. 303, 1–29. https://doi.org/10.3354/meps303001.
Li, H.-Y., Tan, Z., Ma, H., Zhu, Z., Abeshu, G.W., Zhu, S., Cohen, S., Zhou, T., Xu, D., Leung, L.R., 2022. A new large-scale suspended sediment model and its application over the United States. Hydrol. Earth Syst. Sci. 26 (3), 665–688. https://doi.org/10.5194/hess-26-665-2022.
Lin, C.-Y., Yang, Y.-C.-E., 2022. The effects of model complexity on model output uncertainty in co-evolved coupled natural-human systems. Earth's Future 10 (6). https://doi.org/10.1029/2021EF002403.
Lin, C.-Y., Yang, Y.C.E., Malek, K., Adam, J.C., 2022a. An investigation of coupled natural human systems using a two-way coupled agent-based modeling framework. Environ. Model. Softw. 155, 105451 https://doi.org/10.1016/j.envsoft.2022.105451.
Lin, C.-Y., Yang, Y.-C.-E., Wi, S., 2022b. HydroCNHS: A Python package of hydrological model for coupled natural-human systems. J. Water Resour. Plan. Manag. 148 (12), 06022005 https://doi.org/10.1061/(ASCE)WR.1943-5452.0001630.
Liu, T., Bruins, R.J.F., Heberling, M.T., 2018. Factors influencing farmers' adoption of best management practices. A review and synthesis. Sustainability 10 (2), 2. https://doi.org/10.3390/su10020432.
Liu, Y., Engel, B.A., Flanagan, D.C., Gitau, M.W., McMillan, S.K., Chaubey, I., 2017. A review on effectiveness of best management practices in improving hydrology and water quality: Needs and opportunities. Elsevier Enhanced Reader. Sci. Total Environ. 601-602, 580–593. https://doi.org/10.1016/j.scitotenv.2017.05.212.
Liu, H., Ruebeck, C.S., 2020. Knowledge spillover and positive environmental externality in agricultural decision making under performance-based payment programs. Agric. Resour. Econ. Rev. 49 (2), 270–290. https://doi.org/10.1017/age.2020.18.
Lohmann, D., Raschke, E., Nijssen, B., Lettenmaier, D.P., 1998. Regional scale hydrology: I. Formulation of the VIC-2L model coupled to a routing model. Hydrol. Sci. J. 43 (1), 131–141. https://doi.org/10.1080/02626669809492107.
Maille, P., Collins, A.R., 2012. An index approach to performance-based payments for water quality. J. Environ. Manage. 99, 27–35. https://doi.org/10.1016/j.jenvman.2012.01.002.
Mee, L., 2006. Reviving dead zones. Sci. Am. 295 (5), 78–85.
Moyer D.L., Bennett M.R. 2007. Development of Relations of Stream Stage to Channel Geometry and Discharge for Stream Segments Simulated with Hydrologic Simulation Program–Fortran (HSPF), Chesapeake Bay Watershed and Adjacent Parts of Virginia, Maryland, and Delaware. USGS. https://pubs.usgs.gov/sir/2007/5135/.
Muenich, 2017. Pay-for-performance conservation using SWAT highlights need for field-level agricultural conservation. Trans. ASABE 60 (6), 1925–1937. https://doi.org/10.13031/trans.12379.
Müller, B., Bohn, F., Dreßler, G., Groeneveld, J., Klassert, C., Martin, R., Schlüter, M., Schulze, J., Weise, H., Schwarz, N., 2013. Describing human decisions in agent-based models – ODD + D, an extension of the ODD protocol. Environ. Model. Softw. 48, 37–48. https://doi.org/10.1016/j.envsoft.2013.06.003.
Noe, G.B., Cashman, M.J., Skalak, K., Gellis, A., Hopkins, K.G., Moyer, D., Webber, J., Benthem, A., Maloney, K., Brakebill, J., Sekellick, A., Langland, M., Zhang, Q., Shenk, G., Keisman, J., Hupp, C., 2020. Sediment dynamics and implications for management: State of the science from long-term research in the Chesapeake Bay watershed, USA. WIREs Water 7 (4), e1454. https://doi.org/10.1002/wat2.1454.
Palinkas, C.M., Testa, J.M., Cornwell, J.C., Li, M., Sanford, L.P., 2019. Influences of a river dam on delivery and fate of sediments and particulate nutrients to the adjacent estuary: Case study of Conowingo Dam and Chesapeake Bay. Estuar. Coasts 42 (8), 2072–2095. https://doi.org/10.1007/s12237-019-00634-x.
Prokopy, L.S., Floress, K., Klotthor-Weinkauf, D., Baumgart-Getz, A., 2008. Determinants of agricultural best management practice adoption: Evidence from the literature. J. Soil Water Conserv. 63 (5), 300–311. https://doi.org/10.2489/jswc.63.5.300.
Radcliffe, D.E., 2001. Agricultural Cost Share Programs in Kentucky and North Carolina. Georgia Water Resources Conference. Georgia Water Resources Conference. University of Georgia https://smartech.gatech.edu/handle/1853/44083.
Reimer, A.P., Thompson, A.W., Prokopy, L.S., 2012. The multi-dimensional nature of environmental attitudes among farmers in Indiana: Implications for conservation adoption. Agric. Hum. Values 29 (1), 29–40. https://doi.org/10.1007/s10460-011-9308-z.
Ribaudo M.O., Shortle J.S., Blandford D., Horan R.D. 2011. Improving the Efficiency and Effectiveness of Agri-environmental Policies for the Chesapeake Bay. Choices, 26(3). https://www.jstor.org/stable/choices.26.3.03.
Rodriguez, J.M., Molnar, J.J., Fazio, R.A., Sydnor, E., Lowe, M.J., 2009. Barriers to adoption of sustainable agriculture practices: Change agent perspectives. Renewable Agric. Food Syst 24 (1), 60–71. https://doi.org/10.1017/S1742170508002421.
Schrieks T., Botzen W.J.W., Wens M., Haer T., Aerts J.C.J.H. 2021. Integrating behavioral theories in agent-based models for agricultural drought risk assessments. Front. Water, 3. https://www.frontiersin.org/articles/10.3389/frwa.2021.686329.
Schubel, J.R., Pritchard, D.W., 1986. Responses of upper Chesapeake Bay to variations in discharge of the Susquehanna River. Estuaries 9 (4), 236–249. https://doi.org/10.2307/1352096.
Selker J.S., Haith D.A., Reynolds J.E. 1990. Calibration and testing of a daily rainfall erosivity model. Trans. ASAE, 33(5), 1612–1617. https://doi.org/doi: 10.13031/2013.31516.
Stenfert Kroese, J., Batista, P.V.G., Jacobs, S.R., Breuer, L., Quinton, J.N., Rufino, M.C., 2020. Agricultural land is the main source of stream sediments after conversion of an African montane forest. Sci. Rep. 10 (1), 1 https://doi.org/10.1038/s41598-020-71924-9.
Talberth, J., Selman, M., Walker, S., Gray, E., 2015. Pay for Performance: Optimizing public investments in agricultural best management practices in the Chesapeake Bay Watershed. Ecol. Econ. 118, 252–261. https://doi.org/10.1016/j.ecolecon.2015.07.033.
Thomas H.A. 1981. Improved Methods for National Water Assessment. Report, Contract: WR15249270. Washington, D.C.: US Water Resource Council.

Thornton, M.M., Shrestha, R., Wei, Y., Thornton, P.E., Kao, S.-C., Wilson, B.E., 2022. Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 4 R1. ORNL DAAC https://doi.org/10.3334/ORNLDAAC/2129.

Virginia BMP Cost-Share Program | Fauquier County, VA. 2020. https://www.fauquiercounty.gov/government/departments-h-z/john-marshall-soil-water-conservation-district/agricultural-programs/virginia-bmp-cost-share-program.

Zhu, C., Byrd, R.H., Lu, P., Nocedal, J., 1997. Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. ACM Trans. Math. Softw. 23 (4), 550–560. https://doi.org/10.1145/279232.279236.