Journal of Hydrology 584 (2020) 124601

Contents lists available at ScienceDirect

Journal of Hydrology

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

Research papers

SWAT ungauged: Water quality modeling in the Upper Mississippi River Basin

Junyu Qi[a], Xuesong Zhang[a,b,*], Qichuan Yang[c], R. Srinivasan[d], Jeffrey G. Arnold[e], Jia Li[f], Stephanie T. Waldhoff[b], Jefferson Cole[f]

[a] Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD 20740, USA
[b] Joint Global Change Research Institute, Pacific Northwest National Laboratory and University of Maryland, College Park, MD 20740, USA
[c] Department of Infrastructure Engineering, The University of Melbourne, Parkville 3010, Australia
[d] Departments of Ecosystem Science and Management and Biological and Agricultural Engineering, Texas A&M University, College Station, TX 77843, USA
[e] USDA-ARS Grassland Soil and Water Research Laboratory, Temple, TX 76502, USA
[f] U.S. Environmental Protection Agency, 1200 Pennsylvania Avenue, NW (6207 A), Washington, DC 20460, United States

ARTICLE INFO

Keywords:
Stream flow
Sediment
Nitrate
Total nitrogen
Watershed

ABSTRACT

Improving model performance in ungauged basins has been a chronic challenge in watershed model application to understand and assess water quality impacts of agricultural conservation practices, land use change, and climate adaptation measures in large river basins. Here, we evaluate a modified version of SWAT2012 (referred to as SWAT-EC hereafter), which integrates an energy balanced soil temperature module (STM) and the CENTRUY-based soil organic matter algorithm, for simulating water quality parameters in the Upper Mississippi River Basin (UMRB), and compare it against the original SWAT2012. Model evaluation was performed for simulating streamflow, sediment, and nitrate-N (NO_3-N) and total nitrogen (TN) loadings at three stations near the outlets of UMRB. The model comparison was conducted without parameter calibration in order to assess their performance under ungauged conditions. The results indicate that SWAT-EC outperformed SWAT2012 for stream flow and NO_3-N and TN loading simulation on both monthly and annual scales. For sediment, SWAT-EC performed better than SWAT2012 on a monthly time step basis, but no noticeable improvement was found at the annual scale. In addition, the performance of the uncalibrated SWAT-EC was comparable to other calibrated SWAT models reported in previous publications with respect to sediment and NO_3-N loadings. These findings highlight the importance of advancing process representation in physically-based models to improve model credibility, particularly in ungauged basins.

1. Introduction

The design and evaluation of land and water management practices depend on the total hydrologic and biogeochemical performance within a watershed (Parmele, 1972). Watershed-scale models are widely used approaches to predict water quantity and quality accounting for complex physical, chemical, and biological processes of a watershed (Clark et al., 2015). The application of large-scale watershed models requires many spatially variable input data (including weather, topographic and soil characteristics, and land use and management data) which are difficult to obtain (Srinivasan et al., 2010). Even with enough datasets, it will take great effort to incorporate the information into model configuration (Gassman et al., 2006). More importantly, since many physical and biogeochemical processes are conceptually simplified in watershed models, many parameters governing model behavior are not physically-based and need to be calibrated before assessing alternative scenarios and supporting decision-making (Beven and Binley, 1992).

Successful model calibration requires long and high-quality observations of water quantity and quality variables which are always limited on both spatial and temporal scales, especially in ungauged watersheds (Sivapalan, 2003). Using limited observations for model calibration may cause biased estimation of calibrated parameters (Boyle et al., 2000; Doherty and Johnston, 2003). In addition, as physically-based watershed models often contain dozens or even hundreds of parameters to be calibrated, the information contained in observations of streamflow and water quality measurements at the outlet of a watershed is usually not adequate for reliably determining true values of those parameters (Beven and Smith, 2014). Since the initiative of the

* Corresponding author.
E-mail addresses: xzhang14@umd.edu, Xuesong.zhang@pnnl.gov (X. Zhang).

https://doi.org/10.1016/j.jhydrol.2020.124601

Available online 01 February 2020

IAHS Decade on Predictions in Ungauged Basins (PUB) (Sivapalan, 2003), different approaches have been developed and used to enhance watershed model reliability. For example, improvements in model performance has been achieved by transferring parameter information from adjacent gauged watersheds to the ungauged watershed (Burn and Boorman, 1993; Qi et al., 2018a) or improving input data quality using remote sensing and geographic information system (GIS) techniques (Fortin et al., 2001; Sun et al., 2010). Recent studies show that a valid approach to improve model reliability is to enhance model structure and account for well-understood hydrological and biogeochemical processes instead of using simplified processes (Butts et al., 2004; Fenicia et al., 2008; Perrin et al., 2001; Pomeroy et al., 2007).

The Soil and Water Assessment Tool (SWAT) model was originally developed to operate in large-scale ungauged watersheds with minimal calibration efforts (Arnold et al., 1998). It attempts to incorporate spatially and physically distributed watershed inputs to simulate surface and subsurface flow, sediment generation and deposit, and nutrient movement and fate through the watershed system (Gassman et al., 2007). As a continuous and semi-distributed watershed-scale model, SWAT has been successfully used to simulate water quantity and quality in a wide range of gauged or ungauged watersheds across the world (Abbaspour et al., 2017; Gitau and Chaubey, 2010; Lee et al., 2017; Lee et al., 2016; Li et al., 2014; Ndomba et al., 2005; Qi et al., 2017a; Qi et al., 2017b; Srinivasan et al., 2010; Zhang et al., 2008; Zhao et al., 2020). Despite its wide use, SWAT encountered difficulties in regions with climate, soil, and topography characteristics that go beyond the conditions the algorithms in the model have been calibrated for (Chu and Shirmohammadi, 2004; Hülsmann et al., 2015; Mittelstet et al., 2017; Qi et al., 2019a; Qi et al., 2019c; Wagner et al., 2011). One reason explaining the difficulty is that many hydrological and biogeochemical processes in SWAT are described using empirical equations which limited its application on specific conditions. During the past decades, various studies have developed new algorithms with respect to hydrology and sediment and nutrient cycles within SWAT to solve their own concerned environmental issues (Eckhardt et al., 2002; Kim and Lee, 2010; Qi et al., 2016b; Qi et al., 2019b; Qi et al., 2018b; Sakaguchi et al., 2014; Tuppad et al., 2011; Zhang et al., 2017; Zhang et al., 2008). Among those SWAT-modification studies, many were focusing on developing more physically-based modules to replace the original empirically-based algorithms. Most studies were conducted on a site or small-watershed scale to quantify the value of the improvements in physical process representation in SWAT. Note that those studies often calibrated model parameters during the comparison of different versions of SWAT. However, Zhang et al. (2013) showed that even when the structure of SWAT was not well configured, parameter calibration against streamflow observed at the watershed outlet could attain parameter values that meet satisfactory model performance criteria as suggested by Moriasi et al. (2007). Arnold et al. (2015) also pointed out the importance of accurate representation of model processes and its impact on calibration and following scenario analysis. These findings highlight the importance of improving process representation, in addition to parameter calibration, to ensure reliability of watershed model simulation.

The SWAT model was originally designed to simulate large-scale ungauged watersheds (Arnold et al., 1998). Up to date, only a few studies have evaluated SWAT performance under ungauged conditions (e.g. Srinivasan et al. (2010)). In this study, we aimed to evaluate the value of advancing soil energy and biogeochemical processes in SWAT for simulating water quality in a large river basin, i.e., the Upper Mississippi River Basin (UMRB), under ungauged conditions. Here, we evaluated an enhanced version of SWAT2012 (referred to as SWAT-EC hereafter) which has integrated an energy balanced soil temperature module (STM) and the CENTRUY-based soil organic matter algorithm and compared it against the original SWAT2012. To understand which algorithms contribute to the improvements in SWAT-EC, we also investigated performance of SWAT2012 with CENTRUY (referred to as SWAT-C) and SWAT2012 with the energy balanced soil temperature module (referred to as SWAT-E). The four versions of SWAT models were not calibrated and were evaluated against monthly and annual streamflow, sediment, nitrate (NO_3-N), and total-nitrogen (TN) loadings at three gauging stations near the outlet of UMRB. Multiple widely used statistics were used to evaluate the performance difference between different versions of SWAT models. To put the model performance in context, we also compared the uncalibrated SWAT models developed in this study with previous studies that reported performance of SWAT with calibration. We anticipate the results will help understand the value of advancing processes representation in SWAT for water quality modeling in large-scale ungauged basins.

2. Data and model description

2.1. Modifications made to SWAT2012

2.1.1. Physically-based soil temperature module

In SWAT2012, soil temperature T_{soil} is calculated at the center of a soil layer (z) on the hydrologic response unit (HRU) scale with the following equation (Neitsch, 2011),

$$T_{soil}(z) = \gamma T_{soil}^{'}(z) + (1-\gamma)[d(\bar{T}_{Aair} - T_{sur}) + T_{sur}] \tag{1}$$

where γ is the lag coefficient controlling the influence of the previous day's temperature on the current day's temperature, T_{soil} is the soil temperature from the previous day at depth z, d is the depth factor that quantifies the influence of depth on soil temperature, $\bar{T}_{Aair}$ is the average annual air temperature, and T_{sur} is the soil surface temperature on the day. The depth factor is a function of depth at the center of the soil layer (z), maximum damping depth, bulk density, and soil water. The soil surface temperature is a function of the previous day's temperature, the amount of ground cover, and the bare soil temperature. The effects of snow and plant-canopy cover on soil temperature are incorporated empirically as a weighting factor. The bare soil temperature is a function of daily average, minimum, and maximum temperature as well as solar radiation reaching the ground and albedo.

A physically-based and energy balanced STM has been developed and implemented based on heat transfer theory (Qi et al., 2016b),

$$\frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\left(\frac{k}{C}\cdot\frac{\partial T}{\partial x}\right)\frac{s}{C} \tag{2}$$

where T is the temperature (°C), t represents the time step (in days), k is the thermal conductivity (J cm^{-1} d^{-1} °C^{-1}), C is the volumetric heat capacity (J cm^{-3} °C^{-1}), x is the vertical distance from the air–soil or air-snow interface (cm), and s is the latent heat source/sink term (J cm^{-3} d^{-1}). The Eq. 2 is converted to a fully-implicit discretized form and was solved with the tridiagonal-matrix algorithm, described in Patankar (1980). The simulation domain was defined as extending from the air–soil or air-snow interface (upper boundary) to the damping depth (lower boundary), where the impact of air temperature diminishes. The temperature was calculated at the center of each soil layer for individual HRU. When snow accumulated on the ground, the snow cover was treated as a single layer. Heat capacity and thermal conductivity were assumed to be uniform within individual layers.

STM can address the intermediary role of snow cover in the daily evolution of the soil temperature profile which is important for the application of SWAT in cold-climate regions of the world. Instead of considering soil temperature as a function of air temperature by the empirical soil temperature module within SWAT, STM simulates temperature change in snow and soils as a result of heat conduction and latent heat exchange. The new module can estimate snow or soil surface temperature based on energy balance, update thermal properties of snow and soils according to changes in snow density and soil moisture, and simulate freeze–thaw cycles in the soil profile. STM has tested with field observed temperatures from a small experimental watershed in

Atlantic Canada (Qi et al., 2016b) demonstrating a great improvement in prediction of soil temperatures.

2.1.2. CENTURY-based soil organic matter algorithms

In SWAT2012 and previous versions, a PAPRAN model (Seligman and Keulen, 1980) has been used to simulate organic nitrogen cycling. The PAPRAN model considers organic N in plant residue, and active and stable humus pools. The decomposition of each pool is influenced by water, temperature, C:N ratio and intrinsic properties of substrate (Neitsch et al., 2011). One of the drawbacks of the PAPRAN model that it does not explicitly simulate carbon cycling in soil, which is closely coupled with nitrogen cycling. On the other hand, the CENTURY model (Parton et al., 1994) considers both C and N cycling and represents soil organic matter (SOM) and residue in five pools. Plant litter is split into a metabolic carbon pool (e.g. proteins and sugars) and a recalcitrant structural carbon pool (e.g. lignin and cell walls). CENTURY also has an amicrobial biomass pool. Soil humus includes two pools: slow humus pool receives C and N from the decomposition of structural litter, metabolic litter, and microbial biomass and often has a turnover time of 20–50 years, while passive humus pool includes physically- and chemically-stabilized SOM absorbed to clays and may take hundreds to thousands of years to decompose. The decomposition of each pool is determined by multiple factors, such as water, temperature, tillage, C:N ratio, oxygen availability, soil texture, and substrate specific properties.

The major difference between PAPRAN and CENTURY lies in how detailed they represent SOM-residue, how the decomposition rates are calculated, and the coupling between C and N cycling. In general, CENTURY incorporates more detailed biochemical properties and environmental factors in describing the processes relevant to SOM-residue dynamics. Zhang et al. (2013), Yang et al. (2017), and Yang and Zhang (2016) successfully tested the CENTURY algorithm for simulating C and N progress using 16 AmeriFlux agricultural sites (http://public.ornl.gov/ameriflux/) in the US Midwest and three Long-term ecological research (LTER) sites at the Kellogg Biological Station (KBS; https://lter.kbs.msu.edu/) in southern Michigan. However, the CENTURY algorithm in SWAT has not been tested for benefiting nutrients cycling modeling in large ungauged basins.

2.2. Study area

The UMRB drains an area of 431,000 km^2 from northern Minnesota to its confluence with the Ohio River at the southern tip of Illinois (Fig. 1) (Srinivasan et al., 2010). The states of Illinois, Iowa, Minnesota, Missouri, and Wisconsin constitute most of the drainage area of UMRB (Fig. 1) (Arnold et al., 2000), while Indiana, Michigan, and South Dakota together constitute a minor portion (15 percent) of the watershed. Land cover in the basin is diverse, and includes agricultural lands, forest, wetlands, lakes, prairies, and urban areas. Overall, agriculture is the dominant land use and nearly 69% of total land is used for agriculture and pasture with corn, soybeans, and alfalfa as the major crops in the basin (Wu and Tanaka, 2005). Due to its vast and complex landscapes coupled with intensive agricultural activities, landscape management, and widespread use of chemical fertilizers, UMRB is recognized as a major contributor (more than 50 percent) of the nitrogen transported to the Gulf of Mexico (Wu and Tanaka, 2005). The average annual precipitation ranges from 575 mm in the western part of Minnesota to 981 mm in the central part of Illinois, and the average monthly maximum temperature ranging from −9.8 °C in January in central Minnesota to 31.7 °C in July in central Missouri (Wu and Tanaka, 2005). Soil type in the basin ranges from heavy, poorly drained clay soil to light, well-drained sands.

2.3. SWAT model setup

The SWAT model requires a variety of detailed information describing the land use, soil, and topography data of the UMRB. The present study adopted the input data from Srinivasan et al. (2010) and compared the performance of different SWAT models. The UMRB was divided into 131 subbasins according to the eight-digit United States Geological Survey (USGS) hydrologic unit codes (HUCs; Fig. 1). National hydrography Dataset (NHD) stream dataset and a 90 m digital elevation model (DEM) was used to provide watershed configuration and topographic parameter estimation. A land use map was created by combining two sources of information, i.e., the Cropland Data Layer (CDL) and 2001 National Land Cover Data to better define cultivated and non-agricultural land use. The State Soil Geographic (STATSGO) database 1:250,000 scale soil map was used for UMRB. Using a threshold operation of 5% for land use, 10% for soil, and 5% for slope, 14,568 HRUs were generated and the number of HRUs per subbasin ranged from 58 to 216. Management practices such as tile drainage, tillage, crop rotation, and fertilizer application were included in the project according to various sources. To save space, here we only provide brief description of the input data, and detailed model setup information can be found in Srinivasan et al. (2010). The fertilization rates followed the state values reported by Unite States Department of Agriculture Economic Research Service[1]. For tile drainage, we used STATSGO soil databased to identify those locations that are poorly drained, with a slope less than 1%, and under agricultural land use. Tillage intensity information at the county scale is obtained from the Conservation Technology Information Center (CTIC[2]). We represented conservation, reduced tillage and full tillage according to slope deepness. Sorted agricultural HRUs in ascending order according to slope and try to assign full tillage to low slope HRUs and reduced tillage and conservation tillage for HRUs with deep slopes.

2.4. Weather data

The SWAT model requires daily values of precipitation, max/min air temperature, solar radiation, relative humidity, and wind speed as forcing data. Here, we used daily precipitation, max/min air temperature, solar radiation, relative humidity, and wind speed derived from Phase 2 of the North American Land Data Assimilation System (NLDAS2[3]). Climate forcing data has assimilated multiple sources of climate observations and is widely recognized as a high resolution (~1/8°), spatially continuous, and comprehensive dataset that is valuable for water cycling studies (Xia et al., 2012). The data ranges from 01 to 01-1979 to present. The spatial domain, spatial resolution, computational grid, terrain height, and land mask of NLDAS2 are identical to that in NLDAS1, which is described in Mitchell et al. (2004). We aggregated the gridded daily values of NLDAS2 data to the eight-digit subbasins using standard ArcGIS aggregation procedures. As a result, 131 weather datasets, one for each subbasin, were created to input into the SWAT model.

2.5. Streamflow and water quality data

Streamflow data were obtained from US Geological Survey (USGS) for the period of 1981–1997. Monthly sediment, nitrate, and organic nitrogen flux data were obtained from the USGS Upper Midwest Environmental Science Center (UMESC) for the period of 1981–1997. Fig. 1 shows the USGS gauging station locations that provided observed data used in comparisons. Table 1 shows the drainage area indicated by USGS and SWAT and the time period during which observed data were available. Because the USGS gauge location may not always correspond to the outlet of the HUCs, there will be some difference between the two areas. In addition, Table 1 provides the time period of data available for comparison.

[1] https://www.ers.usda.gov/data-products/fertilizer-use-and-price.aspx.

[2] www.ctic.purdue.edu/.

[3] ldas.gsfc.nasa.gov/nldas/.

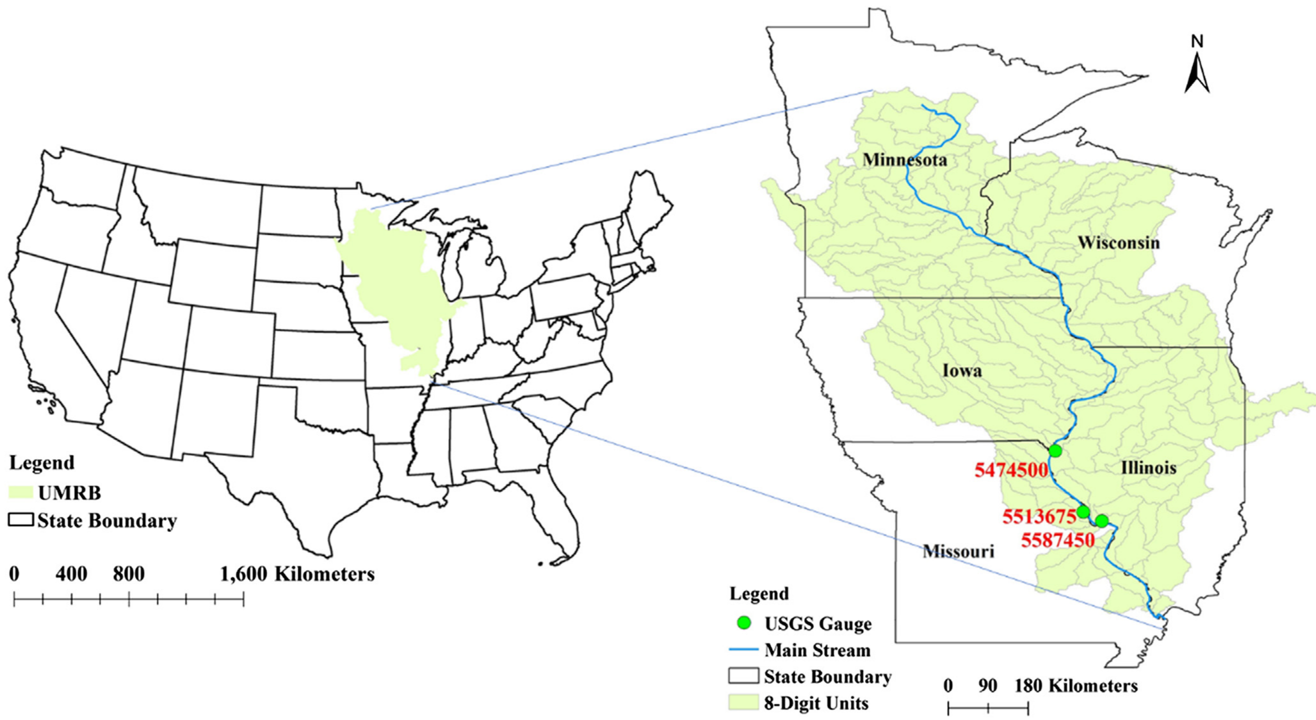

Fig. 1. Location of the Upper Mississippi River Basin (UMRB) and three USGS gauge stations on the main stream (see Table 1 for their corresponding subbasins).

Table 1
The drainage area of each gauging station, the corresponding SWAT simulated drainage area and the time period of observation data used in this study.

Subbasin No.	USGS ID	HUC8	SWAT Area (km^2)	USGS Area (km^2)	SWAT Area/USGS Area	Time Period
63	05474500	07080104	309400	304640	1.02	1980–1987
95	05513675	07110004	368000	363643	1.01	1991–1997
100	05587450	07110009	447500	438528	1.02	1980–1988

Note: HUC8: Hydrologic Unit Catalog at the eight-digit level.

2.6. Model evaluation

In this study, we evaluated four versions of SWAT, i.e., SWAT-EC, SWAT-C, SWAT-E, and SWAT2012 (Fig. 2). SWAT-EC has integrated the energy balanced soil temperature module (STM) and the CENTRUY-based soil organic matter algorithm; SWAT-C only integrated the CENTRUY-based soil organic matter algorithm; SWAT-E only integrated the energy balanced soil temperature module (Fig. 2).

The four versions of SWAT models were not calibrated and were evaluated against monthly and annual streamflow, sediment, nitrate (NO_3-N), and total-nitrogen (TN) loadings at three gauging stations near the outlet of UMRB. Model performance was assessed using three coefficients of accuracy, i.e., percent bias (P_{bias}; %), coefficient of determination (R^2), and Nash-Sutcliffe coefficient (NS) (Nash and Sutcliffe, 1970). The equations used to calculate the above three statistics are given as:

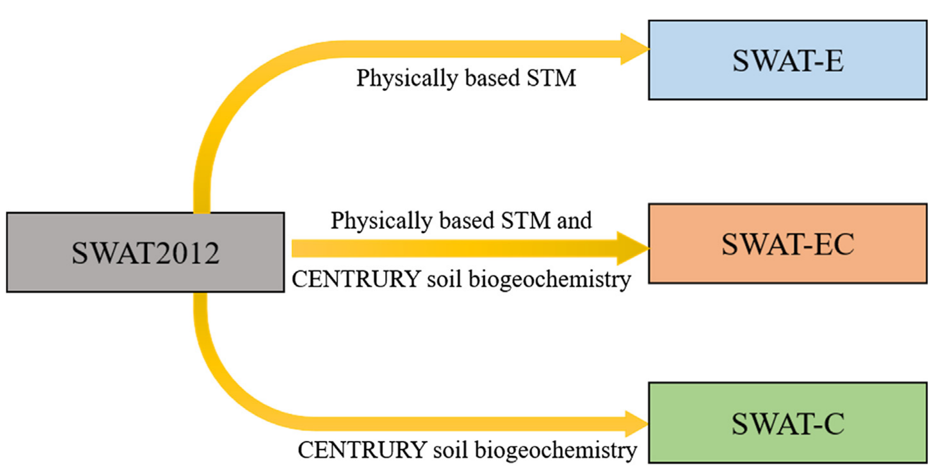

Fig. 2. Model structure difference between SWAT2012, SWAT-E, SWAT-C, and SWAT-EC. STM: energy balanced soil temperature module; CENTRUY: CENTRUY-based soil organic matter algorithm.

$$P_{bias} = 100\bullet\frac{(O_{avg} - P_{avg})}{O_{avg}} \quad (3)$$

$$NS = 1 - \frac{\sum_{i=1}^{n}(O_i - P_i)^2}{\sum_{i=1}^{n}(O_i - O_{avg})^2} \quad (4)$$

$$R^2 = \left(\frac{\sum_{i=1}^{n}(O_i - O_{avg})\bullet(P_i - P_{avg})}{[\sum_{i=1}^{n}(O_i - O_{avg})^2\bullet\sum_{i=1}^{n}(P_i - P_{avg})^2]^{0.5}}\right)^2 \quad (5)$$

where O_i and P_i are the observed and predicted values at a time step i, O_{avg} and P_{avg} are the average of the observed and predicted values. Table 1 shows the basic characteristics of the USGS stations used for model performance evaluation.

3. Results and discussion

3.1. Performance of four versions of SWAT model with respect to different water quantity and quality variables

Model performance evaluation results on stream flow and sediment, NO_3-N, and TN loadings at three USGS gauging stations at monthly and annual scales are shown in Tables 2 and 3, respectively. The evaluation periods for USGS gauging stations corresponding to the outlets of subbasin # 63 and 100 is from 1980 to 1987 and from 1980 to 1988, respectively, while for the station at the outlet of subbasin # 95 is from 1991 to 1997. In the table, shaded numbers indicate the three versions of enhanced SWAT model (i.e., SWAT-C, SWAT-E, and SWAT-EC) performed less than SWAT2012. Average annual basin values for major hydrology and nitrogen cycle components from 1980 to 1997 simulated by four versions of SWAT model are shown in Table 4.

3.1.1. Streamflow

For monthly streamflow simulation, all three enhanced SWAT (i.e. SWAT-C, SWAT-E and SWAT-EC) outperformed SWAT2012 (Table 2). SWAT-C performed modestly better than SWAT2012 in terms of all three statistics and at all three stations. That is, SWAT-C both explained more monthly variation and reduced overall bias. Compared with SWAT2012, SWAT-E achieved much better simulation of monthly streamflow, but slightly increased P_{bias} by 1% for all three stations. Comparison between SWAT-C and SWAT-E indicates that SWAT-C performed better in terms of reducing overall bias, while SWAT-E was better regarding capturing variability of monthly streamflow. SWAT-EC seems to benefit from the advantage of both SWAT-C and SWAT-C, and pronouncedly outperformed SWAT2012 in terms of explaining variations in monthly streamflow and reducing overall bias. On the annual scale, model performance for the four versions of SWAT model in general followed the pattern observed, except that SWAT-E slightly performed less than SWAT2012 in terms of NS at the outlets of subbasin # 95 and 100 (Table 3).

By analyzing the average annual values of key hydrologic budget components (Table 4), we found that SWAT-C simulated lower average annual ET and higher average annual surface runoff, lateral flow, and baseflow than SWAT2012. Even though water flow through different pathways were increased by SWAT-C, average annual NO_3-N loadings transported with those water flows became less than those of SWAT2012 (Table 4), mainly because the CENTURY method generated less nitrate in soils compared with PAPRAN.

SWAT-E pronouncedly improved streamflow simulation compared with SWAT2012 because, with STM, SWAT-E tended to generate more infiltration and less surface runoff during winter and the snowmelt season on the watershed-scale (Qi et al., 2016a). This is also can be seen from Table 4 where average annual surface runoff (85 mm) generated with SWAT-E was less than that (113 mm) of SWAT2012, and average annual lateral flow and baseflow (39 and 109 mm, respectively) generated with SWAT-E were greater than those (30 and 92 mm, respectively) of SWAT2012. The new STM more realistically accounted for snowpack insulation effects on soil temperature allowing surface soil layers to remain unfrozen when snow accumulating on the ground. As a result, SWAT-E improved its simulation on hydrological effects of unfrozen soil during winter and the snowmelt seasons.

3.1.2. Sediment

Like streamflow simulation, all three enhanced versions of SWAT model attained higher R^2 and NS values than the corresponding values of SWAT2012 at all three stations. Notably, SWAT-EC achieved the highest R^2 and NS values among all four models at all three stations, except for R^2 at the outlet of subbasin # 63 (which is slightly less than the R^2 of SWAT-E; Table 2). SWAT-E outperformed SWAT-C with respect to R^2 and NS values. Compared with SWAT2012, SWAT-C reduced P_{bais} values (generated more sediment loadings) while SWAT-E increased P_{bias} values (generated less sediment loadings) which is the major factor explaining the reduced P_{bias} values of SWAT-EC that fall somewhere between those by SWAT-C and SWAT-E (Table 2). The reason for reduced P_{bias} by SWAT-C is that SWAT-C generated more stream flow as explained in the previous section and Table 4, while SWAT-E generated less stream flow and caused increases in P_{bias}. On the annual scale, model performance for the four versions of SWAT model generally follows the pattern observed on the monthly scale, except that SWAT-EC performed less than SWAT2012 at two stations in terms of R^2 and/or NS (Table 3).

Although we did not directly modify algorithms of erosion from uplands and sediment deposition and resuspension processes in

Table 2
Model performance on monthly water quantity and quality by four versions of SWAT model at three monitoring sites in the UMRB.

		SWAT2012			SWAT-C			SWAT-E			SWAT-EC		
Variable	Subbasin	R^2	NS	P_{bias} (%)	R^2	NS	P_{bias}(%)	R^2	NS	P_{bias} (%)	R^2	NS	P_{bias} (%)
Stream Flow	63	0.65	0.31	19	0.68	0.39	14	0.83	0.58	20	0.82	0.62	14
	95	0.71	0.55	16	0.73	0.60	10	0.84	0.71	17	0.84	0.73	11
	100	0.68	0.41	24	0.71	0.50	19	0.84	0.62	25	0.85	0.69	19
Sediment	63	0.28	0.23	34	0.30	0.25	29	0.51	0.30	39	0.47	0.30	36
	95	0.51	0.50	-14	0.54	0.51	-23	0.73	0.71	-6	0.73	0.72	-17
	100	0.57	0.48	32	0.59	0.53	27	0.77	0.57	37	0.77	0.62	32
NO_3-N	63	0.53	0.37	38	0.47	0.27	41	0.68	0.61	15	0.57	0.47	30
	95	0.51	-2.51	-18	0.55	-0.14	19	0.68	-4.48	-56	0.61	0.02	6
	100	0.71	0.57	0	0.61	0.57	16	0.73	-0.18	-35	0.65	0.61	2
Total-N	63	0.56	0.20	-2	0.53	0.36	7	0.70	0.11	-10	0.62	0.54	11
	95	0.59	-1.60	-40	0.56	-0.04	-4	0.66	-2.79	-56	0.63	0.18	-1
	100	0.66	0.08	-18	0.60	0.42	0	0.73	-0.30	-30	0.68	0.59	3

Note: shaded numbers indicate enhanced SWAT has lower R^2 or NS values than SWAT2012 or has higher absolute values of P_{bias} than SWAT2102.

Table 3
Model performance on annual water quantity and quality by four versions of SWAT model at three monitoring sites in the UMRB.

		SWAT2012			SWAT-C			SWAT-E			SWAT-EC		
Variable	**Subbasin**	**R^2**	**NS**	**P_{bias} (%)**	**R^2**	**NS**	**P_{bias} (%)**	**R^2**	**NS**	**P_{bias} (%)**	**R^2**	**NS**	**P_{bias} (%)**
Stream Flow	63	0.98	0.20	20	0.98	0.48	14	0.98	0.20	20	0.99	0.51	14
	95	0.98	0.65	18	0.98	0.82	12	0.98	0.63	18	0.98	0.81	12
	100	0.95	0.18	26	0.96	0.44	21	0.96	0.15	27	0.97	0.43	21
Sediment	63	0.90	0.43	34	0.90	0.51	29	0.92	0.33	39	0.80	0.30	36
	95	0.99	0.94	-9	0.99	0.92	-18	0.99	0.95	-3	0.99	0.93	-14
	100	0.83	0.44	32	0.83	0.55	27	0.85	0.32	37	0.85	0.47	32
NO_3-N	63	0.58	-0.09	38	0.43	-0.38	41	0.67	0.54	15	0.47	0.04	30
	95	0.68	-5.27	-22	0.82	-0.88	17	0.69	-13.01	-55	0.88	-0.02	4
	100	0.82	0.68	0	0.82	0.65	16	0.81	-0.51	-35	0.84	0.83	2
Total-N	63	0.70	0.50	-1	0.59	0.49	7	0.73	0.23	-10	0.58	0.46	11
	95	0.70	-1.04	-36	0.77	0.66	-1	0.71	-2.36	-50	0.83	0.78	1
	100	0.83	0.35	-18	0.83	0.81	0	0.82	-0.42	-30	0.84	0.82	3

Note: shaded numbers indicate enhanced SWAT has lower R^2 or NS values than SWAT2012 or has higher absolute values of P_{bias} than SWAT2102.

Table 4
Average annual basin values for major hydrology and nitrogen cycle components from 1980 to 1997 simulated by four versions of SWAT model.

Variable	SWAT2012	SWAT-C	SWAT-E	SWAT-EC
Precipitation (mm)	871	871	871	871
Evapotranspiration (mm)	641	624	646	629
Surface Runoff (mm)	113	116	85	88
Lateral Flow (mm)	30	33	39	41
Baseflow (mm)	92	102	109	121
Surface Runoff-N (kg ha^{-1})	1.37	1.01	0.92	0.78
Lateral Flow-N (kg ha^{-1})	4.87	3.43	6.23	4.16
Baseflow-N (kg ha^{-1})	3.78	2.20	5.00	2.65

streams, due to the fact that surface runoff and stream flow are closely correlated with sediment generation and transport especially at the monthly scale, improved simulation of surface runoff and stream flow led to improved simulation of monthly variability of sediment loading as indicated by the comparison results.

3.1.3. Nitrate

Different from model performance comparison results for stream flow and sediment loading, no clear pattern was identified for NO_3-N loading. That is, model performance evaluation results are in general mixed. SWAT-C had greater R^2 at the outlet of subbasin # 95 and greater NS at the outlets of subbasin # 63 and 100 than SWAT2012 (Table 2). Although SWAT-E attained greater R^2 for all three stations than SWAT2012, it only had one station with a greater NS than SWAT2012 (Table 2). The absolute values of P_{bias} of SWAT-C and SWAT-E were all larger than those of SWAT2012 except that SWAT-E had lower bias at the outlet of subbasin # 63 (Table 2). With examination of the P_{bias} values of different versions of SWAT model, we found that, as compared with SWAT2012, SWAT-C distinctly increased P_{bias} (less nitrate loadings) while SWAT-E reduced P_{bias} (more nitrate loadings). As SWAT-EC merged both SWAT-C and SWAT-E, its P_{bias} values fall somewhere between those of SWAT-C and SWAT-E, and as a result, the absolute P_{bias} values of SWAT-EC were in general less than those of SWAT2012 (the difference is negligible at the outlet of subbasin # 100, i.e., 0% of SWAT2012 vs. 2% of SWAT-EC; Table 2). In addition, SWAT-EC outperformed SWAT2012 in terms of most of R^2 and NS values at the three stations, except for R^2 value at the outlet of subbasin # 100 (Table 2).

The increased P_{bias} values for SWAT-C is partially explained by lower average annual NO_3-N loadings transported with different water flows compared with those of SWAT2012 (Table 4). Since SWAT-E simulated more infiltration than SWAT2012 in areas with seasonal snow cover during winter and the snowmelt season, more nitrate was leached and exported to streams as compared with SWAT2012. This is consistent with increased average annual NO_3-N loadings transported with lateral flow and baseflow and decreased average annual NO_3-N loadings transported with surface runoff compared with those of SWAT2012 (Table 4). Overall, these changes in SWAT-E caused the reduction in P_{bias} values at the watershed-scale as compared with SWAT2012 (Qi et al., 2016a). In general, SWAT-EC balances the strength and weakness of SWAT-C and SWAT-E and achieved an overall better model performance on NO_3-N loading than SWAT2012 on the monthly scale (Table 2). On the annual scale, model performance results are like the monthly scale findings (Table 3).

3.1.4. Total nitrogen

The four models exhibited similar performance for simulating NO_3-N and TN, mainly because NO_3-N accounts for a major portion of TN. Like the model comparison for nitrate simulation, no single model outperformed the others at all stations and for all evaluation statistics. For instance, SWAT-C had lower values of R^2 but higher NS values than SWAT2012 at all three stations (Table 2). SWAT-E attained higher values of R^2 but lower NS values than SWAT2012 at three stations (Table 2). In general, SWAT2012 and SWAT-E tended to overestimate TN, while SWAT-C exhibited underestimation (Table 2). SWAT-EC better captured seasonal variability of TN and achieved lower bias than SWAT2012 at all three stations, except for higher bias at the outlet of subbasin # 63 (Table 2). The annual results (Table 3) also confirmed the advantage of integrating advanced biogeochemical and energy balanced algorithms for achieving improved simulation of nitrogen loading in large-scale ungauged basins.

3.2. Monthly examination of the results by SWAT2012 and SWAT-EC

The observed and simulated monthly stream flow and sediment, NO_3-N and TN loadings by SWAT2012 and SWAT-EC at the three stations, respectively shown in Figs. 3, 4 and 5. The SWAT-EC model was found to have slightly damped snowmelt flow peaks compared with SWAT2012 leading to improved model performance as indicated in Tables 2 and 3. As stated in Section 3.1.1, SWAT-EC tends to generate more infiltration and less surface runoff during winter and snowmelt seasons, because the STM accounts for snow insulation effects on soil temperature leading to surface soil layers remaining unfrozen when snow cover presents.

Both versions of SWAT model tended to underestimate sediment loading peaks during the spring snowmelt season, especially at the outlets of subbasin # 63 and 100 (Fig. 3b and 5b), which explained the relatively large P_{bias} shown in Tables 2 and 3. Compared with

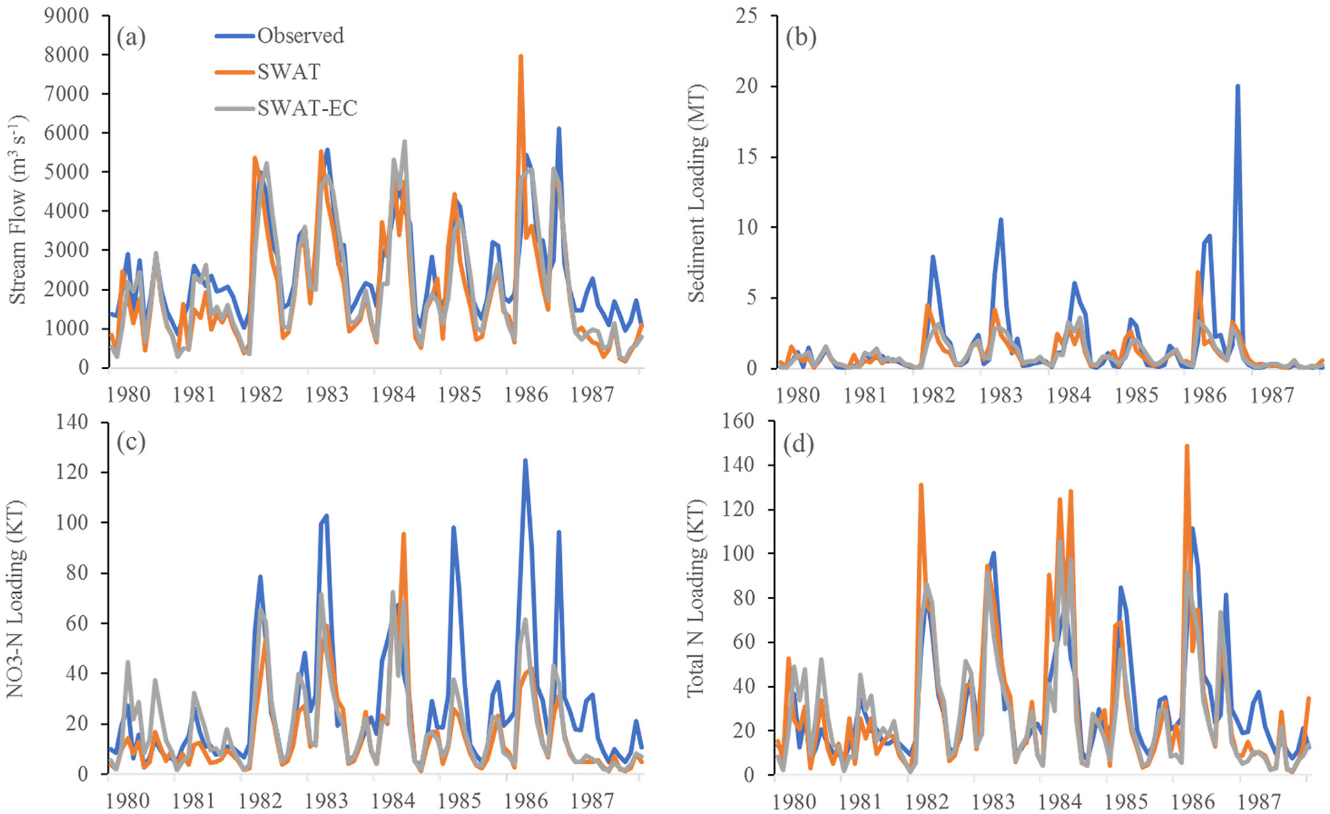

Fig. 3. Observed and simulated stream flow (a) and sediment (b), NO_3-N(c) and TN (d) loadings at USGS gauge # 05474500 (corresponding to subbasin # 63) on the monthly scale for both SWAT2012 and SWAT-EC.

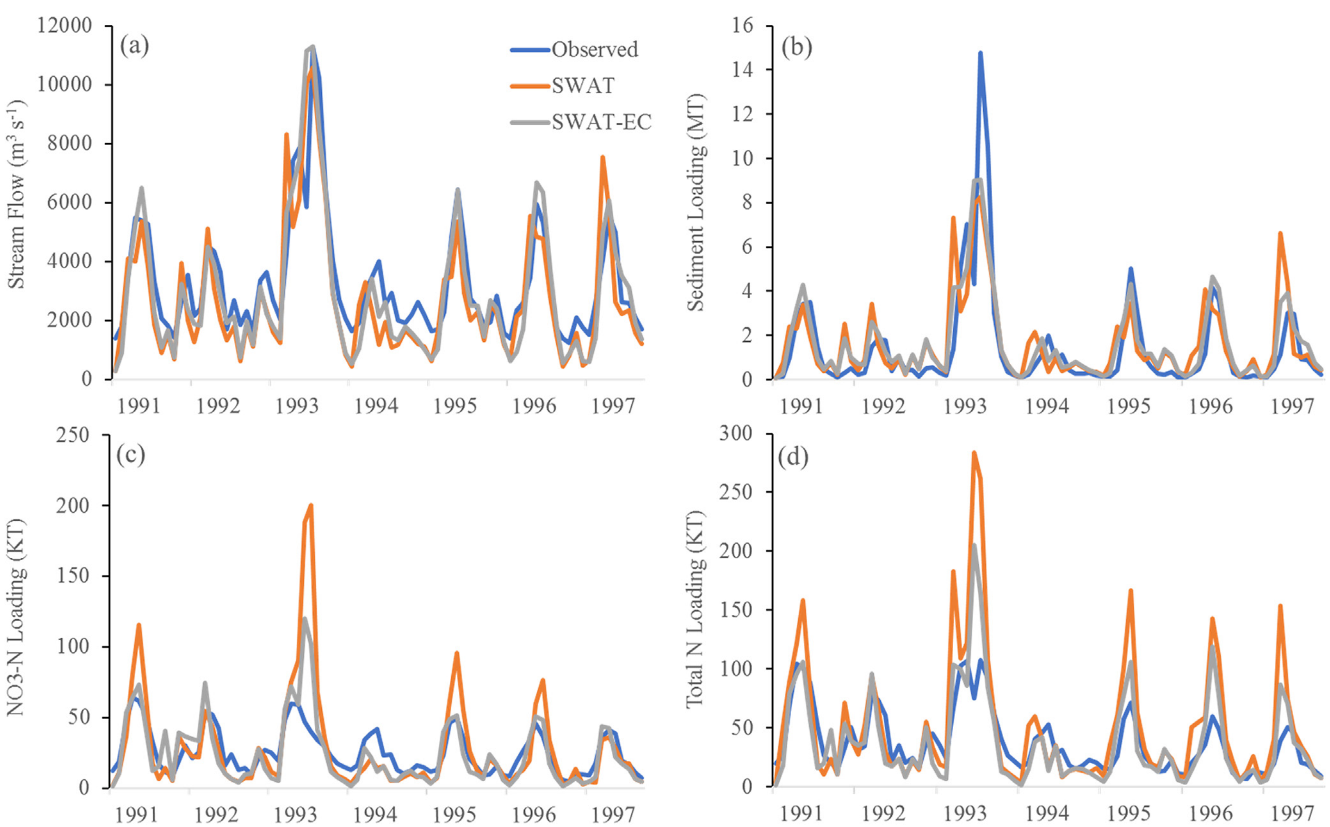

Fig. 4. Observed and simulated stream flow (a) and sediment (b), NO_3-N(c) and TN (d) loadings at USGS gauge # 05513675 (corresponding to subbasin # 95) on the monthly scale for both SWAT2012 and SWAT-EC.

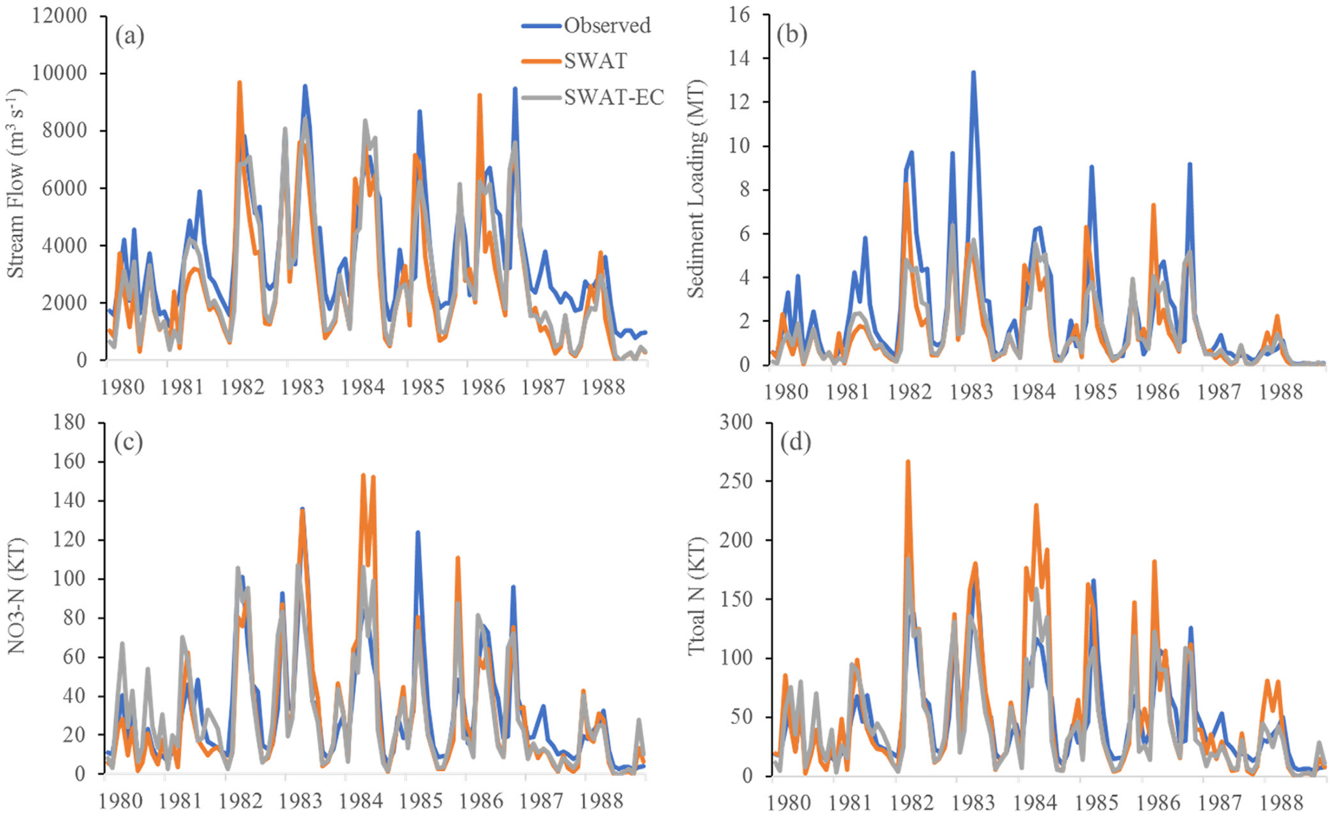

Fig. 5. Observed and simulated stream flow (a) and sediment (b), NO_3-N(c) and Total-N (d) loadings at USGS gauge # 05587450 (corresponding to subbasin # 100) on the monthly scale for both SWAT2012 and SWAT-EC.

SWAT2012, SWAT-EC obtained less sediment loading peaks because it generated less surface runoff (Fig. 3b, 4b, and 5b). Furthermore, SWAT2012 tended to generate sediment loading peaks ahead of observed peaks (about one month earlier), while SWAT-EC simulated sediment loading peaks matched well the timing of observation. Note that, the sediment loading peaks generated by SWAT-EC were less than those generated by the SWAT2012 (Fig. 3b, 4b, and 5b) mainly due to lower surface runoff.

Both versions of SWAT model tended to underestimate NO_3-N loading peaks during snowmelt season at the outlets of subbasin # 63 from 1980 to 1987 and overestimate them at the outlet of subbasin # 95 from 1991 to 1997 (Fig. 3c and 4c). Such a shift in model performance over time may be caused by the lack of detailed fertilization data over years and at fine scales that are required for accurate nutrient loss simulation. Compared with SWAT2012, SWAT-EC generated much more NO_3-N loadings at the outlet of subbasin # 63 and less NO_3-N loadings at the outlet of subbasin # 95 during snowmelt season, which is a major reason explaining the improvement in model performance for NO_3-N loading. Both models performed well at the outlet of subbasin #100 with SWAT-EC performed slightly better (Fig. 5c and Table 2). In addition, SWAT2012 tended to overestimate TN loading peaks during snowmelt period at all three stations, while TN loading peaks simulated by SWAT-EC were more consistent with observations (Fig. 3c, 4c, and 5c), which explained the improvement in model performance of TN loading as shown in Tables 2 and 3.

The overall improvement of SWAT-EC over SWAT2012 for simulating NO_3-N and TN is the result of combining the strengths from SWAT-C and SWAT-E. As shown in Table 2, SWAT-E contributed more to the improvements in SWAT-EC with respect to explaining variability of hydrology and sediment yield, while contribution from SWAT-C modestly improved bias in streamflow and sediment yield simulation. As to nitrate and TN, neither SWAT-C nor SWAT-E clearly outperformed SWAT2012. For example, SWAT-E performed less than SWAT2012 in terms of both NS and P_{bias} in simulating TN for all stations, while SWAT-C had larger bias at all stations from simulation of nitrate. These results demonstrate the complexity and combined efforts of energy and biogeochemical processes on hydrology and water quality, and highlight the importance of advancing the coupled water-energy-biogeochemistry processes within SWAT toward better watershed modeling and assessment.

Based on NS and P_{bias} values generated by SWAT-EC, model performance was satisfactory for stream flow (NS > 0.5 and P_{bais} < 25%) at all three gauging stations on the monthly scale. SWAT-EC performances on sediment loading were also satisfactory (NS > 0.5 and P_{bias} < 55%) at these stations. Statistic values were all falling into satisfactory category for NO_3-N and Total-N simulation on the monthly scale (Moriasi et al., 2007) except for NS for subbasin # 95 (0.02 for NO_3-N and 0.18 for TN). These results demonstrate the potential of advancing process representation in SWAT to achieve satisfactory simulation of water quality in large-scale ungauged basins.

3.3. Assessment of a simple Multi-Model approach

Multi-model approaches often help reduce uncertainty in hydrological predictions and improve model predictability, because a combination of predictions from many single models could synergize the strength of these single models (Li and Sankarasubramanian, 2012). Multi-model predictions are usually obtained by taking a weighted average of the predictions from the single models (the weights sum up to one). The weights can be derived from several multi-model combination methods, including weighted average of single model predictions or using statistical techniques (e.g., multiple linear regression and Bayesian model averaging) (Duan et al., 2007; Krishnamurti et al., 1999; Marshall et al., 2006; Oudin et al., 2006; Shamseldin et al., 1997). In this study, we used the simple equal-weighted method to assess model performance on monthly water quality and quantity

Table 5
Model performance on monthly water quantity and quality by combining simulations from four versions of SWAT model at three monitoring sites in the UMRB.

Variable	Subbasin	R^2	NS	P_{bias} (%)
Stream Flow	63	0.78	0.56	17
	95	0.80	0.68	14
	100	0.79	0.59	22
Sediment	63	0.41	0.28	35
	95	0.64	0.63	−15
	100	0.70	0.57	32
NO_3-N	63	0.72	0.50	32
	95	0.61	−0.98	−11
	100	0.73	0.63	−3
Total-N	63	0.63	0.43	2
	95	0.62	−0.66	−24
	100	0.69	0.37	−10

prediction. Specifically, an ensemble simulation was generated by taking the average of the four simulations by the four versions of SWAT, i.e., SWAT2012, SWAT-E, SWAT-C, and SWAT-EC for streamflow, sediment, NO_3-N, and Total N. Correspondingly, three statistics (i.e., R^2, NS, and P_{bias}) were calculated for the new ensemble simulation (Table 5). In comparison with the statistical values shown Table 3, the ensemble simulation generally performed better than then poorest single model but less than the best single model. This result indicates that a simple average of the four SWAT models may not yield better performance. However, more advanced methods (e.g., multiple linear regression or Bayesian model averaging) deserve research to combine different model structures to achieve more accurate hydrologic and water quality modeling in the future.

3.4. Comparison with previous studies

Multiple studies have been conducted to simulate water quantity and quality in the UMRB using calibrated versions of SWAT (Arnold et al., 2000; Jha et al., 2006; Jha et al., 2004; Jha et al., 2015; Wu et al., 2012). As for water quality variables, only sediment and NO_3-N loading simulation results were reported in two studies (Jha et al., 2006; Jha et al., 2015). We summarized the model performance evaluated based on R^2 and NS from those studies to put the performance of the four versions of SWAT examined in context. The statistics values reported in previous studies are shown in Table 6. As those previous studies only reported results at the gauging station at the outlet of UMRB (i.e., USGS gauge # 05587450), the comparison between our results and previous studies was only conducted at this site.

For sediment loading simulation, compared with Jha et al.'s (2006) results in their calibration period (R^2 = 0.66 and NS = 0.66), SWAT2012 had lower R^2 and NS values (R^2 = 0.57 and NS = 0.48), while SWAT-EC achieved higher R^2 (0.77) and comparable NS (0.62) on the monthly scale. On the annual scale, both SWAT2012 and SWAT-EC attained greater R^2 (0.83 and 0.85, respectively) and less NS (0.44 and 0.47, respectively) values than those (R^2 = 0.77 and NS = 0.69) of Jha et al., 2006 for their calibration period. Considering the different periods of dataset used in model evaluation (1980–1988 for present study vs. 1981–1992 for Jha et al., 2006), it is difficult to draw a definite conclusion. However, it is fair to conclude that our sediment simulation evaluation is comparable to those reported previous.

Since no previous study reported SWAT performance on monthly NO_3-N loading, we only compared our annual results to those reported by Jha et al. (2015). Both SWAT2012 and SWAT-EC models attained higher R^2 (0.82 and 0.84, respectively) and NS (0.68 and 0.83, respectively) values than those (R^2 = 0.62 and NS = 0.46) of Jha et al., 2015 reported for the period of 1981–2003. Again, due to the difference in the periods of data used in this study (1981–1988) and by Jha et al. (2015) (1981–2003), we could not conclude that the two uncalibrated

Table 6
Model performance on monthly and annual water quantity and quality in previous studies at USGS gauge # 05587450 in the UMRB.

	Variable	R^2	NS	Period	Sources
Monthly	Stream	0.63		1960–1980 (calibration)	Arnold et al., 2000
		0.65		1981–1985 (validation)	
		0.75	0.67	1989–1997 (calibration)	Jha et al., 2004
		0.7	0.59	1980–1988 (validation)	
		0.71	0.65	1981–1992 (calibration)	Jha et al., 2006
		> 0.8	> 0.8	1993–2003 (validation)	
		0.79	0.77	1991–2000 (calibration)	
		0.80	0.79	2001–2008 (validation 1)	Wu et al., 2012
		0.75	0.74	1961–1990 (validation 2)	
	Sediment	0.66	0.66	1981–1992 (calibration)	Jha et al., 2006
		0.55	0.54	1993–2003 (validation)	
		0.57	0.48	1980–1988 (uncalibrated)	SWAT2012
		0.77	0.62	1980–1988 (uncalibrated)	SWAT-EC
	NO_3-N	N/A	N/A	N/A	N/A
	TN	N/A	N/A	N/A	N/A
Annual	Stream	0.91	0.91	1989–1997 (calibration)	Jha et al., 2004
		0.89	0.86	1980–1988 (validation)	
		> 0.8	> 0.8	1981–1992 (calibration)	Jha et al., 2006
		> 0.8	> 0.8	1993–2003 (validation)	
		0.94	0.93	1981–2003	Jha et al., 2015
		0.97	0.94	1991–2000 (calibration)	
		0.95	0.94	2001–2008 (validation 1)	Wu et al., 2012
		0.85	0.85	1961–1990 (validation 2)	
	Sediment	0.77	0.69	1981–1992 (calibration)	Jha et al., 2006
		0.90	0.85	1993–2003 (validation)	
		0.83	0.44	1980–1988 (uncalibrated)	SWAT2012
		0.85	0.47	1980–1988 (uncalibrated)	SWAT-EC
	NO_3-N	0.62	0.46	1981–2003	Jha et al., 2015
		0.82	0.68	1980–1988 (uncalibrated)	SWAT2012
		0.84	0.83	1980–1988 (uncalibrated)	SWAT-EC
	TN	N/A	N/A	N/A	N/A

versions of SWAT performed better. But the comparison between the statistics shows the promise of using the latest released SWAT model and the enhanced SWAT with respect to energy and biogeochemistry to achieve satisfactory simulation of nitrate loading in ungauged basin.

4. Conclusions

In the present study, we evaluated SWAT2012 and three enhanced versions (i.e., SWAT-C, SWAT-E, and SWAT-EC, which integrates a CENTRUY-based soil organic matter module, an energy balanced soil temperature module (STM), and both CENTURY and STM, respectively) for simulating water quality variables in the Upper Mississippi River Basin (UMRB) under ungauged conditions. We found that both SWAT-C and SWAT-E improved hydrology and sediment simulations as compared with SWAT2012, where SWAT-C helped reduce overall bias and SWAT-E better captured seasonality, particularly during winter and snowmelt seasons. SWAT-EC combined the strengths of both SWAT-C and SWAT-E, leading to much improved explanation of monthly variation and comparable biases, as compared with SWAT2012.

As to nitrate (NO_3-N) and total-nitrogen (TN), neither SWAT-C nor SWAT-E clearly outperformed SWAT2012, which demonstrate the complexity and combined efforts of energy and biogeochemical processes on hydrology and water quality. In general, SWAT-C performance is similar to that of SWAT2012, while SWAT-E improved simulation of spring fluxes of NO_3-N and TN. Again, SWAT-EC benefited from both SWAT-C and SWAT-E and, for most cases, achieved better results than SWAT2012 in terms of explaining observed data variability and reducing simulation biases.

Based on the model evaluation criteria summarized in Moriasi et al. (2007), the performance of the uncalibrated SWAT-EC model was categorized as "satisfactory", and was comparable to calibrated SWAT models reported in previous studies. Overall, our results demonstrate the value of advancing process representation in physically-based watershed models to improve model credibility, particularly in ungauged basins. We anticipate that the new model development efforts will benefit future watershed modeling and assessment efforts in large, ungauged basins.

CRediT authorship contribution statement

Junyu Qi: Conceptualization, Methodology, Formal analysis, Visualization, Writing - original draft, Writing - review & editing. **Xuesong Zhang:** Supervision, Conceptualization, Methodology, Formal analysis, Visualization, Writing - original draft, Writing - review & editing. **Qichuan Yang:** Writing - review & editing. **R. Srinivasan:** Writing - review & editing. **Jeffrey G. Arnold:** Writing - review & editing. **Jia Li:** . **Stephanie T. Waldholf:** Writing - review & editing. **Jefferson Cole:** Writing - review & editing.

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.

Acknowledgements

The funding support for this project was provided by National Aeronautics and Space Administration (NNH13ZDA001N, NNX17AE66G and 18-CMS18-0052), and United States Department of Agriculture (2017-67003-26485 and 2017-67003-26484), and National Science Foundation (1639327). Funding was also provided in part by the United States Department of Agriculture Natural Resources Conservation Service - Conservation Effects Assessment Project (NRCS-CEAP). The views expressed in this paper are those of the authors and do not necessarily represent the views or policies of the U.S. Government.

References

Abbaspour, K.C., Vaghefi, S.A., Srinivasan, R., 2017. A Guideline for Successful Calibration and Uncertainty Analysis for Soil and Water Assessment: A Review of Papers from the 2016 International SWAT Conference. Multidisciplinary Digital Publishing Institute.

Arnold, J.G., Muttiah, R.S., Srinivasan, R., Allen, P.M., 2000. Regional estimation of base flow and groundwater recharge in the Upper Mississippi river basin. J. Hydrol. 227 (1–4), 21–40.

Arnold, J.G., Srinivasan, R., Muttiah, R.S., Williams, J.R., 1998. Large area hydrologic modeling and assessment part I: Model development. JAWRA 34 (1), 73–89.

Arnold, J.G., et al., 2015. Hydrological processes and model representation: Impact of soft data on calibration. Trans. ASABE 58 (6), 1637–1660.

Beven, K., Binley, A., 1992. The future of distributed models: model calibration and uncertainty prediction. Hydrol. Process. 6 (3), 279–298.

Beven, K., Smith, P., 2014. Concepts of information content and likelihood in parameter calibration for hydrological simulation models. J. Hydrol. Eng. 20 (1), A4014010.

Boyle, D.P., Gupta, H.V., Sorooshian, S., 2000. Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods. Water Resour. Res. 36 (12), 3663–3674.

Burn, D.H., Boorman, D.B., 1993. Estimation of hydrological parameters at ungauged catchments. J. Hydrol. 143 (3–4), 429–454.

Butts, M.B., Payne, J.T., Kristensen, M., Madsen, H., 2004. An evaluation of the impact of model structure on hydrological modelling uncertainty for streamflow simulation. J. Hydrol. 298 (1–4), 242–266.

Chu, T., Shirmohammadi, A., 2004. Evaluation of the SWAT model's hydrology component in the piedmont physiographic region of Maryland. Trans. ASAE 47 (4), 1057.

Clark, M.P., et al., 2015. Improving the representation of hydrologic processes in Earth System Models. Water Resour. Res. 51 (8), 5929–5956.

Doherty, J., Johnston, J.M., 2003. Methodologies for calibration and predictive analysis of a watershed model 1. JAWRA 39 (2), 251–265.

Duan, Q., Ajami, N.K., Gao, X., Sorooshian, S., 2007. Multi-model ensemble hydrologic prediction using Bayesian model averaging. Adv. Water Resour. 30 (5), 1371–1386.

Eckhardt, K., Haverkamp, S., Fohrer, N., Frede, H.-G., 2002. SWAT-G, a version of SWAT99. 2 modified for application to low mountain range catchments. Phys. Chem. Earth, Parts A/B/C 27 (9–10), 641–644.

Fenicia, F., Savenije, H.H., Matgen, P., Pfister, L., 2008. Understanding catchment behavior through stepwise model concept improvement. Water Resour. Res. 44 (1).

Fortin, J.-P., et al., 2001. Distributed watershed model compatible with remote sensing and GIS data. I: Description of model. J. Hydrol. Eng. 6 (2), 91–99.

Gassman, P.W., Reyes, M.R., Green, C.H., Arnold, J.G., 2007. The soil and water assessment tool: historical development, applications, and future research directions. Trans. ASABE 50 (4), 1211–1250.

Gassman, P.W., Secchi, S., Jha, M., Kurkalova, L., 2006. Upper Mississippi River Basin modeling system part 1: SWAT input data requirements and issues. Coast. Hydrol. Process. 103–115.

Gitau, M.W., Chaubey, I., 2010. Regionalization of SWAT model parameters for use in ungauged watersheds. Water 2 (4), 849–871.

Hülsmann, L., Geyer, T., Schweitzer, C., Priess, J., Karthe, D., 2015. The effect of subarctic conditions on water resources: initial results and limitations of the SWAT model applied to the Kharaa River Basin in Northern Mongolia. Environ. Earth Sci. 73 (2), 581–592.

Jha, M., Gassman, P., Secchi, S., Arnold, J., 2006. Upper Mississippi river basin modeling system part 2: baseline simulation results. Coastal Hydrology and Processes. Water Resources Publications, Highlands Ranch, Colorado, pp. 117–126.

Jha, M., Pan, Z., Takle, E.S., Gu, R., 2004. Impacts of climate change on streamflow in the Upper Mississippi River Basin: A regional climate model perspective. J. Geophys. Res.: Atmos. 109 (D9).

Jha, M.K., Gassman, P.W., Panagopoulos, Y., 2015. Regional changes in nitrate loadings in the Upper Mississippi River Basin under predicted mid-century climate. Reg. Environ. Change 15 (3), 449–460.

Kim, N.W., Lee, J., 2010. Enhancement of the channel routing module in SWAT. Hydrol. Process.: Int. J. 24 (1), 96–107.

Krishnamurti, T., et al., 1999. Improved weather and seasonal climate forecasts from multimodel superensemble. Science 285 (5433), 1548–1550.

Lee, S., Sadeghi, A.M., Yeo, I.-Y., McCarty, G.W., Hively, W.D., 2017. Assessing the impacts of future climate conditions on the effectiveness of winter cover crops in reducing nitrate loads into the Chesapeake Bay Watershed using SWAT model. Trans. Am. Soc. Agric. Biol. Eng. 60 (6), 1939–1955. https://doi.org/10.13031/trans.12390.

Lee, S., et al., 2016. Impacts of watershed characteristics and crop rotations on winter cover crop nitrate-nitrogen uptake capacity within agricultural watersheds in the Chesapeake Bay region. PLoS One 11 (6), e0157637.

Li, Q., et al., 2014. An approach for assessing impact of land use and biophysical conditions across landscape on recharge rate and nitrogen loading of groundwater. Agric. Ecosyst. Environ. 196, 114–124. https://doi.org/10.1016/j.agee.2014.06.028.

Li, W., Sankarasubramanian, A., 2012. Reducing hydrologic model uncertainty in monthly streamflow predictions using multimodel combination. Water Resour. Res. 48 (12).

Marshall, L., Sharma, A., Nott, D., 2006. Modeling the catchment via mixtures: Issues of model specification and validation. Water Resour. Res. 42 (11).

Mitchell, K.E., et al., 2004. The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. J. Geophys. Res.: Atmos. 109 (D7).

Mittelstet, A., Storm, D., Fox, G., 2017. Testing of the modified streambank erosion and instream phosphorus routines for the SWAT model. JAWRA 53 (1), 101–114.

Moriasi, D.N., et al., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. ASABE 50 (3), 885–900.

Nash, J., Sutcliffe, J., 1970. River flow forecasting through conceptual models part I—A discussion of principles. J. Hydrol. 10 (3), 282–290.

Ndomba, P., Mtalo, F., Killingtveit, A., 2005. The suitability of SWAT model in sediment yield modeling for ungauged catchments: a case of Simiyu River subcatchment, Tanzania, 3rd International SWAT Conference, pp. 61-69.

Neitsch, S.L., Williams, J.R., Arnold, J.G., Kiniry, J.R., 2011. Soil and water assessment tool theoretical documentation version 2009, Grassland, Soil and Water Research Service, Temple, Texas, USA.

Neitsch, S.L.A., Kiniry, J.G., Williams, J.R., 2011. Soil and Water Assessment Tool Theoretical Documentation Version 2009. Texas Water Resources Institute.

Oudin, L., Andréassian, V., Mathevet, T., Perrin, C., Michel, C., 2006. Dynamic averaging of rainfall-runoff model simulations from complementary model parameterizations. Water Resour. Res. 42 (7).

Parmele, L.H., 1972. Errors in output of hydrologic models due to errors in input potential evapotranspiration. Water Resour. Res. 8 (2), 348–359.

Parton, W.J., Ojima, D.S., Cole, C.V., Schimel, D.S., 1994. A general model for soil organic matter dynamics: sensitivity to litter chemistry, texture and management. Quantitative modeling of soil forming processes (quantitativemod): 147-167.

Patankar, S., 1980. Numerical Heat Transfer and Fluid Flow. CRC Press.
Perrin, C., Michel, C., Andréassian, V., 2001. Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments. J. Hydrol. 242 (3–4), 275–301.
Pomeroy, J., et al., 2007. The cold regions hydrological model: a platform for basing process representation and model structure on physical evidence. Hydrol. Proces: Int. J. 21 (19), 2650–2667.
Qi, J., et al., 2019a. Effects of surface runoff and infiltration partition methods on hydrological modeling: A comparison of four schemes in two watersheds in the northeastern US. J. Hydrol., 124415.
Qi, J., Li, S., Bourque, C.P., Xing, Z., Fan-Rui, M., 2018a. Developing a decision support tool for assessing land use change and BMPs in ungauged watersheds based on decision rules provided by SWAT simulation. Hydrol. Earth Syst. Sci. 22 (7), 3789–3806.
Qi, J., et al., 2017a. Modifying SWAT with an energy balance module to simulate snowmelt for maritime regions. Environ. Modell. Software 93, 146–160.
Qi, J., et al., 2016a. Assessing an enhanced version of SWAT on water quantity and quality simulation in regions with seasonal snow cover. Water Resour. Manage. 30 (14), 5021–5037.
Qi, J., et al., 2016b. A new soil-temperature module for SWAT application in regions with seasonal snow cover. J. Hydrol. 538, 863–877.
Qi, J., Li, S., Yang, Q., Xing, Z., Meng, F.-R., 2017b. SWAT setup with long-term detailed landuse and management records and modification for a micro-watershed influenced by freeze-thaw cycles. Water Resour. Manage. 31 (12), 3953–3974. https://doi.org/10.1007/s11269-017-1718-2.
Qi, J., Zhang, X., Cosh, M.H., 2019b. Modeling soil temperature in a temperate region: A comparison between empirical and physically based methods in SWAT. Ecol. Eng. 129, 134–143.
Qi, J., et al., 2019c. A coupled surface water storage and subsurface water dynamics model in SWAT for characterizing hydroperiod of geographically isolated wetlands. Adv. Water Resour. 131, 103380.
Qi, J., et al., 2018b. Assessing the performance of a physically-based soil moisture module integrated within the Soil and Water Assessment Tool. Environ. Modell. Software 109, 329–341.
Sakaguchi, A., et al., 2014. Development and evaluation of a paddy module for improving hydrological simulation in SWAT. Agric. Water Manag. 137, 116–122.
Seligman, N., Keulen, H., 1980. 4.10 PAPRAN: A simulation model of annual pasture production limited by rainfall and nitrogen. Simulation of nitrogen behaviour of soil-plant systems, 192.
Shamseldin, A.Y., O'Connor, K.M., Liang, G., 1997. Methods for combining the outputs of different rainfall–runoff models. J. Hydrol. 197 (1–4), 203–229.
Sivapalan, M., 2003. Prediction in ungauged basins: a grand challenge for theoretical hydrology. Hydrol. Process. 17 (15), 3163–3170.
Srinivasan, R., Zhang, X., Arnold, J., 2010. SWAT ungauged: hydrological budget and crop yield predictions in the Upper Mississippi River Basin. Trans. ASABE 53 (5), 1533–1546.
Sun, W., Ishidaira, H., Bastola, S., 2010. Towards improving river discharge estimation in ungauged basins: calibration of rainfall-runoff models based on satellite observations of river flow width at basin outlet. Hydrol. Earth Syst. Sci. 14 (10), 2011.
Tuppad, P., Douglas-Mankin, K., Lee, T., Srinivasan, R., Arnold, J., 2011. Soil and Water Assessment Tool (SWAT) hydrologic/water quality model: Extended capability and wider adoption. Trans. ASABE 54 (5), 1677–1684.
Wagner, P., Kumar, S., Fiener, P., Schneider, K., 2011. Hydrological modeling with SWAT in a monsoon-driven environment: experience from the Western Ghats, India. Trans. ASABE 54 (5), 1783–1790.
Wu, J., Tanaka, K., 2005. Reducing nitrogen runoff from the upper Mississippi River basin to control hypoxia in the Gulf of Mexico: easements or taxes? Mar. Resour. Econ. 20 (2), 121–144.
Wu, Y., Liu, S., Abdul-Aziz, O.I., 2012. Hydrological effects of the increased CO 2 and climate change in the Upper Mississippi River Basin using a modified SWAT. Clim. Change 110 (3–4), 977–1003.
Xia, Y., et al., 2012. Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. J. Geophys. Res.: Atmos. 117 (D3).
Yang, Q., Zhang, X., 2016. Improving SWAT for simulating water and carbon fluxes of forest ecosystems. Sci. Total Environ. 569, 1478–1488.
Yang, Q., et al., 2017. Enhancing the soil and water assessment tool model for simulating N2O emissions of three agricultural systems. Ecosyst. Health Sustainability 3 (2), e01259.
Zhang, C., Li, S., Qi, J., Xing, Z., Meng, F., 2017. Assessing impacts of riparian buffer zones on sediment and nutrient loadings into streams at watershed scale using an integrated REMM-SWAT model. Hydrol. Process. 31 (4), 916–924.
Zhang, X., Izaurralde, R.C., Arnold, J.G., Williams, J.R., Srinivasan, R., 2013. Modifying the soil and water assessment tool to simulate cropland carbon flux: model development and initial evaluation. Sci. Total Environ. 463, 810–822.
Zhang, X., Srinivasan, R., Debele, B., Hao, F., 2008. Runoff simulation of the headwaters of the Yellow River using the SWAT model with three snowmelt algorithms. JAWRA 44 (1), 48–61.
Zhao, F., et al., 2020. Multi-environmental impacts of biofuel production in the US Corn Belt: A coupled hydro-biogeochemical modeling approach. J. Cleaner Prod. 251, 119561.