Analysis of Alternative Climate Datasets and Evapotranspiration Methods for the Upper Mississippi River Basin using SWAT within HAWQS

Manyu Chen[1,2], Philip W. Gassman[2,*], Raghavan Srinivasan[3], Yuanlai Cui[1], Raymond Arritt[4]

[1]State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, China

[2]Center for Agricultural and Rural Development, Iowa State University, Ames, Iowa, 50011-1070, USA

[3]Spatial Sciences Laboratory, Department of Ecosystem Science and Management, Texas A&M University, College Station, TX 77843-2120, USA

[4]Department of Agronomy, Iowa State University, Ames, Iowa, 50011-1051, USA

*Contact author: email: pwgassma@iastate.edu; phone: 515-294-6313

Abstract: The Upper Mississippi River Basin (UMRB) is an important ecosystem located in the north central U.S. that is experiencing a range of ecological stresses. Simulation models are a key tool for evaluating UMRB ecosystem-related problems including the Soil and Water Assessment Tool (SWAT). This study focused on the application of SWAT within Hydrologic and Water Quality System (HAWQS) on-line platform, which supports rapid development of SWAT projects for U.S. watersheds. To date, testing of SWAT models developed in HAWQS has been extremely limited including the effects of different climate-related inputs. Thus the focus of this study was to test the effects of the: (1) Hargreaves (HG) and Penman-Monteith (PM) PET methods, and (2) Livneh climate dataset, which exists external to HAWQS, and the National Climatic Data Center (NCDC) and Parameter-elevation Regressions on Independent Slopes Model (PRISM) climate datasets, which are available in HAWQS. The Livneh data was found to result in the highest average annual water yield of 380.6 mm when executed with the PM method but the lowest estimated water yield of 193.9 mm in combination with the HG method in response to 23-year uncalibrated simulations. Higher annual ET and PET values were predicted with HG method versus the PM method for all three weather datasets in response to the uncalibrated simulations, due primarily to higher HG-based estimates during the growing season. Based on these results, it was found that the HG method is the preferred PET option for the UMRB. Initial calibration of SWAT was performed using the Livneh data and HG method for three Mississippi River main stem gauge sites, which was followed by spatial validation at 10 other gauge sites located within the UMRB stream network. Overall satisfactory results were found for the calibration and validation gauge sites, with the majority of R^2 values ranging between 0.61 and 0.82, Nash-Sutcliffe modeling efficiency (NSE) values ranging between 0.50 and 0.79, and Kling-Gupta efficiency (KGE) values ranging between 0.61 to 0.84. The results of an additional experimental suite of six scenarios, which represented different combinations of climate data sets and calibrated parameters, revealed that suggested statistical criteria were again satisfied by the different scenario combinations. Overall, the PRISM data exhibited the strongest overall reliability for the UMRB.

Key words: SWAT model; Climate dataset; UMRB; HAWQS; Hargreaves method; Penman-Monteith method

1. Introduction

The Upper Mississippi River Basin (UMRB) was proclaimed both a “nationally significant ecosystem” and a “nationally significant commercial navigation system” in the Water Resources Development Act that was passed by the U.S. Congress in 1986 (Weitzell et al., 2003; USACE, 2016). However, extensive alterations to the UMRB stream system, which began during the middle of the 19th century to support commercial navigation, continue to conflict with ecosystem services goals (USACE, 2016). These stream system modifications, in conjunction with large-scale land use change throughout much of the region, have resulted in a degraded UMRB ecosystem and loss of native aquatic diversity (Weitzell et al., 2003). Other pervasive ecosystem stresses are prevalent in the UMRB stream system including degraded water quality (Bouska et al., 2018; Christianson et al., 2018; Sprague et al., 2011, Jones et al., 2018), and increasing flood levels and damage (Criss and Shock, 2001; Criss and Luo, 2017).

A wide range approaches have been implemented to support efforts to mitigate habitat decline, pollution, flooding, and/or other UMRB ecosystem related problems, including habitat restoration (USACE, 2016), biological and habitat surveys (Weitzell et al., 2003), nutrient loss reduction strategies (Christianson et al., 2018) and in-stream monitoring (Royer et al., 2006; Sprague et al., 2011; Jones, 2018). Applications of simulation models have also emerged as key tools in evaluating UMRB ecosystem-related problems including The Soil and Water Assessment Tool (SWAT) ecohydrological model (Arnold et al., 1998; 2012; Williams et al., 2008; Bieger et al., 2017) which has been extensively applied worldwide for a wide range of watershed scales, environmental conditions and water resource problems (Gassman et al., 2007; 2014; Tuppad et al., 2011; Krysanova and White, 2015; Bressiani et al., 2015; Tan et al., 2019; CARD, 2019). SWAT has been used to analyze several UMRB-focused water quantity and/or water quality issues as described in numerous previous studies (e.g., Jha et al., 2004; 2006; Demissie et al., 2012a; Srinivasan et al., 2010; Rabotyagov et al., 2010; Kling et al., 2014; Qi et al., 2019). The model has also been applied in dozens of studies for smaller stream systems within the UMRB (e.g., Hanratty and Stefan, 1998; Vaché et al., 2002; Kirsch et al.., 2002; Chaplot et al., 2004; Green and

Wang, 2008; Jha et al., 2010; Beeson et al., 2014; Almendinger et al., 2014; Teshager et al., 2016; Getahun and Keefer, 2016; Gassman et al., 2017; Schilling et al., 2019).

SWAT has been adopted within several decision support tools (DSTs; e.g., Barnhart et al., 2018) including the USEPA Hydrologic and Water Quality System (HAWQS) on-line platform (Yen et al., 2016; USEPA, 2019). The HAWQS platform provides the capability of building SWAT projects relatively quickly for U.S. watersheds of any scale using pre-loaded climate, land use, management, soil, topographic and other pertinent data layers. SWAT models are constructed in HAWQS by using hydrologic unit codes (HUCs) to delineate a study region that have been defined by the USGS and other federal agencies (USGS, 2013). The HUC8, HUC10 or HUC12 levels (USGS, 2013) can be chosen in HAWQS (Yen et al., 2016) which are commonly referred to as 8-, 10- or 12-digit watersheds. Yen et al. (2016) note that "preliminary calibration" have been conducted for parameters incorporated in HAWQS and Barnhart et al. (2018) state that HAWQS is an example of a "widely used and tested watershed-based DST." However, only three studies to date report results of SWAT applications built in HAWQS (Yen et al., 2016; Fant et al., 2017; Yuan et al., 2018), which were supported with limited model testing. The Initial testing results of these HAWQS-based SWAT models strongly suggest that more in-depth testing is needed to better establish ideal input parameter values and/or sources for different regions including the UMRB, for SWAT projects built in HAWQS.

Thus it is extremely relevant in this study to investigate further testing of a UMRB SWAT model built in HAWQS, in the context of U.S. Department of Energy (USDOE) sponsored research focused on assessing climate simulations in conjunction with the energy-land-water nexus (USDOE, 2019). An important component of this research is the evaluation of different baseline (historical) measured climate data sources and potential evapotranspiration (PET) methods, within the HAWQS-based SWAT model created for the UMRB. The evaluation of inputs from alternative climate data sources is fundamentally important in determining the accuracy of SWAT for representing hydrological processes of a given watershed system as documented in numerous previous studies (e.g., Roth and Lemann, 2016; Tan et al., 2017; Vu et al., 2018; Qi et al., 2019). Assessment of simulated evapotranspiration (ET) is also critical due to the influence of ET on the overall hydrologic balance, crop yields and agricultural management at the watershed scale (Alemayehu et al., 2016; Aouissi et al., 2016; Tie et al., 2018; Valle Júnior et al., 2020). The assessment of climate data sources and PET methods in this study is foundational for future applications of HAWQS-based UMRB SWAT models within the USDOE sponsored research project and other potential applications focused on various ecosystem related problems as described above.

Previous evaluations of climate data effects on SWAT UMRB hydrologic predictions are limited to the research reported by Qi et al. (2019), who compared the impacts on estimated streamflow between National Climatic Data Center (NCDC) data (NOAA, 2019) data and two versions of the NASA North American Land Data Assimilation System Phase Two (NLDAS2) data (Xia et al., 2012). The evaluation of climate data sources in this study included three data sets that spanned > 20 years: (1) NCDC data and the Parameter-elevation Regressions on Independent Slopes Model (PRISM) data (PCG, 2019; USDA-NRCS, 2019), which are both available in HAWQS, and (2) Livneh data (Livneh, 2013), which must be accessed external to HAWQS (ESRL, 2019). In addition, two other key inputs to the SWAT UMRB assessment were evaluated in this study. First, the effects of two PET methods on streamflow and ET estimates were also investigated which are standard options available in SWAT: Hargreaves (Hargreaves and Samani, 1985) and Penman-Monteith (Monteith, 1965; Allen et al., 2006). Second, the influence of different land use types on UMRB hydrology were analyzed for baseline conditions, in combination with the three climate data sets and two PET methods,.

In summary, the specific objectives of this research were to compare: (1) the temporal and spatial differences of the NCDC, PRISM and Livneh climate data sets, (2) the impacts of using the Hargreaves (HG) versus the Penman-Monteith (PM) PET methods on SWAT UMRB streamflow and ET estimates, (3) the impacts of different land use types on UMRB baseline hydrology, and (4) the effects of the three different climate datasets in combination with the two PET methods on UMRB streamflow estimates.

2. Description of the Study Region

The UMRB originates from Lake Itasca in northern Minnesota and outlets at the confluence of the Ohio and Mississippi Rivers near the town of Cairo in southern Illinois (Fig. 1). The UMRB stream system drains a total of 491,700 km^2, which includes large portions of five states (Illinois, Iowa, Minnesota, Missouri and Wisconsin) and small portions of three other states (Indiana, Michigan and South Dakota). Nearly 1,400 km of the Upper Mississippi River is commercially navigable, between St. Paul, Minnesota and the confluence with the Ohio River, which is facilitated by a system of 29 locks/dams and dredging to maintain a minimum channel depth of 2.7 m (UMRBA, 2019). The region is designated as code 07 at the 2-digit watershed level (USGS, 2013) and is further delineated into 131 8-digit subwatersheds and 5,729 12-digit subwatersheds (Panagopoulos et al., 2015). The basin outlet is often assumed to be a gauge site located near Grafton, Illinois in SWAT modeling studies, which drains an area of 447,802 km^2 (119 8-digit watersheds) and is located just upstream of the confluence of the Mississippi and Missouri Rivers.

The major land use in the UMRB is cropland (44.7%), which is dominated by rotations of corn (27.5%) and soybean (17.2%). There are smaller areas of wheat, oats and other crops in the UMRB but those are excluded in this HAWQS modeling framework. Other important land use categories include forest (20.1%), grassland (16.2%), water and wetlands (9.9%) and urban/developed areas (9.1%). Annual precipitation averaged over 830 mm across the UMRB during the 23-year simulation period (1983 to 2005) used in this study, and ranged from < 600 mm in the northwest part of the basin to > 1,000 mm in the southern area of the basin, depending on the source of precipitation data (Fig. 2). Average daily temperatures in the region generally range from 4.0 °C to 5.5 °C in the northern part of the basin to 11.5 °C to 13.0 °C in the south (Fig. 3). However, distinctly colder average temperatures are indicated by the NCDC dataset for a subset of specific 8-digit watersheds in the northwest and central parts of the UMRB, as compared to the PRISM and Livneh datasets (these lower temperatures are likely an anomaly in the NCDC data as discussed in section 4.2). The soil types range from heavy, poorly drained clay soil to light, well-drained sands, with silty loam and loam soils covering about 66% of the total UMRB area

Figure 1. The location of the Upper Mississippi River Basin (UMRB) with 8-digit watersheds, and gauge sites over the study region.

(Demissie et al., 2012b). The topography is characterized by flat to gently rolling terrain, with 55% of the area having less than a 2% slope and an average elevation of 280 m.

3. Previous SWAT Applications Reported for the UMRB

A total of 41 previous studies (Table 1) were documented (CARD, 2019) that reported simulation of the UMRB using SWAT. Four of the studies report SWAT applications for the entire MARB that included the UMRB as a major subregion (Kannan et al., 2019; Santhi et al., 2014; White et al., 2014; Yuan et al., 2018) and three other studies report Corn Belt region results that include SWAT analyses for the combined UMRB and Ohio-Tennessee River Basin (OTRB) systems (Kling et al., 2014; Panagopoulos et al., 2015; 2017). The studies listed in Table 1 focused on a range of themes including

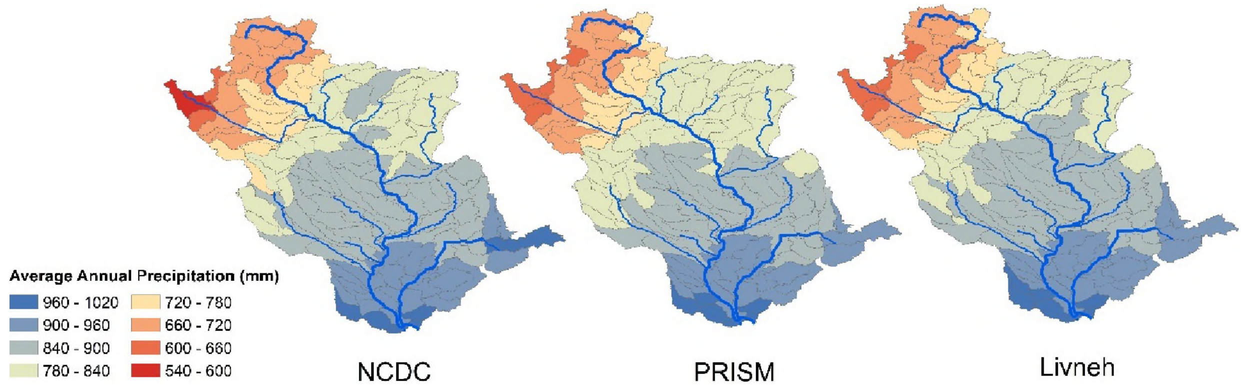

Figure 2. Distribution of average annual precipitation amounts values by climate data set and 8-digit watersheds across the UMRB for the time period of 1983 to 2005.

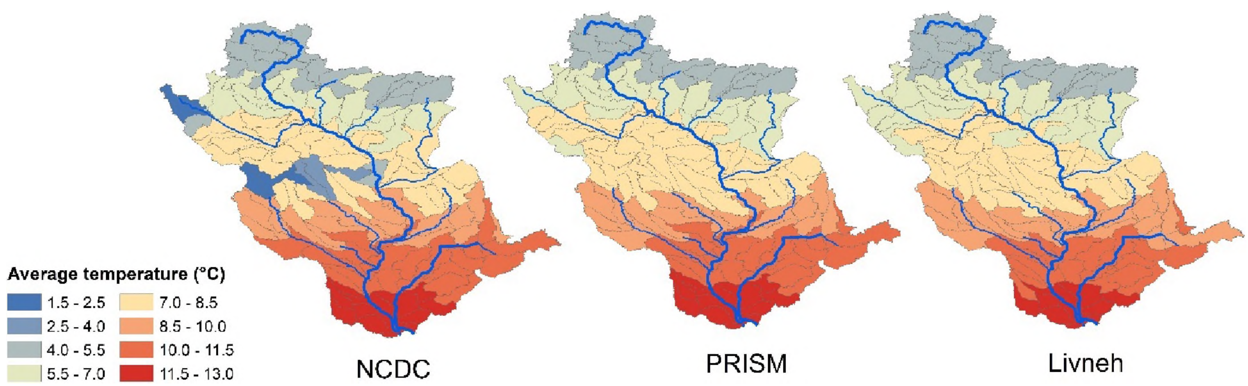

Figure 3. Distribution of average daily temperature values by climate data set and 8-digit watersheds across the UMRB for the time period of 1983 to 2005.

model testing, climate change impacts on streamflow or water quality, and BMP, land use change or biofuel cropping production impacts on water quality. . The only UMRB application that reported the effects of different climate sources on UMRB streamflow was Qi et al. (2019) and none of the studies reported the impacts of different PET methods.

There is a considerable distribution of reported SWAT subwatershed and HRU structures (Table 1), ranging from a relatively coarse delineation of 119 subwatersheds and 474 HRUs (Takle et al., 2005) to an extremely detailed subdivision of 5,732 subwatersheds and 136,079 HRUs (Feng et al., 2018). Most of the studies state that model testing was performed using streamflow data collected at a gauge site located near Grafton, IL. A smaller subset of studies report expanded calibration and/or validation at

Table 1. Focus, structure, SWAT version and total gauge sites used for model testing in previous SWAT-based UMRB studies

Study	Focus of study	Subwatersheds/ HRUs	SWAT version	Testing at Grafton, IL	Streamflow testing sites Cal/Val	Pollutant testing sites Cal/Val
Arnold et al., 2000	Hydrologic testing and evaluation	131/NR[a,b]	NR	No[c]	0/1	0/0
Deb et al., 2015	Biofuel crops and climate change	131/14,568	2009	Yes	0/1	0/1
Demissie et al., 2012a[d]	Biofuel crop impacts on water quality	131/14,200	2005	Yes	3/7	3/7
Eisner et al., 2017[e]	Climate change impacts from 9 hydrologic models used for major river systems	NR	NR	No[c]	1/1	0/0
Feng et al., 2017	Suitability of marginal land for biofuel crop production	NR	NR	No	0/0	0/0
Feng et al., 2018	Biofuel crop production on marginal land	5,732/136,079	2012	Yes	13/13	0/0
Gu et al., 2015	Biofuel crop impacts on water quality	131/2,730	2005[f]	No[c]	0/0[g]	0/0
Gosling et al., 2017[e]	Predicted changes in runoff due to multiple global warming scenarios	NR	NR	No[c]	1/1	0/0
Hattermann et al., 2017[e]	Climate change impacts on hydrological model output for major river systems	NR	NR	No[c]	1/1	0/0
Huang et al., 2017[e]	Comparison of 9 hydrologic models that were applied to major river systems	NR	NR	No[c]	1/1	0/0
Jha et al., 2004	Climate change impacts on streamflow	119[g]/474[h]	NR	Yes	1/1	0/0
Jha et al., 2006	Climate change impacts on streamflow	119[g]/NR	2000[f]	Yes	1/1	0/0
Jha et al., 2015	Climate change impacts on nitrate loads	131/18,000	2005	Yes	1/0	0/0
Kannan et al., 2008	Automatic calibration approach	131/NR	NR	No	0/0	0/0
Kannan et al., 2019	Calibration approaches/issues	NR	2000[f]	Yes	5/5	0/0
Kling et al., 2014	BMP impacts on water quality	5,279/5,279[i]	2009[f]	No[j]	0/0[j]	0/0[j]
Krysanova & Hattermann, 2017[e]	Summary of comparing hydrologic and climate models for 12 major river systems	NR	NR	No[c]	1/1	0/0
Li et al., 2017	Drought impacts on ecosystem services	157/6,686	2009[f]	Yes	13/13	0/0

Table 1. Continued

Study	Focus of study	Subwatersheds/ HRUs	SWAT version	Testing at Grafton, IL	Streamflow testing sites Cal/Val	Pollutant testing sites Cal/Val
Li et al., 2019	Climate change impacts on BMP effects	NR	2009[f]	Yes	13/13	6/6
Panagopoulos et al., 2014	Climate change impacts on BMP effects	5,279/5,279[i]	2012	Yes	12/0	0/0
Panagopoulos et al., 2015	Calibration and validation approach	5,279/5,279[i]	2012	Yes	12/12	6/6
Panagopoulos et al., 2017	Biofuel crop impacts on water quality	5,279/5,279[i]	2012	No[j]	0/0[j]	0/0[j]
Qi et al., 2019a	Climate source impacts on streamflow	131/NR	2012	Yes	11/11	0/0
Qi et al., 2019b	Enhanced freeze-thaw cycle processes	131/14,568	2012[k]	Yes	1/0	0/0
Qi et al., 2020	Water quality testing and evaluation	131/14,568	2012[l]	Yes	0/3[m]	0/3[m]
Rabotyagov et al., 2010	Analyses of least cost of BMPs	131/NR	NR	No	0/0	0/0
Rajib & Merweade, 2017	Land use change impacts on hydrology	260/NR	NR	Yes	10/2	0/0
Santhi et al. 2008	Calibration and validation approach	131/NR	NR	No	0/0[n]	0/0
Santhi et al., 2014[o]	Phosphorous transport in stream system	NR[p]	2005[q]	Yes	5/5	5/5
Secchi et al., 2011	Land use change impacts on water quality	131/2,730	2005[f]	Yes	1/1	1/1
Srinivasan et al., 2010	Uncalibrated baseline streamflow testing	131/14,568	2009[r]	Yes	0/11[m]	0/0
Takle et al., 2005	Climate change impacts on streamflow	119/474	2000[f]	Yes	0/0[s]	0/0
Takle et al., 2010	Climate change impacts on streamflow	NR	2000[f]	Yes	0/0[s]	0/0
Vetter et al., 2017[e]	Climate change uncertainty within hydrologic models used for major rivers	NR	NR	No[c]	1/1	0/0
Wang et al., 2011[o]	Crop production impacts on sediment loss	NR[p]	2005	Yes	5/5	0/0
White et al., 2014[o]	Nutrient loads delivered to stream system, including effects of BMPs	131[t]/NR[b]	2005[o]	Yes	7/4	7/4
Whittaker et al., 2015	Land use change impacts on water quality	131/NR	2009[u]	Yes	0/0/	0/1
Wu et al., 2012a[d]	Biofuel crop impacts on water quality	131/14,200	2005	Yes	3/7	3/7
Wu et al., 2012b	Climate change impacts on streamflow	187[g]/972	99.2[v]	Yes	1/1	0/0

Table 1. Continued

Study	Focus of study	Subwatersheds/ HRUs	SWAT version	Testing at Grafton, IL	Streamflow testing sites Cal/Val	Pollutant testing sites Cal/Val
Wu & Tanaka, 2005	Reducing nitrate loads in stream system	118[g]/1,410	2000	Yes	0/1	0/1
Yuan et al., 2018	Estimating nitrate loads for Mississippi stream system using multiple models	131[t]/NR	2012[w]	NR	NR	NR

[a]NR = not reported.
[b]Specific HRU data are not reported in these studies; however Arnold et al. (2000) note that approximately 16 HRUs were delineated per subwatershed and White et al. (2014) state that a range of 40 to 99 HRUs were delineated in a given subwatershed.
[c]The gauge used for model testing in these studies was located near Alton, IL, which is located several km south of Grafton, IL (and below the confluence of the Mississippi and Missouri Rivers) and captures a drainage area of 444,185 km^2 (Huang et al., 2017).
[d]The model structure, gauge testing sites and model testing results used in these studies were reported in Demissie et al. (2012b).
[e]These six studies were part of a special issue published in Climatic Change (https://link.springer.com/journal/10584/141/3/page/1). Results of applying SWAT for the UMRB are reported in these six studies. SWAT testing statistics are reported in supporting documentation that can be accessed at Huang et al., 2017b.
[f]The SWAT model version was inferred from citations to SWAT documentation reported in the respective study.
[g]These SWAT models were constructed with the outlet at Grafton, IL and thus excluded the subwatersheds that drain to the Mississippi River below Grafton.
[h]Inferred from information reported in Takle et al. (2005).
[i]A dominant HRU approach was used that resulted in one HRU per subwatershed.
[j]Model testing was based on the results reported in Panagopoulos et al. (2015).
[k]Standard SWAT model used in study was SWAT2012, Revision 664; modified version of SWAT called TSWAT.
[l]A standard SWAT2012 version
[m]Uncalibrated simulations were performed in these studies; Srinivasan et al. (2010) list "calibrated statistics" in Table 9 of their study for comparison purposes.
[n]Streamflow testing was not reported although mean NSE and R^2 statistics were reported for water balance indicators determined for all 131 subwatersheds.
[o]An interface between the Agricultural Policy/Environmental eXtender (APEX) model (Gassman et al., 2010) and SWAT was used in these studies.
[p]It is inferred that the model structure used in these studies is based on what is reported in White et al. (2014).
[q]Model version based on personal communication with M. White, Grassland Soil and Water Research Laboratory, USDA-ARS, Temple, TX.
[r]Model version was not directly reported in study but confirmed in later study published by Deb et al. (2014).
[s]Model testing was based on previously reported information in either Jha et al. (2004) or Jha et al. (2006).
[t]Almost all of the entire MARB were simulated in these studies; i.e., White et al. (2014) simulated a total of 848 USGS 8-digit watersheds (USGS, 2013) and Yuan et al. simulated total of 821 USGS 8-digit subwatersheds. It is assumed that 131 of the 8-digit watershed were used to represent the UMRB in both studies.
[u]SWAT model structure based on previously developed model described by Srinivasan et al. (2010).
[v]The authors report using a modified version of SWAT 99.2 in their study.
[v]The model version was not directly reported in the study but the modeling system was constructed via the Hydrologic and Water Quality System (HAWQS) that currently provides the option of using four different releases of the SWAT2012 model (Srinivasan, 2019).

additional gauge sites (Table 1). Ten of the studies report some level of pollutant load testing at specific gauge sites (Table 1). Some of the studies provide comparisons of SWAT predicted loads versus observed or other estimated loads at the 8-digit watershed level (e.g., Arnold et al., 2000; Kannan et al., 2008; White et al., 2014; Jha et al., 2015).

The majority of the studies report both calibration and validation testing results at Grafton and/or other gauge locations, and include evaluations based on the Coefficient of Determination (R^2) and/or Nash-Sutcliffe Efficiency (NSE) statistics (Krause et al., 2005). Tabulation of computed R^2 and NSE statistics, which measure how accurately simulated streamflows replicated measured streamflow, are shown by frequency in Table 2 for a daily time step (usual SWAT time step), and aggregated monthly and annual time periods. Roughly 90% of the NSE and R^2 statistics represented in Table 2 exceed 0.5 and 0.6, respectively, which satisfies satisfactory or better model performance criteria suggested by Moriasi et al. (2007; 2015). The distribution of statistics in Table 2 also generally mirror previous similar statistical compilations reported in several review studies (Gassman et al., 2007, 2014; Tuppad et al., 2011; Bressiani et al., 2015; Tan et al., 2019). Some of the weaker validation statistics reflect more stringent applications of un-calibrated SWAT models reported by Srinivasan et al. (2010) and Qi et al. (2020).

The composite results of previous studies (Table 2) confirm that applications of different versions of SWAT have been generally successful in replicating observed streamflows at Grafton, IL and at other gauge sites, for both calibration and validation. The majority of model testing was performed using a split-time approach (Arnold et al., 2012), where calibration and validation were conducted for the same gauge locations based on observed streamflow data collected during two different time periods. Spatial validation, where calibration is performed for different gauge sites versus the gauges used for validation, was performed only in support of the analyses by White et al. (2014), and for Demissie et al. (2012a) and Wu et al. (2012a) as reported in their corresponding supporting documentation (Demissie et al., 2012b). A spatial validation approach was adopted in this study as described in more detail below.

Table 2. Distribution of statistics comparing simulated streamflows versus measured streamflows that were reported in SWAT UMRB studies by time steps and frequency ranges[a,b]

	Daily				Monthly				Annual			
Frequency	Calibration		Validation		Calibration		Validation		Calibration		Validation	
Range	NSE	R^2	NSE	R^2	NSE	R^2	NSE	R^2	NSE	R^2	NSE	R^2
0.90-1.00					1	3	3	5	7	11	6	27
0.80-0.89					12	23	23	35	6	6	8	9
0.70-0.79	1	1	1	3	43	44	27	29	1		6	3
0.60-0.69		1	2		28	24	19	19			7	
0.50-0.59	1				17	4	13	8	2		4	
0.40-0.49					3	4	10	7			3	
0.30-0.39					1	2	5	2				
0.20-0.29					1		2	1			2	
0.10-0.19							3				2	
0.00-0.09							1					
<0							1					

[a]Data based on the following studies: Arnold et al. (2000), Deb et al. (2015), Feng et al. (2018), Huang et al. (2017b), Jha et al. (2004), Jha et al. (2006), Jha et al. (2015), Kannan et al. (2019), Li et al. (2017), Panagopoulos et al. (2014), Panagopoulos et al. (2015), Qi et al. (2019a; 2019b; 2020); Rajib & Merweade (2017), Santhi et al. (2014), Secchi et al. (2011), Srinivasan et al. (2010), Wang et al. (2011), Wu et al. (2012b)
[b]Calibration was not performed by Srinivasan et al. (2010) and Qi et al. (2020); statistic from those two studies are reported here as validation

4. Methods and Materials

4.1. SWAT/HAWQS model configuration and simulation scenarios

The development of the UMRB SWAT model was performed in HAWQS, which provides interactive web interfaces, maps and preloaded data layers including stream network, land use and land management, soil, climatic, point sources, historical climate, future climate projections, atmospheric deposition and reservoir data (Srinivasan, 2019). The sources of these input data and the date (month/year) are listed in HAWQS (2017). Users can assign preferred parameter values at HRU, subwatershed and/or overall basin levels, respectively. In addition, HAWQS is technically capable of providing preliminarily calibrated parameters as default values, although the level of testing supporting these parameter values is very limited as previously noted. In this study, the default parameters values set by HAWQS for the UMRB SWAT model were considered to be uncalibrated baseline data (Table 3), and output from this baseline model are referred to as uncalibrated results. The files created for this uncalibrated baseline UMRB SWAT model were also downloaded from HAWQS after the initial model construction, which allows additional parameter modification using the SWAT editor program or other external software.

The SWAT model was configured for the UMRB at the 8-digit watershed level within the HAWQS platform, resulting in 119 8-digit watersheds that encompass the previously described 447,802 km^2 area that drains to Grafton, IL (the outlet is the 8-digit watershed identified as HUC07110009). A total of 34,630 HRUs were initially configured within the 119 subwatersheds when the UMRB SWAT model was first constructed within HAWQS. HRU thresholds of 1 km^2 were then applied to the land use, soil type and slope classes to eliminate minor land uses, soils, and slopes in each subwatershed. The application of the thresholds resulted in a total of 30,812 HRUs for the baseline UMRB SWAT model and subsequent calibrated versions of the UMRB model.

The SWAT simulations were performed from 1981 to 2005; the first 2 years served as an initialization period. This 25-year simulation period reflects a consistent time period available in all three data sets (see Section 4.3). Two sets of six simulations each were then performed as scenarios (Table 4) The first set of simulations were executed without calibration using the baseline UMRB model, to provide an initial comparison of water balance and streamflow estimates between the three climate data sets and two ET methods that were not influenced by any adjustments in SWAT input parameters. These scenarios were based on the baseline UMRB SWAT model that was executed with the HAWQS default input parameters (Table 3). This allowed the weather inputs (including daily precipitation, daily maximum temperature and minimum temperature) to be held constant while varying the PET methods, so that the effects of the PET methods on the hydrologic outputs can be discerned. The initial uncalibrated SWAT model was constructed using the Penman-Monteith (PM) PET method, which is the default PET option that is used in HAWQS (scenarios PRISM(PM), NCDC(PM) and Livneh(PM) in Table 4). The uncalibrated SWAT model was then executed with the HG PET method (scenarios PRISM(HG), NCDC(HG) and Livneh(HG) in Table 4), to provide a further basis of comparison (Section 5.1).

The second phase of simulations were based on calibration and validation, which was initially performed using the Livneh climate data and HG PET method. The Livneh data were chosen for the initial calibration due to the fact that the data set has served in a historical climate role for a suite of downscaled climate projections (Pierce, 2016), that may be analyzed as part of the broader research initiative (DOE, 2019). The HG method was selected because the annual ratio of ET/precipitation was > 0.7 for HG method versus approximately 0.6 for PM method. The HG method ET/precipitation ratio of 0.7 was more consistent with the UMRB region ratio reported by Liu et al. (2013), who estimated ET and runoff for the major basins that contribute steamflow to the Gulf of Mexico. Calibration was performed for the three calibration gauge sites shown in Fig 1 and Table 5: St. Paul, Clinton and Grafton. Spatial validation was then performed by performing an additional simulation with the calibrated model, without any further adjustments to the SWAT input parameters, and comparing simulated versus observed streamflows at the other 10 “hydrologically independent” gauge sites (Fig 1 and Table 5). Each of the 10 hydrologically independent sub-regions corresponds to either the most upstream part of the main stem (Royalton for Mississippi River) or a major tributary flowing into the main stem (i.e., the Skunk, St. Croix, Chippewa, Rock, Wisconsin, Iowa, Des Moines, Minnesota or Illinois Rivers). Table 5 summarizes the information related to the monitoring points.

Following calibration and validation, six additional experimental scenarios (Table 4) were evaluated, which were again based on monthly streamflow output from 1983 to 2005. These scenarios provided further assessment of the performance of SWAT in response to the three climate datasets, HG method, and calibrated parameters listed in Table 3 (that are described in more detail in Section 5.3). These second set of six scenarios were further split into two subsets, which were demarcated as follows: (1) the first subset of three scenarios was based on the calibrated parameters (Table 3) obtained with the previously described Livneh climate dataset and HG method (scenario Livneh-calibrated, in Table 4), versus (2) a second set of three scenarios that were performed using calibrated parameters (Table 3)

Table 3. Calibration input parameters, default values in HAWQS, allowable ranges, type of calibration adjustment and final calibrated values.

Parameters	Description	Type of change[a]	Default values in HAWQS	Allowable ranges		Calibrated value for Livneh	Calibrated value for PRISM
				Min	Max		
CN2	Initial SCS runoff curve number for moisture condition Ⅱ	R	22 - 84	-0.1	0.1	0.0985	0.0526
ALPHA_BF	Baseflow alpha factor (1/days)	V	0.023 - 0.85	0	1	0.9	0.999
GW_DELAY	Groundwater delay time (days)	A	25 - 323	-30	90	-12	6.880
GWQMN	Threshold depth of water in shallow aquifer required for return flow to occur (mm)	A	0.7 - 900	-1000	1000	-422	-577
GW_REVAP	Groundwater "revap" coefficient	V	0.01 - 0.1066	0.02	0.1	0.04	0.053
RCHRG_DP	Deep aquifer percolation fraction	A	0.01 - 0.33	-0.05	0.05	0.027	0.044
REVAPMN	Threshold depth of water in the shallow aquifer for "revap" or percolation to the deep aquifer to occur (mm)	A	264.6; 500	-750	750	206	-24
ALPHA_BF_D	Alpha factor for groudwater recession curve of the deep aquifer (1/day)	V	0	0	1	0.25	0.32
ESCO	Soil evaporation compensation factor	V	0.808 - 0.98	0.6	0.1	0.92-0.95	0.93-0.98
CANMX	Maximum canopy storage (mm)	V	15.4	0	25	2[b]	2[b]
SLSOIL	Slope length for lateral subsurface flow (m)	V	0	0	150	6.375	28
LAT_TTIME	Lateral flow travel time (days)	V	0	0	200	153	186
SOL_AWC	Available water capacity of the soil layer (mm H_2O/mm soil)	R	0.01 - 0.42	-0.05	0.05	0.038[c]	-0.034[c]
SFTMP	Snowfall temperature (°C)	V	1	-5	1	0.175	3.53
SMTMP	Snow melt base temperature (°C)	V	0.5	0	3	0.68	0.29
SMFMX	Melt factor for snow on June 21 (mmH_20/°C -day)	V	4.5	2	4.5	3.99	4.42
SMFMN	Melt factor for snow on December 21 (mm H_2O/°C -day)	V	4.5	0	2.5	0.07	2.37
TIMP	Snow pack temperature lag factor	V	1	0	1	0.55	0.38

[a]R indicates that an existing parameter value is multiplied by (1+ a given value), V indicates that the existing parameter value is to be replaced by a given value, and A indicates that a given value is added to the existing parameter value.
[b]For various forest (FRSD, FRSE and FRST) landuse.
[c]For the first soil layer.

Table 4. Description of the SWAT scenarios executed with baseline (HAWQS default parameters) versus scenarios that were calibrated parameters.

Scenario name	Climate dataset	PET method	Input parameters	Calibration dataset
PRISM(PM)	PRISM	PM	HAWQS default	NA[a]
NCDC(PM)	NCDC	PM	HAWQS default	NA
Livneh(PM)	Livneh	PM	HAWQS default	NA
PRISM(HG)	PRISM	HG	HAWQS default	NA
NCDC(HG)	NCDC	HG	HAWQS default	NA
Livneh(HG)	Livneh	HG	HAWQS default	NA
Livneh-calibrated	Livneh	HG	Calibrated	Livneh
PRISM-Livneh	PRISM	HG	Calibrated	Livneh
NCDC-Livneh	NCDC	HG	Calibrated	Livneh
PRISM-calibrated	PRISM	HG	Calibrated	PRISM
Livneh-PRISM	Livneh	HG	Calibrated	PRISM
NCDC-PRISM	NCDC	HG	Calibrated	PRISM

[a]NA = not applicable

Table 5. The USGS gauge sites used for streamflow calibration and validation in this study, including location, gauge site IDs, hydrologic units and reported drainage area.

Gauge site	River	State	Used for	USGS Station	Hydrologic Unit	Drainage Area (km^2)
St. Paul	Mississippi	MN	Calibration	05331000	7010206	95312
Clinton	Mississippi	IA	Calibration	05420500	7080101	221703
Grafton	Mississippi	IL	Calibration	05587450	7110009	443665
Augusta	Skunk	IA	Validation	05474000	7080107	11168
St. Croix Falls	St. Croix	WI	Validation	05340500	7030005	16162
Durand	Chippewa	WI	Validation	05369500	7050005	23336
Joslin	Rock	IL	Validation	05446500	7090005	24732
Muscoda	Wisconsin	WI	Validation	05407000	7070005	26936
Royalton	Mississippi	MN	Validation	05446500	7010201	30044
Wapello	Iowa	IA	Validation	05446500	7080209	32375
Keosauqua	Des Moines	IA	Validation	05446500	7100009	36358
Jordan	Minnesota	MN	Validation	05446500	7020012	41958
Valley City	Illinois	IL	Validation	05446500	7130011	69264

obtained with the PRISM climate dataset in combination with the HG method (scenarios PRISM-calibrated, Livneh-PRISM and NCDC-PRISM in Table 4).

This suite of scenarios thus provided an approach of further testing SWAT with the three different climate data sets using two different sets of calibration parameters that represent two different potential baseline climate data sets: Livneh versus PRISM. The Livneh-calibrated and PRISM-calibrated scenarios depict obvious conventional SWAT calibration simulations using the climate data sets that the model was calibrated with. However, the PRISM-Livneh, NCDC-Livneh, Livneh-PRISM and NCDC-PRISM scenarios reflect atypical SWAT simulations that consist of executing the model with a different climate data set than was used for the calibration process. These additional scenarios provide additional insight into the sensitivity and performance of SWAT in response to different climate inputs for the UMRB.

4.2. Description of Climate Datasets

Daily precipitation and temperature data obtained from the NCDC, PRISM and Livneh climate datasets were used to simulate UMRB streamflow. Brief summaries of these datasets are provided below followed by further analysis of the apparent anomalies in the NCDC temperature data revealed by Fig 3.

(1) NCDC: NCDC dataset consists of daily weather variables from the Global Historical Climatology Network (GHCN)-Daily of land-based weather stations. The dataset was developed via processing steps of data collection, quality control, and archival and removal of biases associated with factors such as urbanization and changes in instrumentation through time (Menne et al., 2012). The NCDC dataset in HAWQS spans the time period of 1961 to 2010.

(2) PRISM: PRISM was developed by the PRISM Climate Group at Oregon State University (PCG, 2019) and is officially endorsed by the U.S. Department of Agriculture Natural Resources Conservation Service (USDA-NRCS, 2019). PRISM data are defined on a 2.5 min degree grid, which calculates a climate-elevation regression for each grid cell of digital elevation model (DEM). Stations included in the regression are assigned weights based primarily on the similarity of physiographic characteristics (Daly et al., 2008; Gao et al., 2017). PRISM data are available in HAWQS for the time period of 1981 to 2015.

(3) Livneh: Livneh dataset is derived from observations at NCDC cooperative observer (COOP) stations across the continental United States (CONUS). Both temperature and precipitation were gridded to 1/16° using the synergraphic mapping system (SYMAP) algorithm (Livneh et al., 2013; ESRL, 2019). Long-term daily climatic data are provided in the Livneh dataset for 1981 to 2010 (ESRL, 2019).

4.2.1. UMRB climate data distributions and NCDC temperature data anomalies

The spatial distributions of average annual precipitation and air temperature from 1983 to 2005 (Fig. 2 and Fig. 3) provide further insights regarding the differences between the weather datasets. The trends in spatial distribution of precipitation across the UMRB are similar between the NCDC, PRISM and Livneh datasets (Fig. 2), with highest annual precipitations occurring in the southeast versus the lowest annual precipitation in the northwest. The trends in spatial distribution of annual average temperature across the UMRB are also similar among three weather datasets (Fig.3), with a clear gradient of increasing temperature from the north to south. However, the distribution of the NCDC temperature data reveals that some subwatersheds in the northwest and central part of the UMRB manifest cooler average annual temperatures versus subwatersheds in the most northern part of the region; i.e., subwatersheds located in far eastern South Dakota, southeast Minnesota, northern Iowa and southwest Wisconsin (Fig. 3). These "cooler subwatersheds" do not manifest in the PRISM and Livneh data. Thus, it is likely that these cooler subwatersheds are anomalies in the NCDC data that may be due to errors in the original measured observations and/or that occurred during the interpolation and averaging of the data to create pseudo-stations at the 8-digit watershed level, both of which would have occurred prior to inclusion within HAWQS. In contrast, the PRISM and Livneh data were processed for each subwatershed by using their gridded cell values, which provides a more continuous temperature surface for creating a single set of subwatershed temperature data. The revelation of the apparent NCDC temperature anomalies warrants further review and probable correction of the data. However, it is unlikely that these errors greatly affected estimates of UMRB streamflow.

4.3. PET estimation methods

The PET concept was introduced by Thornthwaite (1948) as part of a climate classification scheme. PET was defined as the rate of evapotranspiration without any limits imposed by the supply of water. Numerous methods have been developed to estimate PET. The PM method (Monteith, 1965; Allen et al., 2006) and HG method (Hargreaves and Samani, 1985) are two of the PET options available in SWAT and were tested in this study.

The PM equation combines components that account for the energy needed to sustain evaporation, the strength of mechanism required to remove the water vapor, and aerodynamic and surface resistance terms. The PM equation is:

$$\lambda E = \frac{\Delta \cdot (H_{\mathrm{net}} - G) + \rho_{air} \cdot c_p \cdot \left[e_z^o - e_z \right] / r_a}{\Delta + \gamma \cdot (1 + r_c / r_a)} \qquad (1)$$

where λE is the latent heat flux density (MJ m^{-2} d^{-1}), E is the depth rate evaporation (mm d^{-1}), Δ is the slope of the saturation vapor pressure-temperature curve de/dT (kPa °C^{-1}), H_{net} is the net radiation (MJ m^{-2} d^{-1}), G is the heat flux density to the ground (MJ m^{-2} d^{-1}), ρ_{air} is the air density (kg m^{-3}), c_p is the specific heat at constant pressure (MJ kg^{-1} °C^{-1}), e_z^o is the saturation vapor pressure of air at height z (kPa), e_z is the water vapor pressure of air at height z (kPa), γ is the psychrometric constant (kPa °C^{-1}), r_c is the plant canopy resistance (s m^{-1}), and r_a is the diffusion resistance of the air layer (aerodynamic resistance; s m^{-1}).

The original HG equation (Hargreaves and Samani, 1985) is the form used in SWAT as follows:

$$\lambda E_0 = \alpha_{pet} \cdot \frac{\Delta}{\Delta + \gamma} \cdot (H_{net} - G) \qquad (2)$$

where λ is the latent heat of vaporization (MJ kg^{-1}), E_0 is the potential evapotranspiration (mm d^{-1}), α_{pet} is a coefficient, Δ is the slope of the saturation vapor pressure-temperature curve de/dT (kPa °C^{-1}), γ is the psychrometric constant (kPa °C^{-1}), H_{net} is the net radiation (MJ m^{-2} d^{-1}), and G is the heat flux density to the ground (MJ m^{-2} d^{-1}).

The PM and HG methods vary considerably in the amount of required inputs. The PM method requires solar radiation, air temperature, relative humidity and wind speed but the HG method requires air temperature only. Daily solar radiation, relative humidity and wind speed inputs were generated by the weather generator within SWAT, because daily precipitation and air temperature are the only measured climatic data available.

4.4. Calibration approach and evaluation criteria

The SWAT-CUP platform (Abbaspour, 2015; SWAT, 2019) is a software package with a web-based interface, which facilitates automatic calibration and/or uncertainty analyses for SWAT applications via manipulation of the large number of text files associated with a typical SWAT project (prior to the release of SWAT+; see Bieger et al., 2017). SWAT-CUP allows users to control the initial range of parameters and supports the most accurate identification of the parameter optimum values by automatic or manual calibration of SWAT projects. There are several algorithms incorporated in SWAT-CUP to help with the automatic calibration process: Sequential Uncertainty Fitting (SUFI-2), Particle Swarm Optimization (PSO), Generalized Likelihood Uncertainty Estimation (GLUE), Parameter Solution (ParaSol), and Markov Chain Monte Carlo (MCMC). The SUFI-2 algorithm was used for model calibration in this study. This algorithm can map all of the uncertainties for each parameter (expressed as uniform distributions or ranges) and attempts to capture most of the measured data within the 95% prediction uncertainty (95PPU) of the model in an iterative process. It requires fewer simulations to complete a calibration/uncertainty project (Yang et al., 2008) and is highly recommended for the calibration of SWAT models (Arnold et al., 2012).

Parallel processing was also used in this study, since it can speed up the calibration process by parallelizing the simulations in SUFI-2. The speed of the parallel processing depends on the characteristics of the computer. For example, if the computer has 8 central processing units (CPUs), the parallel processing module can utilize all 8 CPUs so that a 200-run iteration can be divided into 8 simultaneous runs of 25 each per CPU. For a large-scale SWAT model, the utilization of the parallel processing option results in substantially faster overall simulation times versus using just a single 200-run CPU submission.

The SUFI-2 algorithm was set to identify the optimum parameters by using the Nash-Sutcliffe modeling efficiency (NSE) statistic (Krause et al., 2005) as the key objective function. However, the results were also evaluated according to the coefficient of determination (R^2; Krause et al., 2005), percent bias (PBIAS; Moriasi et al., 2007) and Kling-Gupta efficiency (KGE; Gupta et al., 2009). Values for NSE, R^2 and PBIAS on a monthly scale were evaluated per criteria suggested by Moriasi et al. (2007; 2015); i.e., NSE values $\geq$0.50, R^2 values $\geq$ 0.6 and PBIAS values $\leq\pm$25% (Moriasi et al., 2007) or $\leqslant\pm$ 15% (Moriasi et al., 2015) are judged to be satisfactory. Patil and Stieglitz (2015) implied that simulated values could be regarded as satisfactory with a KGE value >0.6. The KGE statistics are designed to provide an improved criterion by incorporating error compensation for the bias and variability components (Roy et al., 2014; Zhu, et al., 2016). Graphical comparisons between the simulated and measured streamflow values were also used to assess the accuracy of the model output.

5. Results and Discussion

5.1. Climate dataset analysis

Table 6 lists the daily mean temperature (°C), average annual precipitation (mm) and other uncalibrated average annual water balance components (mm) that were predicted for the UMRB using the three weather datasets and two PET methods during the 23-year simulation period. The Livneh data set was found to have the largest average annual amount of precipitation (837.2 mm) and highest daily mean temperature (8.2 °C) among the three weather datasets. In contrast, the respective average annual precipitation for the NCDC and PRISM data sets was 836.1 mm and 831.5 mm, and the respective daily mean NCDC and PRISM temperatures were 8.0 °C and 7.9 °C. Overall, there were small differences in the annual average precipitation (5.7 mm maximum) and temperature (0.3°C maximum).

Table 6. Average annual values (mm) of hydrological components and daily mean temperature (°C) over the UMRB, for the different combinations of climate data and PET methods, based on the applications of the uncalibrated HAWQS SWAT model during 1983 to 2005.

Scenario	Precipitation	Daily mean temperature	surface runoff	lateral flow	Groundwater flow	Soil water	ET	PET	Water yield	ET/P[a]	WY/P[b]
PRISM(PM)	831.5	8.0	204.6	40.9	25.7	267.9	508.2	856.9	323.7	0.56	0.43
NCDC(PM)	836.1	7.9	227.2	46.7	30.5	285.5	470.6	764.2	363.1	0.61	0.39
Livneh(PM)	837.2	8.2	235.7	49.6	35.0	295.2	451.7	707.8	380.6	0.54	0.45
PRISM(HG)	831.5	8.0	141.9	24.0	18.3	231.4	620.4	958.6	212.4	0.74	0.26
NCDC(HG)	836.1	7.9	145.3	26.2	18.7	238.3	615.9	934.9	219.7	0.75	0.26
Livneh(HG)	837.2	8.2	128.1	22.7	18.8	228.9	641.7	961.3	193.9	0.77	0.23

[a]ET/P = the ratio of annual ET/Precipitation
[b]WY/P= the ratio of annual Water yield/Precipitation

The Livneh data also generated the largest amounts of surface runoff, lateral flow, groundwater, soil water and water yield when simulated in combination with the PM method (Table 6). The predicted annual average water yield for the Livneh data was 380.6 mm, versus 323.7 mm for PRISM and 363.1 mm for NCDC. This was due primarily because the Livneh data produced the lowest estimated annual average PET of 707.8 mm among the three weather datasets, as compared to 856.9 mm and 746.2 mm for PRISM and NCDC, respectively. The different weather data set inputs resulted in maximum differences of 56.9 mm in water yield and 149.2 mm in PET for the uncalibrated simulations. These trends are also reflected in the ratios of annual ET/precipitation (ET/P) and annual water yield/precipitation (WY/P) reported in Table 6; e.g., the lowest ET/P and WY/P ratios were found for Livneh(PM) and Livneh(HG), respectively,

Considerably higher annual ET and PET values were estimated when the three weather data sets were simulated in combination with the HG method, resulting in much lower predicted water yield and key water yield components; i.e., surface runoff, lateral flow and groundwater flow (Table 6). There was also considerable variation in the responses of the three weather datasets to the two PET methods, especially for the Livneh data set. The Livneh data resulted in the lowest estimated ET and PET when used in combination with the PM method, but produced the highest ET and PET estimates when simulated with the HG method (Table 6). Consequently, the Livneh data generated the highest and lowest water yields when executed with the PM and HG methods, respectively.

In addition to comparing the spatial distribution of average annual precipitation and temperature, the differences between the monthly mean precipitation and daily mean temperature during the 23-year uncalibrated SWAT simulations are presented in Fig.4. The PRISM data generated smaller amounts of precipitation as compared to the NCDC data in most months, especially during the May to September growing season. The cumulative difference between PRISM and NCDC during the growing season accounted for 85% of the total annual average difference between the two data sets. However, the Livneh data precipitation amounts were larger versus NCDC for most months, except for March, August, November and December. The Livneh data precipitation amount was greater during the growing season,

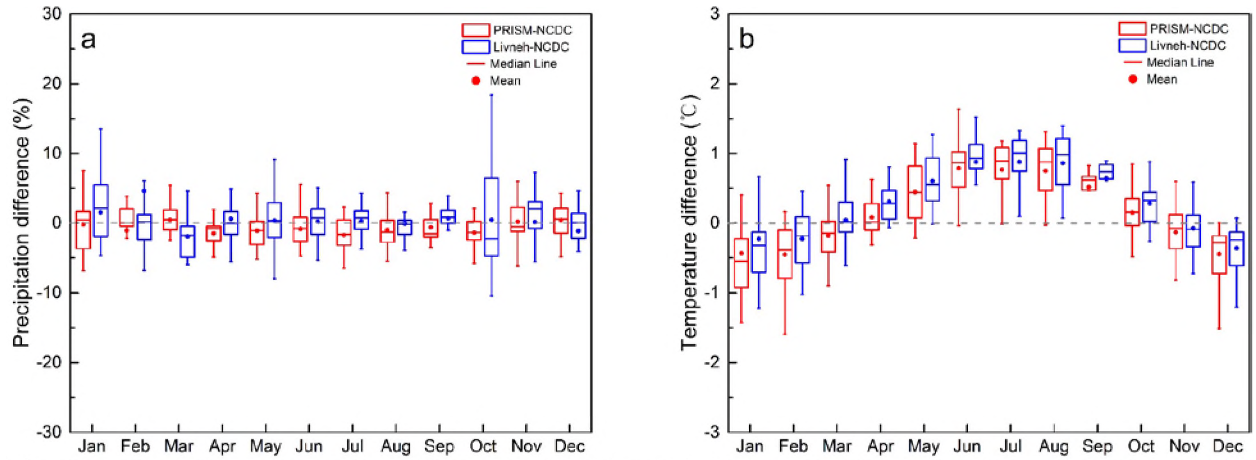

Figure 4. Box plots of monthly differences for (a) precipitation and (b) daily mean temperature. For precipitation, percent difference (%) is displayed. Temperature values displayed are the absolute difference (°C).

while it was slightly smaller during the non-growing season, relative to the NCDC data. With regard to daily temperature, there are distinct seasonal variations between the monthly differences. Both PRISM and Livneh tend to be warmer during the summer months and colder during the winter period, as compared to the NCDC data. Fig. 4 further shows that the Livneh data had the highest daily temperature among three weather datasets in spring and summer, while the PRISM data set results in the lowest temperature in winter.

5.2. Land use data analysis

Table 7 represents average annual precipitation (mm), ET (mm) and water yield to the reach (mm) predicted for different UMRB land uses using the three weather datasets and two PET methods during the 23-year simulation period. The average annual amount of precipitation for urban areas was 847.6 mm, which was greater than the corresponding annual average precipitation levels of 843.7 mm for cropland, 832.1 mm for grassland and 820.8 mm for forest. The differences in the annual average precipitation (from 3.9 to 26.8 mm) are primarily caused by the uneven spatial distribution of both precipitation and land use.

The highest ET levels were predicted for the composite urban areas when the PM method was simulated in combination with the three weather datasets. On average, the annual ET for the urban areas was 526.6 mm, versus 513.1 mm for cropland, 493.4 mm for grassland and 383.5 mm for forest (Table 7).

Table 7. Average annual values (mm) of hydrological components for different land uses, based on the application of the uncalibrated HAWQS SWAT model during 1983 to 2005.

PET method	Land use	Precipitation				ET				Water yield			
		PRISM	NCDC	Livneh	Average	PRISM	NCDC	Livneh	Average	PRISM	NCDC	Livneh	Average
PM	Cropland	843.7	848.0	848.8	846.9	546.9	502.3	490.2	513.1	296.7	344.6	356.5	332.6
	Forest	820.8	826.4	827.5	824.9	409.3	381.5	359.7	383.5	408.5	440.7	461.6	436.9
	Grassland	832.1	835.7	838.0	835.3	525.5	488.6	466.0	493.4	304.4	343.1	366.1	337.9
	Urban area	847.6	852.8	853.5	851.3	556.5	521.4	502.0	526.6	293.4	332.4	351.5	325.8
HG	Cropland	843.7	848.0	848.8	846.9	613.9	602.5	643.4	620.0	229.6	245.1	205.4	226.7
	Forest	820.8	826.4	827.5	824.9	674.5	682.4	695.9	684.3	147.3	144.9	132.6	141.6
	Grassland	832.1	835.7	838.0	835.3	586.8	587.9	613.8	596.2	243.2	245.4	221.9	236.8
	Urban area	847.6	852.8	853.5	851.3	630.3	632.3	660.9	641.2	219.6	222.5	194.8	212.3

In contrast, forest was predicted to have the highest annual water yield of 436.9 mm, followed by grassland (337.9 mm), cropland (332.6 mm) and urban area (325.8 mm). Also, the forest areas tend to generate more runoff in response to the Livneh climate data as compared to the NCDC or PRISM climate data. This is in accord with the variation of runoff for the whole UMRB in Table 6.

Among all land use types, the predicted ET amounts ranked from high to low as follows: forest (684.3 mm) > urban areas (641.2 mm) > cropland (620.0 mm) > grassland (596.2 mm). The highest annual average water yield was produced by grassland (236.8 mm), as compared to 226.7 mm, 212.3 mm and 141.6 mm for cropland, urban areas and forest, respectively. Overall, Table 7 shows that the estimated impacts of land use on the hydrology varied considerably in response to the different climate data sets and/or PET methods. For example, the highest and lowest annual average water yields were estimated to be generated by forested areas when simulated with PM and HG methods, respectively. With HG method, it is indicated that cropland (dominated by corn and soybean) may increase streamflow because of decreased evapotranspiration. The results of HG method are consistent with the work of Zhang and Schilling (2006) that assessed the effect of land use on streamflow in Mississippi River.

5.3. PET methods analysis

Fig 5 shows the monthly variations of ET, PET and water yield predicted by the six uncalibrated SWAT scenarios (Table 4), which again are averaged over the period of 1983 to 2005. The experimental results show that the predicted annual distribution of ET and PET vary quite similarly in response to the two PET methods. Both the ET and PET start rising after January in the winter period, peak in July, and then descend during the remaining fall and winter months (Fig 5a-b). During the growing season (May to September), the SWAT-predicted ET and PET amounts calculated with HG are considerably higher versus the corresponding PM-based estimates. However, the gap between the HG- and PM-estimates is much smaller during the non-growing season and become virtually negligible in winter. The predicted water yield patterns are similar for the two PET methods, except that the peaks occur in June and the ascents to and declines from the peaks are more gradual (Fig 5c). The HG method generated smaller

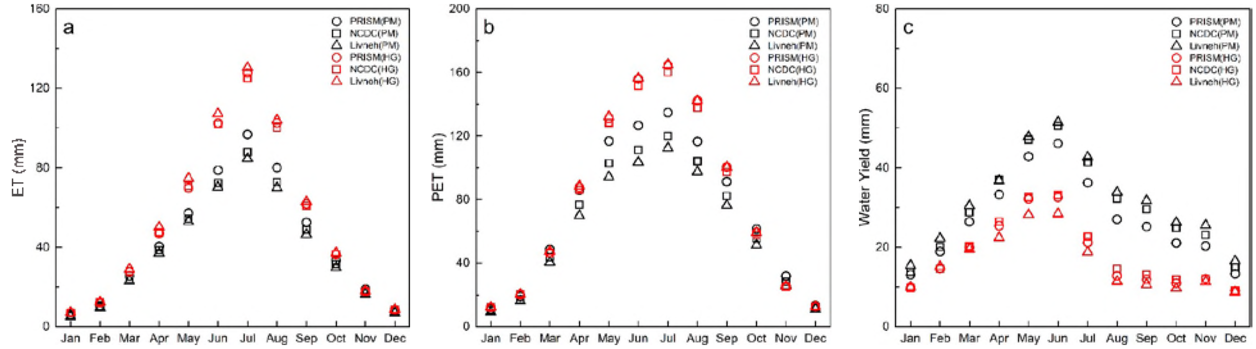

Figure 5. Monthly variations for (a) ET, (b) PET, and (c) Water Yield, based on the applications of the uncalibrated HAWQS SWAT model during 1983-2005. Colors denote the PET method used: black are Penman-Monteith method, red are Hargreaves method.

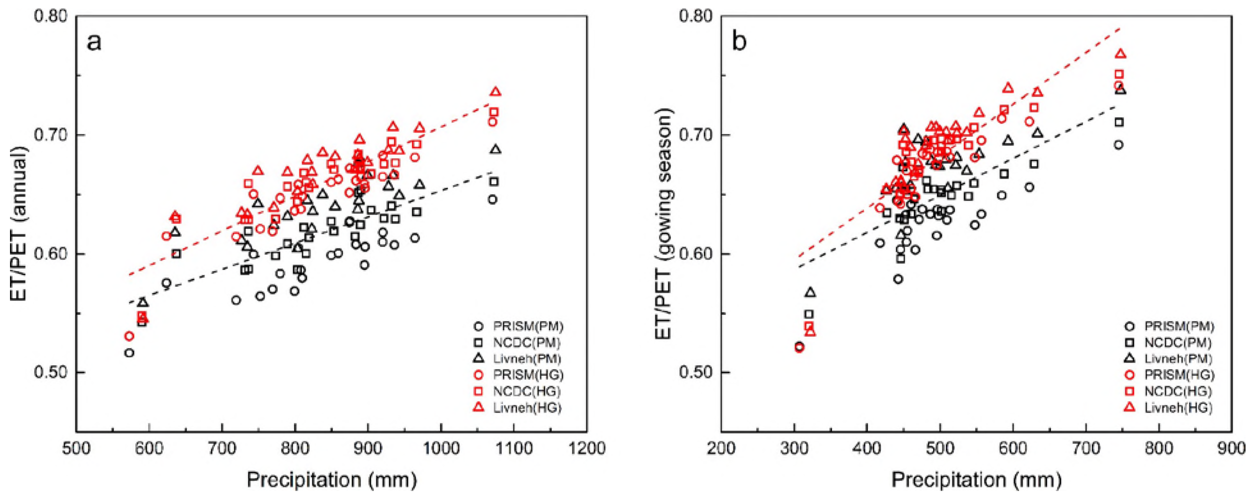

Figure 6. Ratio of ET/PET along with precipitation based on the applications of the uncalibrated HAWQS SWAT model from 1983 to 2005, on an (a)annual basis, and (b) growing season basis. Colors denote the PET method used: black are Penman-Monteith method, red are Hargreaves method.

water yields as compared to the PM method, but the differences are greater during the growing season. The ratio of ET/PET relative to corresponding precipitation from 1983 to 2005 are presented respectively on an annual basis and growing season basis in Fig. 6. Fig. 6a reveals that the ET/PET ratios steadily increased as precipitation increased. Consistently higher ratios of ET/PET were predicted with the HG method across the full range of precipitation amounts. Ratios of ET/PET estimated with the PM method range from 0.52 to 0.69 on annual scales while ratios of ET/PET predicted with the HG method range from 0.53 to 0.74. This result underscores that the HG method results in a higher rainfall use efficiency for the SWAT UMRB model. The ET/PET ratios reveal a similar tendency during the growing season periods (Fig. 6b). The ET/PET ratios are predicted to be higher during the growing season due to the growth of the crops. In general, lower ET/PET ratios imply that crops and other vegetation are not supplied with sufficient water needed for ET and growth, and thus may experience greater water stresses (Chen et al., 2010).

5.4. Flow calibration and validation

Table 3 lists the allowable ranges and types of calibration adjustments that were performed for the selected calibration parameters. Surface runoff and baseflow were calibrated simultaneously. The primary calibration parameters adjusted for surface runoff were the curve numbers (CN2), which represented different land conditions. Seven parameters related to groundwater (ALPHA_BF, GW_DELAY, GWQMN, GW_REVAP, RCHRG_DP and REVAPMN) were adjusted to improve the agreement between observed and simulated streamflows (Table 3). Five snow parameters (SFTMP, SMTMP, SMFMX, SMFMN and TIMP) were also adjusted in this study (Table 3) to better reflect snowmelt magnitude and hydrograph shapes.

The Last two columns in Table 5 represent the two different sets of calibrated parameter values that were obtained for the respective Livneh-calibrated and PRISM-calibrated scenarios. The subset of calibration parameters and allowable ranges were the same for the Livneh- and PRISM-based calibration processes. Because the performance of the PRISM and Livneh data sets were similar within SWAT (Fig. 2 and Fig. 3), consistent adjustment trends occurred for the majority of the parameters. For example, values of CN2, ALPHA_BF and RCHRG_DP increased for both scenario calibrations relative to default values in HAWQS. However, different trends in the final calibrated values resulted for a smaller subset of parameters between the two calibration phases; e.g. GW_DELAY and REVAPMN, where the calibrated values decreased for one calibration phase versus increased values for the other calibration phase (both calibration processes resulted in positive values for both parameters). This does not mean those two sets of parameters are contradictory. It should be noted that the goal of the SUFI-2 algorithm application is not to find the so-called “best simulation” in such a stochastic procedure but instead to find the 95PPU that brackets some or most of the observed data (Abbaspour, 2015). Hence, the calibrated values (Table 3) do not represent the "best parameter" but rather the fitted value within a certain range.

Time series plots of measured versus simulated total streamflow on an aggregated monthly time scale for the three calibration sites (Fig. 1) in response to the Livneh-calibrated scenario are presented in Fig. 7. The solid blue triangles represent the measured monthly streamflow that was derived from daily measured streamflows (USGS, 2019). The black solid line represents the simulated flow based on the uncalibrated baseline SWAT model (using default HAWQS input parameters). The baseline SWAT model generally tracked the seasonal variance pattern including the peaks and recessions, although there is an obvious underestimation of the observed streamflows by the simulated streamflows for all three calibration gauge sites (St. Paul, Clinton and Grafton). The red solid line represents the predicted monthly streamflow after calibration was completed. The calibration process resulted in increased predicted streamflows including peak streamflow estimates that are more consistent with observed peak streamflows during the summer periods, although some peak streamflows were still underestimated (especially for Grafton). Winter low streamflow periods were generally still underpredicted, especially versus the observed streamflows for Clinton during November to February. Overall, the magnitude and temporal variation of the simulated streamflows matched the measured streamflows, indicating a realistic representation of the observed hydrographs by the model.

Table 7 presents the statistical results for comparison of the SWAT simulated monthly streamflows versus corresponding observed streamflows for both the calibration and validation gauge sites under Livneh-calibrated scenario. The results indicate satisfactory monthly NS values (>0.5 per the criteria suggested by Moriasi et al., 2007; 2015) for all the 3 calibration gauges and most of the 10 validation gauge sites within the UMRB. However, NS values were <0.5 for the two UMRB subregions that drain to Muscoda and Royalton (Table 7). Weaknesses were also reflected in the other statistics calculated for these two regions. This may be due in part to an under-representation of the impact of

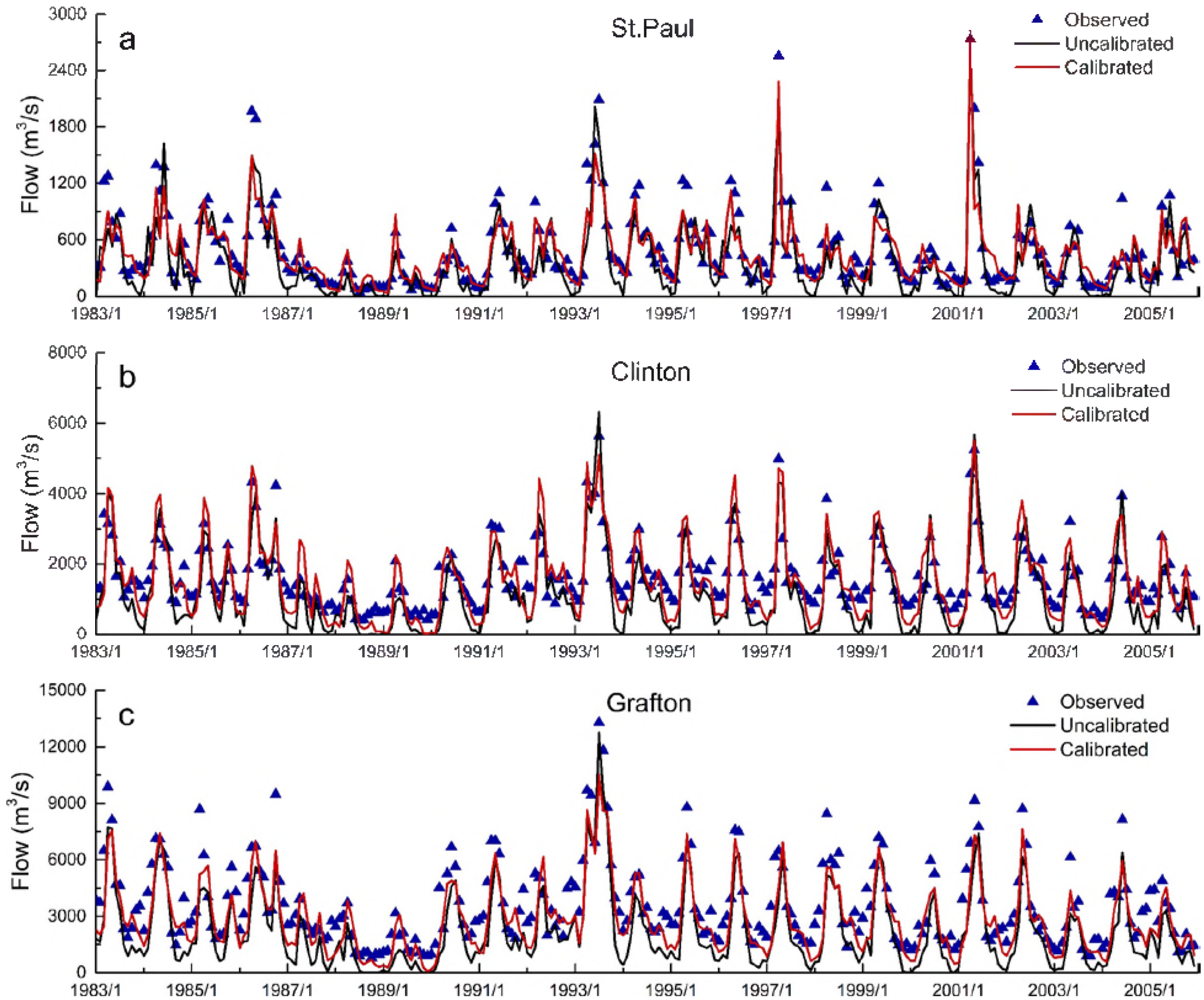

Figure 7. Monthly flows at calibration gauge sites. Observed are measured flow data from USGS stations, uncalibrated are flow outputs of uncalibrated HAWQS/SWAT model when using Livneh dataset, and calibrated are simulated flow after calibrating.

natural lakes and/or wetlands in the two regions, which can attenuate peak streamflows and maintain considerable storage of streamflows in low-flow periods. The percentages of these lake and wetland land uses, which are not captured well by the current HAWQS wetlands parameterization, are the highest for the Muscoda and Royalton drainage areas among the 10 different validation gauge sites.

The R^2 statistics ranged from 0.54 to 0.81, which indicates that the majority of the simulated streamflow trends replicated the counterpart observed streamflows well, considering the R^2 criteria of 0.6 proposed by Moriasi et al. (2015). Almost all of the PBIAS results (Table 8) are acceptable per the criterion of ±25% deviation suggested by Moriasi (2007), except for Valley City (37.30%). Most of

Table 8. Monthly streamflow statistics for calibration and validation gauge sites.

Gauge site	Used for	R^2	NSE	PBIAS	KGE
St.Paul	Calibration	0.78	0.77	5.95	0.76
Clinton	Calibration	0.76	0.58	7.07	0.70
Grafton	Calibration	0.76	0.66	19.74	0.73
Augusta	Validation	0.80	0.73	-8.36	0.62
St.Croix Falls	Validation	0.54	0.54	-0.37	0.61
Durand	Validation	0.66	0.60	13.22	0.75
Joslin	Validation	0.75	0.70	6.64	0.84
Muscoda	Validation	0.61	0.30	17.12	0.64
Royalton	Validation	0.65	0.29	-20.31	0.58
Wapello	Validation	0.82	0.77	3.11	0.66
Keosauqua	Validation	0.68	0.62	-14.74	0.57
Jordan	Validation	0.81	0.79	6.30	0.73
Valley City	Validation	0.81	0.50	37.30	0.52

the PBIAS results also meet the more stringent criteria of ±15% proposed by Moriasi et al. (2015), with the exception being Grafton (19.74%), Muscoda (17.12%), Royalton (-20.31%) and Valley City (37.30%). The positive PBIAS that was calculated for the majority of gauge sites reveals that there was an underestimation bias for the simulated streamflows. The KGE values for three calibration stations were >0.7; i.e., 0.76, 0.70 and 0.73 for St. Paul, Clinton and Grafton, respectively (Table 7). For the validation locations, the lowest KGE value was 0.52 for Valley City while the highest KGE was 0.84 for Joslin. All of the computed KGE statistics met the criteria of 0.6 suggested by Patil and Stieglitz (2015), except the KGE values determined for Royalton (0.58), Keosauqua (0.57) and Valley City (0.52). Overall, the validation statistics verify the calibration process and were even stronger for some gauge sites.

5.5. Comparison of Model performance evaluation

Fig. 8 summarizes all of the evaluated criteria values for six calibrated scenarios (Table 4) and the ensemble mean at the 10 validation gauge sites (Fig. 1 and Table 5). Statistical values that are considered "satisfactory" lie within the rose color background in Fig. 8a-d. Almost all of the R^2 values are acceptable

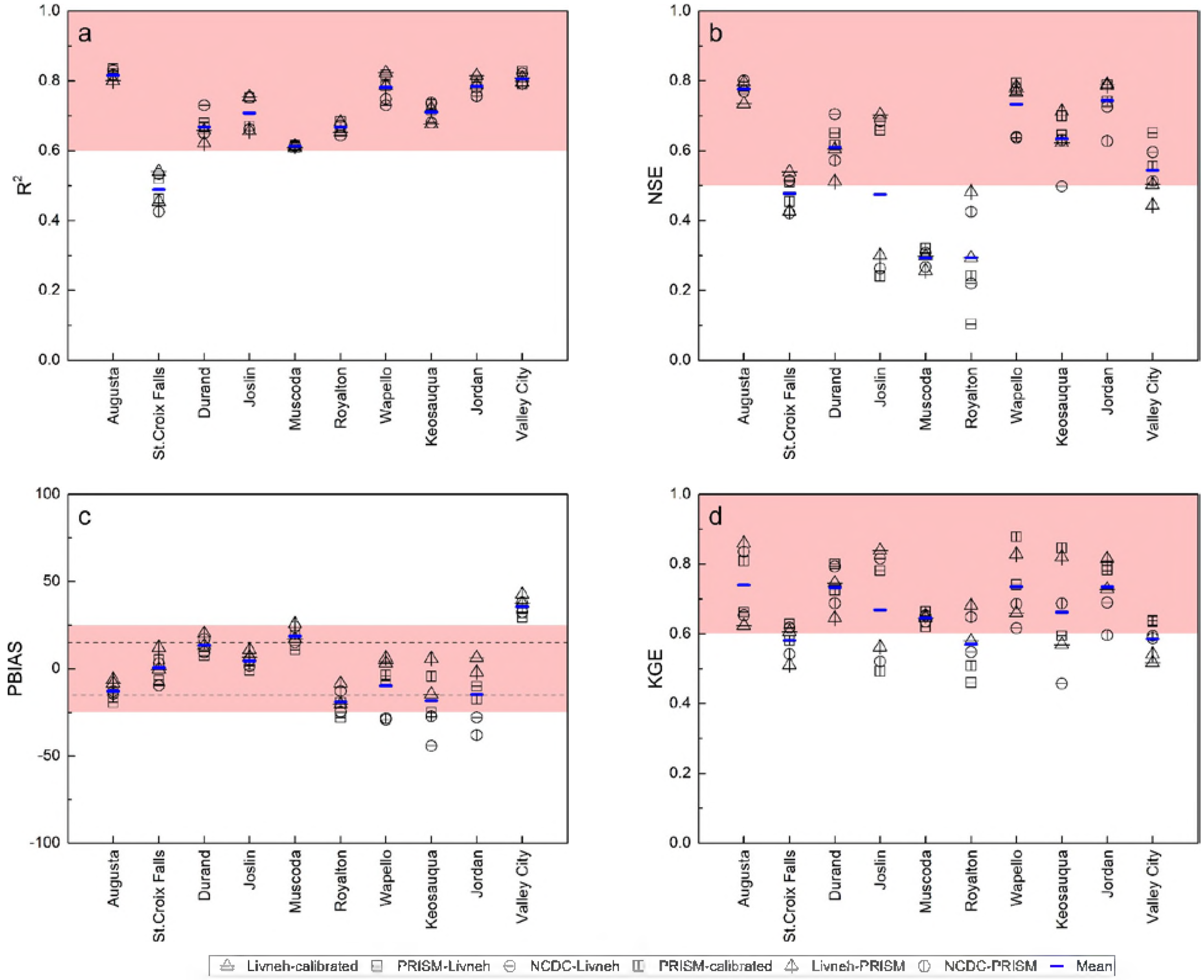

Figure 8. The summary of 4 criteria values for 10 validation gauge sites. The rose background indicates the "Satisfactory" performance range for each criterion. The rose background in 8c depicts the ±25% PBIAS criteria suggested by Moriasi et al. (2007) while the dashed lines in 8c represent the ±15% PBIAS criteria reported by Moriasi et al. (2015).

(>0.6) with the exception of the St. Croix Falls station in Wisconsin (Fig. 8a). For Augusta and Valley City, the mean R2 values are > 0.8, which indicates a strong linear relationship between observed flow and simulation flow. The NSE values determined for the sites of Augusta, Durand, Wapello, Keosauqua, Jordan and Valley City are all satisfactory (>0.5) as shown in Fig. 8b. For Muscoda and Royalton, the NSE values were found to be unacceptable for most of the scenarios, which indicates that additional calibration is likely required for these two independent basins. The majority of PBIAS values are within the "satisfactory" range (≤±25%) as suggested by Moriasi et al. (2007) except for the gauge site located at

Valley City, where streamflows are considerably underestimated resulting in a PBIAS > 40% (Fig. 8c). However, Fig. 8c also shows that some of the computed PBIAS values for seven other gauge sites lie outside of the more stringent criteria of ±15% proposed by Moriasi et al. (2015). Fig. 8d reveals that most of the KGE mean values for the six scenarios are > 0.6 except for St.Croix Falls, Royalton and Valley City. There are also some large differences between scenarios in the KGE criterion for specific gauge sites. For example, the KGE statistics range from 0.75 to 0.85 for the Livneh-calibrated, PRISM-Livneh and NCDC-Livneh scenarios versus 0.5 to 0.6 for the PRISM-calibrated, Livneh-PRISM and NCDC-PRISM scenarios at Joslin.

Most of the statistics for the PRISM-Livneh and NCDC-Livneh scenarios are very close to the corresponding statistics found for the Livneh-calibrated scenario. This implies that the PRISM and NCDC data are adaptable to the model that was calibrated with the Livneh data despite some minor differences in the calculated statistics (Fig. 8). Likewise, application of the NCDC data and Livneh data also result in similar effects for the calibrated model driven by PRISM data. The model performance was very strong at some gauge sites in response to different climate datasets. For example, the statistics determined for the Augusta and Wapello stations all lie in the satisfactory range. However, unacceptable results, according to the criteria suggested by Moriasi et al. (2007; 2015) occurred for some stations for one or more evaluation criteria; e.g., R^2 and NSE values at St.Croix Falls, NSE at Muscoda, and NSE and KGE at Royalton. This may have been caused by the weakness of model adaptability to spatial variability at the subwatershed level. It should also be noted that for some gauge sites, the model estimated streamflow value was even more precise with the alternative climate datasets than with the driving dataset; i.e., NSE and KGE values for the PRISM-Livneh and NCDC-Livneh scenarios are both higher than the Livneh-calibrated scenario at Augusta, Durand and Valley City. These differences in model performance that occurred between the three climate datasets are likely due primarily to differences in the spatial distribution of precipitation and temperature (Fig. 3). However, it is possible that accuracy in precipitation and temperature measurements also effect the ability of SWAT to replicate UMRB streamflows; e.g., the previously described apparent errors in the NCDC temperature for the subset of subwatersheds in the western and central part of the basin.

It can be concluded that the execution of the UMRB SWAT model resulted in different performance levels across all of the calibrated experimental scenarios. However, it is difficult to compare or rank the six different scenarios in a straight forward way. To overcome this issue and to provide additional insights regarding the outcomes of statistical analysis, a Global Performance Indicator (GPI) is introduced to assess the combined effects of the individual statistical indicators (Behar et al., 2015; Despotovic et al., 2015; Jamil and Akhtar, 2017). The values of all the statistical indicator are scaled between 0 and 1.

These scaled indicators are then subtracted from their corresponding median values respectively. Finally, the obtained differences are summed up using appropriate weight factors. The GPI indicator *i* is defined as

$$GPI_i = \sum_{j=1}^{4} \alpha_j (\tilde{y}_j - y_{ij}) \qquad (3)$$

where $\tilde{y}_j$ is the median of scaled values of indicator *j*, y_{ij} is the scaled value of indicator *j* for scenario *i*, α_j equals 1 for the indicator PBIAS, and equals -1 for other 3 indicators. As illustrated in Eq. (3), the GPI in this study is a multiplication of four statistical factors: R^2, NSE, PBIAS and KGE. A higher value of GPI indicates improved accuracy of a scenario between the observed data and simulated data.

The GPI rankings of the six calibrated experimental scenarios at the 10 validation gauge sites are reported in Table 8. For instance, the NCDC-PRISM scenario was ranked first at Augusta while the Livneh-calibrated scenario ranked sixth. The best overall performing climate dataset based on the highest consistent rank was PRISM, due to average rankings of the PRISM-Livneh scenario and PRISM-calibrated scenarios of 2.2 and 2.8 (Table 8), respectively. In contrast, the NCDC-Livneh and NCDC-PRISM scenarios were ranked 3.5 and 4.0, respectively. The Livneh data was ranked in the last positions among the three climate datasets on average, with 4.1 for Livneh-calibrated and 4.4 for Livneh-PRISM.

Table 8. Ranking of scenarios according to GPI at 10 validation gauge sites.

Scenario	Augusta	St.Croix Falls	Durand	Joslin	Muscoda	Royalton	Wapello	Keosauqua	Jordan	Valley City	Average ranking
Livneh-calibrated	6	3	3	3	4	4	4	5	3	5	4.0
PRISM-Livneh	2	2	2	1	3	5	1	4	1	1	2.2
NCDC-Livneh	4	1	1	2	2	6	6	6	5	2	3.5
PRISM-calibrated	3	4	4	4	1	3	2	1	4	3	2.9
Livneh-PRISM	5	6	6	6	6	1	3	3	2	6	4.4
NCDC-PRISM	1	5	5	5	5	2	5	2	6	4	4.0

This does not mean that the Livneh data can not be applied in the UMRB SWAT model but it may result in a weaker performance than PRISM or NCDC in some subwatersheds. The Livneh data is also highly ranked at some stations; i.e., it was ranked in the third place at St.Croix Falls, Joslin and Jordan for the Livneh-calibrated scenario, and in first place at Royalton for the Livneh-PRISM scenario.

5.6. Reflections on results relative to previous UMRB SWAT studies

The NSE and R^2 calibration/validation statistical results computed between the SWAT-simulated and measured streamflows in this study compare favorably with corresponding statistics reported in previous studies that mostly ranged between 0.5 and 0.9 (Table 2). The primary exceptions were validation gauge sites located in the northern part of the UMRB system (Joslin, Muscoda, Royalton and St. Croix Falls) that manifested weaker statistics (Table 5 and Figure 8). This was likely due in part to the more rigorous spatial validation approach used in this study in which the calibrated parameters (Table 3), that were determined for 3 gauge sites (Figure 1 and Table 5), were then simulated for the 10 validation sites (Figure 1 and Table 5) without any further adjustments. In addition, two other reasons may have contributed to the weaker results within the HAWQS-based SWAT simulations of these northern subregions: (1) the lack of accounting for ponds, wetlands and other non-stream water bodies (as previously noted), which may have particularly affected the water balance results at these specific gauge sites, and (2) inadequate representation of forest growth parameters and algorithms, which has been documented as a weakness in previously reported SWAT applications (Yang et al., 2018; Yang and Zhang, 2016) and would be of particular importance in these northern subregions because forest is a dominant land use in the areas that drain to these gauge sites.

The most directly comparable previous study to the application reported here was the analysis described by Qi et al. (2019), who compared the effects of the NCDC, NLDAS2 and partial-NLDAS2 climate data sets on SWAT streamflow predictions for 11 gauge sites within the UMRB. They found that all three climate data sets resulted in satisfactory replication of UMRB measured streamflows, but that the NLDAS2 data set produced the most accurate results relative to the other two data sets which was likely due to the inclusion of measured solar radiation, relative humidity, and wind speed data (versus just measured precipitation, minimum temperature and maximum temperature data). Their findings are similar to what was found in this study; i.e, all three climate data sets produced acceptable results but the PRISM climate data set generated the most accurate SWAT-streamflow predictions compared to the NCDC and Livneh climate data sets. Overall, the Qi et al. (2019) statistical results were generally stronger than the comparative statistics computed in this study. This may have been partly due to the fact that Qi et al. calibrated and validated SWAT streamflows using a split-time approach for each of the 11 gauge sites included in their application. Qi et al. also accounted for subsurface tile drainage in UMRB subregions that are characterized by low slope and poorly drained soils; tile drainage was not incorporated in the HAWQS-based SWAT models developed for this study. Subsurface tile drains are primary sources of discharge water and soluble nutrients (e.g., nitrate) to stream networks in intensely tile areas as documented in several previous studies that focused on UMRB subwatersheds (Jha et al., 2010; Beeson et al., 2014; Panagopoulos et al., 2015; Teshager et al., 2016; Gassman et al., 2017; Jones et al., 2018; Schilling et al., 2019).

In summary, the SWAT models that were developed relatively rapidly within HAWQS for this study were successful in replicating UMRB streamflows for most of the gauge sites that were evaluated within the calibration or spatial validation phases. However, future improvements are needed to better represent specific aspects of the UMRB system including incorporation of non-stream water bodies and subsurface tile drainage. These and other improvements can provide improved estimates of streamflow as well as more accurate depiction of nutrient and other pollutant transport in the region.

6. Conclusions

The SWAT model was developed for the UMRB by using the on-line data and other resources provided by HAWQS. The uncalibrated model was used to evaluate the impacts of three spatial climate datasets (PRISM, NCDC and Livneh) and two PET estimation methods (HG and PM) on UMRB hydrologic processes. A comparison of climate datasets showed that the Livneh data had the highest precipitation and temperature levels during the growing season from May to September. The differences in precipitation and temperature inputs between the three climate data sets results are a primary factor in the SWAT-estimated differences in streamflow, ET and other hydrological outputs. Regarding the impact of the two PET methods, higher annual ET and PET values were calculated with HG method versus the PM method for all three climate data sets. This is because the SWAT-predicted ET and PET amounts are considerably higher with HG method versus the corresponding PM-based estimates during the growing season.

The UMRB SWAT scenario performances were evaluated on a monthly time step according to four statistics: coefficient of determination (R^2), percent bias (PBIAS), Nash-Sutcliffe modeling efficiency (NSE) and Kling-Gupta efficiency (KGE). Parallel processing and spatial validation were used in the calibration and validation of such a large hydrologic system, which improved the execution speed greatly and captured the spatial variation in runoff. The results of the calibration and validation phases showed that the SWAT model based on the Livneh dataset and HG method replicated streamflows well at most of the monitoring stations (three calibration points and ten validation points), indicating that the model could adequately predict long-term water yield in UMRB. After replacing the Livneh dataset with PRISM and NCDC, the model performances for validation points are still satisfactory on the whole despite some differences that occurred per the computed statistics. This substitutability between weather datasets also revealed that the calibrated SWAT model, which was based on the PRISM data, resulted in in mostly satisfactory results. In addition, the Global Performance Indicator (GPI) was used so that all six of the experimental scenarios that were based on a calibrated version of the model could be evaluated with a single parameter and easily ranked. Based on the ranking of GPI, the PRISM data was found to be the strongest climate data set among the three climate data sets.

However, uncertainties in the available climate data and variations in other spatial data need to be further evaluated and improved for large-scale watershed modeling such as the UMRB system simulated here. This is especially true for the NCDC climate data which exhibited unexpected anomalies in the temperature data (Figure 3) that should be resolved in future versions of HAWQS. In addition, incorporation of non-stream water bodies, subsurface tile drainage and other aspects of the UMRB system, that were not simulated in this study, are needed to more accurately simulate streamflows and pollutant transport throughout the stream network.

Based on the results of this study, the HG method would be recommended to be applied in the UMRB SWAT model because it resulted in a higher range of predicted ET/precipitation ratios which is more consistent with the limited estimates reported for the region (Liu et al., 2013). It is also recommended that the PRISM climate data be selected for UMRB SWAT applications built in HAWQS based on the results obtained in this study. The results of this study also confirm that future users of HAWQS should conduct testing of any SWAT models built in the system, regardless of the watershed that is being analyzed.

Acknowledgements

This research was partially funded by U.S. Department of Energy Initiative, Award No. DESC0016438, "A Hierarchical Evaluation Framework for Assessing Climate Simulations Relevant to the Energy-Water-Land Nexus", and by U.S. Department of Energy Initiative, Award No. DESC0016605, "An Integrated Assessment of Regional Climate-Water-Energy-Land-Decision Modeling." Appreciation is also expressed to the China Scholarship Council for providing support to lead author Chen to conduct research at Iowa State University from September 2018 to September 2019 and to the USEPA Office of Water for their ongoing support of different aspects related to HAWQS. Finally, we acknowledge our late colleague Raymond W. Arritt, whose scientific leadership and curiosity greatly influenced the design and implementation of this study before his untimely death occurred in October, 2018.

References

Abbaspour, K.C. 2015. SWAT-CUP: SWAT calibration and uncertainty programs – a user manual. Eawag, Swiss Federal Institute of Aquatic Science and Technology: Dübendorf, Switzerland. Available at: https://swat.tamu.edu/software/swat-cup/.

Alemayehu, T., van Griensven, A., Bauwens, W., 2016. Evaluating CFSR and WATCH data as input to SWAT for the estimation of the potential evapotranspiration in a data-scarce eastern-African catchment. Journal of Hydrologic Engineering. 21(3). https://doi.org/10.1061/(ASCE)HE.1943-5584.00013.

Allen, R.G., W.O. Pruitt, J.L. Wright, T.A. Howell, F. Ventura, R. Snyder, D. Itenfisu, P. Steduto, J. Berengena, J.B. Yrisarry, M. Smith, L.S. Pereira, D. Raes A. Perrier, I. Alves, I. Walter and R. Elliott. 2006. A recommendation on standardized surface resistance for hourly calculation of reference ETo by the FAO56 Penman-Monteith method. Agricultural Water Management. 81: 1-22. Doi: 10.1016/j.agwat.2005.03.007.

Almendinger, J.E., M.S. Murphy and J.S. Ulrich. 2014. Use of SWAT to scale sediment delivery from field to watershed in an agricultural landscape with topographic depressions. Journal of Environmental Quality. 43(1): 9-17. Doi: 10.2134/jeq2011.0340.

Aouissi, J., Benabdallah, S., Lili Chabaâne, Z. and Cudennec, C., 2016. Evaluation of potential evapotranspiration assessment methods for hydrological modelling with SWAT—Application in data-scarce rural Tunisia. Agricultural Water Management, 174: 39-51. https://doi.org/10.1016/j.agwat.2016.03.004

Arnold, J.G., D.N. Moriasi, P.W. Gassman, K.C. Abbaspour, M.J. White, R. Srinivasan, C. Santhi, R.D. Harmel, A. van Griensven, M.W. Van Liew, N. Kannan and M.K. Jha. 2012. SWAT: Model use, calibration, and validation. Transactions of the ASABE. 55(4): 1491-1508. Doi: 10.13031/2013.42256.

Arnold, J.G., R.S. Muttiah, R. Srinivasan and P.M. Allen. 2000. Regional estimation of base flow and groundwater recharge in the Upper Mississippi River Basin. Journal of Hydrology. 227(1-4): 21-40. Doi: 10.1016/S0022-1694(99)00139-0.

Arnold, J.G., R. Srinivasan, R.S. Muttiah and J.R. Williams. 1998. Large area hydrologic modeling and assessment part I: Model development. Journal of the American Water Resources Association. 34(1): 73-89. Doi: 10.1111/j.1752-1688.1998.tb05961.x.

Barnhart, B.L., H.E. Golden, J.R. Kasprzyk, J.J. Pauer, C.E. Jones, K.A. Sawicz, N. Hoghooghi, M. Simon, R.B. McKane, P.M. Mayer, A.N. Piscopo, D.L. Ficklin, J.J. Halama, P.B. Pettus and B. Rashleigh. 2018. Embedding co-production and addressing uncertainty in watershed modeling decision-support tools: Successes and challenges. Environmental Modelling & Software. 109: 368-379. Doi: 10.1016/j.envsoft.2018.08.025.

Beeson, P.C., A.M. Sadeghi, M.W. Lang, M.D. Tomer and C.S.T. Daughtry. 2014. Sediment delivery estimates in water quality models altered by resolution and source of topographic data. Journal of Environmental Quality. 43(1): 26-36. Doi: 10.2134/jeq2012.0148.

Behar, O., Khellaf, A. and Mohammedi, K., 2015. Comparison of solar radiation models and their validation under Algerian climate – The case of direct irradiance. Energy Conversion and Management, 98: 236-251. Doi: 10.1016/j.enconman.2015.03.067.

Bieger, K., J.G. Arnold, H. Rathjens, M.J. White, D.D. Bosch, P.M. Allen, M. Volk and R. Srinivasan. 2017. Introduction to SWAT+, a completely restructured version of the Soil and Water Assessment Tool. Journal of the American Water Resources Association. 53(1): 115–130. Doi: 10.1111/1752-1688.12482.

Bouska, K.L., J. Houser, N.R. De Jager and J. Hendrickson. 2018. Developing a shared understanding of the Upper Mississippi River: The foundation of an ecological resilience assessment. Ecology and Society 23(2): 6. Doi: 10.5751/ES-10014-230206.

Bressiani, D. de A., P.W. Gassman, J.G. Fernandes, L.H.P. Garbossa, R. Srinivasan, N.B. Bonumá and E.M. Mendiondo. 2015. A review of Soil and Water Assessment Tool (SWAT) applications in Brazil: Challenges and prospects. International Journal of Agricultural and Biological Engineering. 8(3): 9-35. Doi: 10.3965/j.ijabe.20150803.1765.

CARD. 2019. SWAT literature database for peer-reviewed journal articles. Center for Agricultural and Rural Development, Iowa State University, Ames, Iowa. Available at: https://www.card.iastate.edu/swat_articles/.

Cherkauer, K.A. and D.P. Lettenmaier. 1999. Hydrologic effects of frozen soils in the Upper Mississippi River basin. Journal of Geophysical Research: Atmospheres. 104(D16): 19599-19610. Doi: 10.1029/1999JD900337.

Chaplot, V., A. Saleh, D.B. Jaynes and J. Arnold. 2004. Predicting water, sediment and NO3-N loads under scenarios of land-use and management practices in a flat watershed. Water, Air, and Soil Pollution. 154 (1-4): 271-293. Doi: 10.1023/B:WATE.0000022973.60928.30.

Chen, L., Wang, J., Wei, W., Fu, B. and Wu, D., 2010. Effects of landscape restoration on soil water storage and water use in the Loess Plateau Region, China. Forest Ecology and Management, 259(7): 1291-1298. Doi: 10.1016/j.foreco.2009.10.025.

Christianson, R., L. Christianson, C. Wong, M. Helmers, G. McIsaac, D. Mulla and M. McDonald. 2018. Beyond the nutrient strategies: Common ground to accelerate agricultural water quality improvement in the upper Midwest. Journal of Environmental Management. 206:1072-1080. Doi: 10.1016/j.jenvman.2017.11.051.

Criss, R.E. and M. Luo. 2017. Increasing risk and uncertainty of flooding in the Mississippi River Basin. Hydrological Processes. 31: 1283–1292. Doi: 10.1002/hyp.11097.

Criss, R.E. and E.L. Shock. 2001. Flood enhancement through flood control. Geology. 29(10): 875–878. Doi: 10.1130/0091-7613(2001)029<0875:FETFC>2.0.CO;2.

Daly, C. et al., 2008. Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. INTERNATIONAL JOURNAL OF CLIMATOLOGY, 28(15): 2031-2064. Doi: 10.1002/joc.1688.

Deb, D., P. Tuppad, P. Daggupati, R. Srinivasan and D. Varma. 2015. Spatio-temporal impacts of biofuel production and climate variability on water quantity and quality in Upper Mississippi River Basin. Water. 7(4): 3283-3305. Doi: 10.3390/w7073283.

Demissie, Y., E. Yan and M. Wu. 2012a. Assessing regional hydrology and water quality implications of large-scale biofuel feedstock production in the upper Mississippi River basin. Environmental Science & Technology. 46: 9174-9182. Doi: 10.1021/es300769k.

Demissie, Y., E. Yan, M. Wu and Z. Zhang. 2012b. Watershed modeling of potential impacts of biofuel feedstock production in the Upper Mississippi River Basin. Report ANL/EVS/AGEM/TR-12-07. Argonne National Laboratory: Argonne, IL. Available at: http://www.ipd.anl.gov/anlpubs/2012/08/73898.pdf.

Despotovic, M., Nedic, V., Despotovic, D. and Cvetanovic, S., 2015. Review and statistical analysis of different global solar radiation sunshine models. Renewable and Sustainable Energy Reviews, 52: 1869-1880. Doi: 10.1016/j.rser.2015.08.035.

Eisner, S., M. Flörke, A. Chamorro, P. Daggupati, C. Donnelly, J. Huang, Y. Hundecha, H. Koch, A. Kalugin, I. Krylenko, V. Mishra, M. Piniewski, L. Samaniego, O. Seidou, M. Wallner and V. Krysanova. 2017. An ensemble analysis of climate change impacts on streamflow seasonality across 11 large river basins. Climatic Change. 141(3): 401-417. Doi: 10.1007/s10584-016-1844-5.

Ellet, C. 1852. Report on the overflows of the delta of the Mississippi. War Department: Washington, D.C. Available at: https://biotech.law.lsu.edu/la/geology/REPORT-ON-THE-OVERFLOWS-OF-THE-DELTA-OF-THE-MISSISSIPPI.pdf.

ESRL. 2019. Livneh daily CONUS near-surface gridded meteorological and derived hydrometeorological data. Earth System Research Laboratory, National Oceanic and Atmospheric Administration, U.S. Department of Commerce: Boulder, CO. Available at: https://www.esrl.noaa.gov/psd/data/gridded/data.livneh.html.

Fant, C., R. Srinivasan, B. Boehlert, L. Rennels, S.C. Chapra, K.M. Strzepek, J. Corona, A. Allen and J. Martinich. 2017. Climate change impacts on US water quality using two models: HAWQS and US Basins. Water. 9(2): 118. Doi: 10.3390/w9020118.

Feng, Q., I. Chaubey, R. Cibin, B. Engel, K.P. Sudheer, J. Volenec and N. Omani. 2018. Perennial biomass production from marginal land in the Upper Mississippi River Basin. Land Degradation & Development. 29(6): 1748-1755. Doi: 10.1002/ldr.2971.

Feng, Q., I. Chaubey, B. Engel, R. Cibin, K.P. Sudheer and J. Volenec. 2017. Marginal land suitability for switchgrass, Miscanthus and hybrid poplar in the Upper Mississippi River Basin (UMRB). Environmental Modelling & Software. 93: 356-365. Doi: 10.1016/j.envsoft.2017.03.027.

Gao, J., Sheshukov, A.Y., Yen, H. and White, M.J., 2017. Impacts of alternative climate information on hydrologic processes with SWAT: A comparison of NCDC, PRISM and NEXRAD datasets. CATENA, 156: 353-364. Doi: 10.1016/j.catena.2017.04.010.

García, A.M., R.B. Alexander, J.G. Arnold, L. Norfleet, M.J. White, D.M. Robertson, and G. Schwarz. 2016. Regional effects of agricultural conservation practices on nutrient transport in the Upper Mississippi River Basin. Environmental Science Technology. 50(13): 6991−7000 Doi: 10.1021/acs.est.5b03543.

Gassman, P.W., M.R. Reyes, C.H. Green and J.G. Arnold. 2007. The Soil and Water Assessment Tool: Historical development, applications, and future research directions. Transactions of the ASABE. 50(4): 1211-1250. Doi: 10.13031/2013.23634.

Gassman, P.W., A.M. Sadeghi and R. Srinivasan. 2014. Applications of the SWAT Model Special Section: Overview and Insights. Journal of Environmental Quality. 43(1): 1-8. Doi: 10.2134/jeq2013.11.0466.

Gassman, P.W., A.M. Valcu-Lisman, C.L. Kling, S.K. Mickelson, Y. Panagopoulos, R. Cibin, I. Chaubey, C.F. Wolter and K.E. Schilling. 2017. Assessment of bioenergy cropping scenarios for the Boone River watershed in north central Iowa, United States. Journal of the American Water Resources Association. 53(6): 1336-1359. Doi: 10.1111/1752-1688.12593.

Gassman, P.W., J.R. Williams, X. Wang, A. Saleh, E. Osei, L.M. Hauck, R.C. Izaurralde and J.D. Flowers. 2010. The Agricultural Policy/Enviromental EXtender (APEX) model: An emerging tool for landscape and watershed environmental analyses. Transactions of the ASABE. 53(3): 711-740. Doi: 10.13031/2013.30078.

Getahun, E. and L. Keefer. 2016. Integrated modeling system for evaluating water quality benefits of agricultural watershed management practices: Case study in the Midwest. Sustainability of Water Quality and Ecology. 8: 14-29. Doi: 10.1016/j.swaqe.2016.06.002.

Gosling, S.N., J. Zaherpour, N.J. Mount, F.F. Hattermann, R. Dankers, B. Arheimer, L. Breuer, J. Ding, I. Haddeland, R. Kumar, D. Kundu, J. Liu, A. van Griensven, T.L.E. Veldkamp, T. Vetter, X. Wang and X. Zhang. 2017. A comparison of changes in river runoff from multiple global and catchment-scale hydrological models under global warming scenarios of 1 °C, 2 °C and 3 °C. Climatic Change. 141(3): 577-595. Doi: 10.1007/s10584-016-1773-3

Green, M.B. and D. Wang. 2008. Watershed flow paths and stream water nitrogen-to-phosphorus ratios under simulated precipitation regimes. Water Resources Research. 44: 1-13. Doi: 10.1029/2007WR006139.

Gu, R.R., M.K. Sahu and M.K. Jha. 2015. Simulating the impacts of bio-fuel crop production on nonpoint source pollution in the Upper Mississippi River Basin. Ecological Engineering. 74: 223-229. Doi: 10.1016/j.ecoleng.2014.10.010.

Gupta, H.V., Kling, H., Yilmaz, K.K. and Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. Journal of Hydrology, 377(1-2): 80-91. Doi: 10.1016/j.jhydrol.2009.08.003.

Hanratty, M.P. and H.G. Stefan. 1998. Simulating climate change effects in a Minnesota agricultural watershed. Journal of Environmental Quality. 27: 1524-1532. Doi: 10.2134/jeq1998.00472425002700060032x.

Hargreaves G.H. and Z.A. Samani. 1985. Reference crop evapotranspiration from temperature. Applied Engineering in Agriculture. 1(2): 96-99. Doi: 10.13031/2013.26773.

Hattermann, F.F., V. Krysanova, S N. Gosling, R. Dankers, P. Daggupati, C. Donnelly, M. Flörke, S. Huang, Y. Motovilov, S. Buda, T. Yang, C. Müller, G. Leng, Q. Tang, F.T. Portmann, S. Hagemann, D. Gerten, Y. Wada, Y. Masaki, T. Alemayehu, Y. Satoh and L. Samaniego. 2017. Cross-scale intercomparison of climate change impacts simulated by regional and global hydrological models in eleven large river basins. Climatic Change. 141(3): 561-576. Doi: 10.1007/s10584-016-1829-4.

HAWQS. 2017. HAWQS v1.0: Inputs. Texas A&M AgriLife Research, Texas A&M University: College Station, Texas: Available at: https://hawqs.tamu.edu/content/docs/HAWQS-Input-Database-Citation.pdf.

Huang, S., R. Kumar, M. Flörke, T. Yang, Y. Hundecha, P. Kraft, C. Gao, A. Gelfan, S. Liersch, A. Lobanova, M. Strauch, F. van Ogtrop, J. Reinhardt, U. Haberlandt and V. Krysanova. 2017. Evaluation of an ensemble of regional hydrological models in 12 large-scale river basins worldwide. Climatic Change. 141(3): 381-397. Doi: 10.1007/s10584-016-1841-8.

Jamil, B. and Akhtar, N., 2017. Estimation of diffuse solar radiation in humid-subtropical climatic region of India: Comparison of diffuse fraction and diffusion coefficient models. Energy, 131: 149-164. Doi: 10.1016/j.energy.2017.05.018.

Jha, M., J.G. Arnold, P.W. Gassman, F. Giorgi and R.R. Gu. 2006. Climate change sensitivity assessment on Upper Mississippi River Basin streamflows using SWAT. Journal of the American Water Resources Association. 42(4): 997-1016. Doi: 10.1111/j.1752-1688.2006.tb04510.x.

Jha, M.K., P.W. Gassman and Y. Panagopoulos. 2015. Regional changes in nitrate loadings in the Upper Mississippi River Basin under predicted mid-century climate. Regional Environmental Change. 15(3): 449-460. Doi: 10.1007/s10113-013-0539-y.

Jha, M., Z. Pan, E.S. Takle and R. Gu. 2004. Impacts of climate change on streamflow in the Upper Mississippi River Basin: A regional climate model perspective. Journal of Geophysical Research. 109: D09105. Doi: 10.1029/2003JD003686.

Jha, M., K.E. Schilling, P.W. Gassman and C.F. Wolter. 2010. Targeting land-use change for nitrate-nitrogen load reductions in an agricultural watershed. Journal of Soil and Water Conservation. 65(6): 342-352. Doi: 10.2489/jswc.65.6.342.

Jones, C.S., J.K. Nielsen, K.E. Schilling and L.J. Weber. 2018. Iowa stream nitrate and the Gulf of Mexico. PLoS ONE 13(4): e0195930. Doi: 10.1371/journal.pone.0195930.

Kannan, N., C. Santhi and J.G. Arnold. 2008. Development of an automated procedure for estimation of the spatial variation of runoff in large river basins. Journal of Hydrology. 359(1-2): 1-15. Doi: 10.1016/j.jhydrol.2008.06.001.

Kannan, K., C. Santhi, M.J. White, S. Mehan, J.G. Arnold and P.W. Gassman. 2019. Some challenges in hydrologic model calibration for large-scale studies: A case study of SWAT model application to Mississippi-Atchafalaya River Basin. Hydrology. 6(1): 17. Doi: 10.3390/hydrology6010017.

Kirsch, K., A. Kirsch, and J.G. Arnold. 2002. Predicting sediment and phosphorus loads in the Rock River Basin using SWAT. Transactions of the ASAE. 45(6): 1757-1769. Doi: 10.13031/2013.11427.

Kling, C.L., Y. Panagopoulos, S.S. Rabotyagov, A.M. Valcu, P.W. Gassman, T. Campbell, M.J. White, J.G. Arnold, R. Srinivasan, M.K. Jha, J.J. Richardson, L.M. Moskal, R.E. Turner and N.N. Rabalais. 2014. LUMINATE: Linking agricultural land use, local water quality and Gulf of Mexico hypoxia. European Review of Agricultural Economics. 41(3): 431–459. Doi: 10.1093/erae/jbu009.

Krause, P., D.P. Boyle and F. Bäse. 2005. Comparison of different efficiency criteria for hydrological model assessment. Advances in Geosciences. 5: 89–97. Doi: 10.5194/adgeo-5-89-2005.

Krysanova, V. and F.F. Hattermann. 2017. Intercomparison of climate change impacts in 12 large river basins: Overview of methods and summary of results. Climatic Change. 141(3): 363-379. Doi: 10.1007/s10584-017-1919-y.

Krysanova V. and M. White. 2015. Advances in water resources assessment with SWAT—an overview. Hydrological Sciences Journal. 60(5): 771-783. Doi: 10.1080/02626667.2015.1029482.

Li, P., R.L Muenich, I. Chaubey and X. Wei. 2019. Evaluating agricultural BMP effectiveness in improving freshwater provisioning under changing climate. Water Resources Management. Doi: 10.1007/s11269-018-2098-y.

Li, P., N. Omani, I. Chaubey and X. Wei. 2017. Evaluation of drought implications on ecosystem services: Freshwater provisioning and food provisioning in the Upper Mississippi River Basin. International Journal of Environmental Research and Public Health. 14(5): 496. Doi: 10.3390/ijerph14050496.

Liu, M., H. Tian, Q. Yang, J. Yang, X. Song, S.E. Lohrenz and W.-J. Cai. 2013. Long-term trends in evapotranspiration and runoff over the drainage basins of the Gulf of Mexico during 1901–2008. Water Resources Research. 49:1988–2012. Doi: 10.1002/wrcr.20180.

Livneh B., E.A. Rosenberg, C. Lin, B. Nijssen, V. Mishra, K.M. Andreadis, E.P. Maurer and D.P. Lettenmaier. 2013. A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States: Update and extensions, Journal of Climate. 26(23): 9384–9392. Doi: 10.1175/JCLI-D-12-00508.1.

McLellan, E., D. Robertson, K. Schilling, M. Tomer, J. Kostel, D. Smith and K. King. 2015. Reducing nitrogen export from the Corn Belt to the Gulf of Mexico: Agricultural strategies for remediating hypoxia. Journal of the American Water Resources Association. 51(1): 263-289. Doi: 10.1111/jawr.12246.

Menne, M.J., Durre, I., Vose, R.S., Gleason, B.E. and Houston, T.G., 2012. An Overview of the Global Historical Climatology Network-Daily Database. Journal of Atmospheric and Oceanic Technology, 29(7): 897-910. Doi: 10.1175/JTECH-D-11-00103.1.

Monteith, J.L. 1965. Evaporation and environment. In: The state and movement of water in living organisms, XIXth Symposium. Society for Experimental Biology, Swansea, United Kingdom. Cambridge University Press: London, United Kingdom.

Moriasi, D.N., J.G. Arnold, M.W. Van Liew, R.L. Bingner, R.D. Harmel and T. Veith. 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Transactions of the ASABE. 50(3): 885-900. Doi: 10.13031/2013.23153.

Moriasi, D.N., M.W. Gitau, N. Pai and P. Daggupati. 2015. Hydrologic and water quality models: performance measures and evaluation criteria. Transactions of the ASABE. 58(6): 1763-1785. Doi: 10.13031/trans.58.10715.

NOAA. 2019. Climate data online. National Centers for Environmental Information, National Oceanic and Atmospheric Administration: Ashville, NC. Available at: https://www.ncdc.noaa.gov/cdo-web/.

Panagopoulos, Y., P.W. Gassman, R.W. Arritt, D.E. Herzmann, T.D. Campbell, M.K. Jha, C.L. Kling, R. Srinivasan, M. White and J. G. Arnold. 2014. Surface water quality and cropping systems sustainability under a changing climate in the Upper Mississippi River Basin. Journal of Soil and Water Conservation. 69(6): 483-494. Doi: 10.2489/jswc.69.6.483.

Panagopoulos, Y., P.W. Gassman, M.K. Jha, C.L. Kling, T. Campbell, R. Srinivasan, M. White and J.G. Arnold. 2015. A refined regional modeling approach for the Corn Belt - experiences and recommendations for large-scale integrated modeling. Journal of Hydrology. 524: 348–366. Doi: 10.1016/j.jhydrol.2015.02.039.

Panagopoulos, Y., P.W. Gassman, C.L. Kling, R. Cibin and I. Chaubey. 2017. Water quality assessment of large-scale bioenergy cropping scenarios for the Upper Mississippi and Ohio-Tennessee River basins. Journal of the American Water Resources Association. 53(6): 1355-1367. Doi: 10.1111/1752-1688.12594.

Patil, S.D. and Stieglitz, M., 2015. Comparing spatial and temporal transferability of hydrological model parameters. Journal of Hydrology, 525: 409-417. Doi: 10.1016/j.jhydrol.2015.04.003.

PCG. 2019. PRISM Climate Data. PRISM Climate Group, Northwest Alliance for Computational Science and Engineering, Oregon State University: Corvallis, OR. Available at: http://www.prism.oregonstate.edu/.

Pierce, D.W. 2016. LOCA statistical downscaling (localized constructed analogs): Training (observed) data sets. Scripps Institution of Oceanography: La Jolla, CA. Available at: http://loca.ucsd.edu/training-observed-data-sets/.

Qi, J., Q. Wang and X. Zhang. 2019a. On the use of NLDAS2 weather data for hydrologic modeling in the Upper Mississippi River Basin. Water. 11(5): 960. Doi: 10.3390/w11050960.

Qi, J., X. Zhang and Q. Wang. 2019b. Improving hydrological simulation in the Upper Mississippi River Basin through enhanced freeze-thaw cycle representation. Journal of Hydrology. 571: 605-618. Doi: 10.1016/j.jhydrol.2019.02.020.

Rabotyagov, S., T. Campbell, M. Jha, P.W. Gassman, J. Arnold, L. Kurkalova, S. Secchi, H. Feng and C.L. Kling. 2010. Least-cost control of agricultural nutrient contributions to the Gulf of Mexico hypoxic zone. Ecological Applications. 20(6): 1542-1555. Doi: 10.1890/08-0680.1.

Rabotyagov, S.S., C.L. Kling, P.W. Gassman, N.N. Rabalais, and R.E. Turner. 2014. The economics of dead zones: Linking externalities from the land to their consequences in the sea. Review of Environmental Economics and Policy. 8(1): 58-79. Doi: 10.1093/reep/ret024.

Rajib, A. and V. Merwade. 2017. Hydrologic response to future land use change in the Upper Mississippi River Basin by the end of 21st century. Hydrological Processes. 31: 3645-3661. Doi: 10.1002/hyp.11282.

Roy, S., Banerjee, R. and Bose, P.K., 2014. Performance and exhaust emissions prediction of a CRDI assisted single cylinder diesel engine coupled with EGR using artificial neural network. Applied Energy, 119: 330-340. Doi: 10.1016/j.apenergy.2014.01.044.

Roth, V. and Lemann. T., 2016. Comparing CFSR and conventional weather data for discharge and soil oss modelling with SWAT in small catchments in the Ethiopian Highlands. Hydrology and Earth System Sciences. 20: 921-934. https://doi.org/10.5194/hess-20-921-2016.

Royer, T.V., M.B. David and L.E. Gentry. 2006. Timing of riverine export of nitrate and phosphorus from agricultural watersheds in Illinois: Implications for reducing nutrient loading to the Mississippi River. Environmental Science & Technology. 40(13): 4126-4131. Doi: 10.1021/es052573n.

Santhi, C., Kannan, N., Arnold, J.G. and Di Luzio, M., 2008. Spatial calibration and temporal validation of flow for regional scale hydrologic modeling. JOURNAL OF THE AMERICAN WATER RESOURCES ASSOCIATION, 44(4): 829-846. Doi: 10.1111/j.1752-1688.2008.00207.x.

Santhi, C., M. White, J.G. Arnold, L. Norfleet, J. Atwood, R. Kellogg, N. Kannan, X. Wang, M. Di Luzio, J.R. Williams and T. Gerik. 2014. Estimating the effects of agricultural conservation practices on phosphorus loads in the Mississippi-Atchafalaya River Basin. Transactions of the ASABE. 57(5): 1339-1357. Doi: 10.13031/trans.57.10458.

Schilling, K.E., P.W. Gassman, A. Arenas-Amado, C.S. Jones and J. Arnold. 2019. Quantifying the contribution of tile drainage to basin-scale water yield using analytical and numerical models. Science of the Total Environment. 657: 297-309. Doi: 10.1016/j.scitotenv.2018.11.340.

Secchi, S., P.W. Gassman, M. Jha, L. Kurkalova and C.L. Kling. 2011. Potential water quality changes due to corn expansion in the Upper Mississippi River Basin. Ecological Applications. 21(4): 1068-1084. Doi: 10.1890/09-0619.1.

Sprague, L.A., R.M. Hirsch, and B.T. Aulenbach, 2011. Nitrate in the Mississippi River and Its Tributaries. 1980 to 2008: Are We Making Progress? Environmental Science & Technology. 45 (17): 7209-7216. Doi: 10.1021/es201221s.

Srinivasan, R. X. Zhang and J. Arnold. 2010. SWAT ungauged: Hydrological budget and crop yield predictions in the Upper Mississippi River Basin. Transactions of the ASABE. 53(5): 1533-1546. Doi: 10.13031/2013.34903.

Srinivasan, R. 2019. HAWQS user guide: Version 1.1. Spatial Sciences Laboratory, Texas A&M AgriLife Research: College Station, TX. Available at: https://epahawqs.tamu.edu/.

SWAT. 2019. SWAT-CUP: SWAT-CUP is a calibration/uncertainty or sensitivity program interface for SWAT. Texas A&M AgriLife Research, Texas A&M University: College Station, Texas: Available at: https://swat.tamu.edu/software/swat-cup/.

Takle, E.S., M. Jha and C.J. Anderson. 2005. Hydrological cycle in the Upper Mississippi River basin: 20th century simulations by multiple GCMs. Geophysical Research Letters. 32(18): L18407. Doi: 10.1029/2005GL023630.

Takle, E.S., M. Jha, E. Lu, R.W. Arritt, W.J. Gutowski and the NARCCAP Team. 2010. Streamflow in the upper Mississippi river basin as simulated by SWAT driven by 20th Century contemporary results of global climate models and NARCCAP regional climate models. Meteorologische Zeitschrift. 19(4): 341-346. Doi: 10.1127/0941-2948/2010/0464.

Tan, M.L., Gassman, P.W. and Cracknell, A.P., 2017. Assessment of three long-term gridded climate products for hydro-climatic simulations in tropical river basins. Water. 9(3): 229. https://doi.org/10.3390/w9030229.

Tan, M.L., P.W. Gassman, R. Srinivasan, J.G. Arnold and X. Yang. 2019. A review of SWAT studies in Southeast Asia: Applications, challenges and future directions. Water. 11(5): 914. Doi: 10.3390/w11050914.

Teshager, A.D., P.W. Gassman, J.T. Schoof and S. Secchi. 2016. Assessment of impacts of agricultural and climate change scenarios on watershed water quantity and quality, and crop production. Hydrology and Earth System Sciences. 20(8): 3325-3342. Doi: 10.5194/hess-20-3325-2016.

Thornthwaite, C.W. 1948. An approach toward a rational classification of climate. Geographical Review 38: 55-94. Doi: 10.2307/210739.

Tie, Q., Hu, H., Tian, F. and Holbrook, N.M., 2018. Comparing different methods for determining forest evapotranspiration and its components at multiple temporal scales. Science of The Total Environment, 633: 12-29. https://doi.org/10.1016/j.scitotenv.2018.03.082.

Tuppad, P., K.R. Douglas-Mankin, T. Lee, R. Srinivasan and J.G. Arnold. 2011. Soil and Water Assessment Tool (SWAT) hydrologic/water quality model: Extended capability and wider adoption. Transactions of the ASABE. 54(5): 1677-1684. Doi: 10.13031/2013.34915.

Turner, R.E. and N.N. Rabalais. 2019. The Gulf of Mexico. In: World Seas: An Environmental Evaluation (Ed.: C. Sheppard). Academic Press: Amsterdam, The Netherlands. Doi: 10.1016/B978-0-12-805068-2.00022-X.

UMRBA. 2019. River and Basin Facts. Upper Mississippi River Basin Association: Bloomington, MN. Available at: http://www.umrba.org/facts.htm.

USACE. 2016. Report to congress: Upper Mississippi River restoration program. U.S. Army Corps of Engineers: Rock Island District, Rock Island, IL. Available at: https://www.umesc.usgs.gov/ltrmp/documents/2016_umrr_rtc_%20final.pdf.

USDA-NRCS. 2019. PRISM. National Water and Climate Center, Natural Resources Conservation Service, U.S. Department of Agriculture: Portland, Oregon. Available at: https://www.wcc.nrcs.usda.gov/climate/prism.html.

USDOE. 2019. A hierarchical evaluation framework for assessing climate simulations relevant to the energy-water-land nexus. U.S. Department of Energy, Office of Science, Climate and Environmental Sciences Division: Washington, D.C. Available at: https://climatemodeling.science.energy.gov/projects/hierarchical-evaluation-framework-assessing-climate-simulations-relevant-energy-water-land.

USEPA. 2019. Hydrologic and Water Quality System: A national watershed and water quality assessment tool. U.S. Environmental Protection Agency: Washington, D.C. Available at: https://epahawqs.tamu.edu/.

USGS, 2013. Federal Standards and Procedures for the National Watershed Boundary Dataset (WBD). Techniques and Methods 11-A3, Chapter 3 of Section A, Federal Standards Book 11, Collection and Delineation of Spatial Data, Fourth Edition. U.S. Department of the Interior, U.S. Geological Survey, Reston, VA and U.S. Department of Agriculture, Natural Resources Conservation Service, Washington, D.C., 63 pp. Available at: http://pubs.usgs.gov/tm/11/a3/.

USGS. 2019. USGS Surface-Water Daily Data for the Nation. U.S. Department of the Interior, U.S. Geological Survey, Reston, VA. Available at: https://waterdata.usgs.gov/nwis/dv/?referred_module=sw.

Vaché, K.B., J.M. Eilers and M.V. Santelman. 2002. Water quality modeling of alternative agricultural scenarios in the U.S. Corn Belt. Journal of the American Water Resources Association. 38(3): 773-787. Doi: 10.1111/j.1752-1688.2002.tb00996.x.

Valle Júnior, L. C. G., Ventura, T. M., Gomes, R. S. R., de S. Nogueira, J., de A. Lobo, F., Vourlitis, G. L. and Rodrigues, T. R., 2020. Comparative assessment of modelled and empirical reference evapotranspiration methods for a brazilian savanna. Agricultural Water Management, 232: 106040. https://doi.org/10.1016/j.agwat.2020.106040.

Van Meter, K.J., P. Van Cappellen and N.B. Basu. 2018. Legacy nitrogen may prevent achievement of water quality goals in the Gulf of Mexico. Science. 360(6387): 427–430. Doi: 10.1126/science.aar4462.

Vetter, T., J.Reinhardt, M. Flörke, A. van Griensven, F. Hattermann, S. Huang, H. Koch, I. G. Pechlivanidis, S. Plötner, O. Seidou, B. Su, R.W. Vervoort and V. Krysanova. 2017. Evaluation of sources of uncertainty in projected hydrological changes under climate change in 12 large-scale river basins. Climatic Change. 141(3): 419-433. Doi: 10.1007/s10584-016-1794-y.

Vu, T., Li, L. and Jun, K., 2018. Evaluation of Multi-Satellite Precipitation Products for Streamflow Simulations: A Case Study for the Han River Basin in the Korean Peninsula, East Asia. Water, 10(5): 642. https://doi.org/10.3390/w10050642.

Wang, X., N. Kannan, C. Santhi, S.R. Potter, J.R. Williams and J.G. Arnold. 2011. Integrating APEX output for cultivated cropland with SWAT simulation for regional modeling. Transactions of the ASABE. 54(4): 1281-1298. Doi: 10.13031/2013.39031.

Weitzell, R.E, M.L. Khoury, P. Gagnon, B. Schreurs, D. Grossman and J. Higgins. 2003. Conservation Priorities for Freshwater Biodiversity in the Upper Mississippi River Basin. Nature Serve and The Nature Conservancy: Arlington, VA. Available at: https://www.conservationgateway.org/Files/Pages/conservation-priorities-f.aspx.

White, M.J., C. Santhi, N. Kannan, J.G. Arnold, D. Harmel, L. Norfleet, P. Allen, M. DiLuzio, X. Wang, J. Atwood, E. Haney and M. Vaughn Johnson. 2014. Nutrient delivery from the Mississippi River to the Gulf of Mexico and effects of cropland conservation. Journal of Soil and Water Conservation. 69(1): 26-40. Doi: 10.2489/jswc.69.1.26.

Whittaker, G., B.L. Barnhart, R. Srinivasan and J.G. Arnold. 2015. Cost of areal reduction of gulf hypoxia through agricultural practice. Science of the Total Environment. 505: 149-153. Doi: 10.1016/j.scitotenv.2014.09.101.

Williams, J.R., J.G. Arnold, J.R. Kiniry, P.W. Gassman and C.H. Green. 2008. History of model development at Temple, Texas. Hydrological Sciences Journal. 53(5): 948-960. Doi: 10.1623/hysj.53.5.948.

Wu, M., Y. Demissie and E. Yan. 2012a. Simulated impact of future biofuel production on water quality and water cycle dynamics in the Upper Mississippi river basin. Biomass & Bioenergy. 41: 44-56. Doi: 10.1016/j.biombioe.2012.01.030.

Wu, Y., S. Liu and O.I. Abdul-Aziz. 2012b. Hydrological effects of the increased CO2 and climate change in the Upper Mississippi River Basin using a modified SWAT. Climatic Change. 110(3-4): 977-1003. Doi: 10.1007/s10584-011-0087-8.

Wu, J. and K. Tanaka. 2005. Reducing nitrogen runoff from the Upper Mississippi River Basin to control hypoxia in the Gulf of Mexico: Easements or taxes?. Marine Resource Economics. 20(2): 121-144.

Xia, Y., K. Mitchell, M. Ek, J. Sheeld, B. Cosgrove, E. Wood, L. Luo, C. Alonge, H. Wei and J. Meng. 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. Journal of Geophysical Research Atmospheres. 117(D3): D03110. Doi: 10.1029/2011JD016051.

Yang, Q., J.E. Almendinger, X. Zhang, M. Huang, X. Chen, G. Leng, Y. Zhou, K. Zhao, G.R. Asrar, R. Srinivasan and X. Li. 2018. Enhancing SWAT simulation of forest ecosystems for water resource assessment: A case study in the St. Croix River basin. *Ecological Engineering*. 120: 422-431. Doi: 10.1016/j.ecoleng.2018.06.020.

Yang, J., Reichert, P., Abbaspour, K.C., Xia, J. and Yang, H., 2008. Comparing uncertainty analysis techniques for a SWAT application to the Chaohe Basin in China. Journal of Hydrology, 358(1-2): 1-23. Doi: 10.1016/j.jhydrol.2008.05.012.

Yang, Q. and X. Zhang. 2016. Improving SWAT for simulating water and carbon fluxes of forest ecosystems. Science of the Total Environment. 569–570: 1478–1488. Doi: 10.1016/j.scitotenv.2016.06.238.

Yen, H., P. Daggupati, M.J. White, R. Srinivasan, A. Gossel, D. Wells and J.G. Arnold. 2016. Application of large-scale, multi-resolution watershed modeling framework using the Hydrologic and Water Quality System (HAWQS). Water. 8(4): 164. Doi: 10.3390/w8040164.

Yuan, Y., R. Wang, E. Cooter, L. Ran, P. Daggupati, D. Yang, R. Srinivasan and A. Jalowska. 2018. Integrating multimedia models to assess nitrogen losses from the Mississippi River Basin to the Gulf of Mexico. Biogeosciences. 15(23): 7059-7076. Doi: 10.5194/bg-15-7059-2018.

Zhang, Y.K. and Schilling, K.E., 2006. Increasing streamflow and baseflow in Mississippi River since the 1940s: Effect of land use change. Journal of Hydrology, 324(1-4): 412-422. Doi: 10.1016/j.jhydrol.2005.09.033.

Zhu, Z., Chen, Z., Chen, X. and He, P., 2016. Approach for evaluating inundation risks in urban drainage systems. Science of The Total Environment, 553: 1-12. Doi: 10.1016/j.scitotenv.2016.02.025.