

Article

Some Challenges in Hydrologic Model Calibration for Large-Scale Studies: A Case Study of SWAT Model Application to Mississippi-Atchafalaya River Basin

Narayanan Kannan ^{1,*}, Chinnasamy Santhi ², Michael J. White ³, Sushant Mehan ⁴, Jeffrey G. Arnold ⁵ and Philip W. Gassman ⁶

¹ Texas Institute for Applied Environmental Research, Tarleton State University, Stephenville, TX 76401, USA

² Independent Researcher, Frisco, TX 75035, USA; csanthi@yahoo.com

³ United States Department of Agriculture-Agricultural Research Service, Grassland, Soil and Water Research Laboratory, Temple, TX 76502, USA; mike.white@ars.usda.gov

⁴ Formation Environmental LLC, Sacramento, CA 95816, USA; sushantmehan@gmail.com

⁵ United States Department of Agriculture-Agricultural Research Service, Grassland, Soil and Water Research Laboratory, Temple, TX 76502, USA; jeff.arnold@ars.usda.gov

⁶ Center for Agriculture and Rural Development, Iowa State University, Ames, IA 50011, USA; pwgassma@iastate.edu

* Correspondence: kannan@tarleton.edu; Tel.: +1-254-968-9691

Received: 23 December 2018; Accepted: 7 February 2019; Published: 10 February 2019



Abstract: This study is a part of the Conservation Effects Assessment Project (CEAP) aimed to quantify the environmental and economic benefits of conservation practices implemented in the cultivated cropland throughout the United States. The Soil and Water Assessment Tool (SWAT) model under the Hydrologic United Modeling of the United States (HUMUS) framework was used in the study. An automated flow calibration procedure was developed and used to calibrate runoff for each 8-digit watershed (within 20% of calibration target) and the partitioning of runoff into surface and sub-surface flow components (within 10% of calibration target). Streamflow was validated at selected gauging stations along major rivers within the river basin with a target R^2 of >0.6 and Nash and Sutcliffe Efficiency of >0.5 . The study area covered the entire Mississippi and Atchafalaya River Basin (MARB). Based on the results obtained, our analysis pointed out multiple challenges to calibration such as: (1) availability of good quality data, (2) accounting for multiple reservoirs within a sub-watershed, (3) inadequate accounting of elevation and slopes in mountainous regions, (4) poor representation of carrying capacity of channels, (5) inadequate capturing of the irrigation return flows, (6) inadequate representation of vegetative cover, and (7) poor representation of water abstractions (both surface and groundwater). Additional outstanding challenges to large-scale hydrologic model calibration were the coarse spatial scale of soils, land cover, and topography.

Keywords: MARB; river basin; base flow; CEAP; conservation effects assessment project; validation; water yield; runoff; streamflow

1. Introduction

A calibrated model is able to produce most precise and accurate simulated results when compared with observed data [1]. Validation is the process of demonstrating that a given site-specific calibrated model can make sufficiently accurate simulations in a new modeling situation, although “sufficiently accurate” can vary based on project goals [2]. Development of a single, consistent, coherent calibration

and validation strategy that is applicable across all hydrological and water quality models, and biogeochemical, topographic, and climate settings has ranged from challenging to daunting.

The calibration and validation approaches used by many hydrologic and water quality modelers are ad-hoc, piecemeal, incomplete, and often inadequate for the task of achieving or demonstrating that model performance meets the needs of the modeling application [3]. The simplest approach is based on temporal split to compare the single-site data using automated calibration methods (Parameter estimation (PEST), dynamically dimensioned search (DDS), shuffled complex evolution (SCE), etc.). Other common software that are used for calibration and validation of the Soil and Water Assessment Tool (SWAT) ecohydrological model [4–7] include SWAT-CUP [8] or Latin hypercube sampling within the ArcGIS SWAT (ArcSWAT) interface [9]. Model calibration at one site is the most common approach used which may result in undesirable interpretations because the same set of calibrated parameters are used, even in non-homogenous large watersheds, which may result in over-estimates or under-estimates of key hydrologic variables [10,11]. In order to account for complex physical processes within large watersheds, a multi-site or spatial calibration with additional sites is preferred [12].

SWAT is a comprehensive watershed- and river basin-scale model that is designed to support assessments of a wide variety of water quantity and water quality problems. The model has been extensively used worldwide across a broad spectrum of watershed scales and environmental conditions, as documented in more than 3000 individual applications [13] and in reviews of selected subsets of SWAT modeling studies [14–18]. A core component of the majority of these previously published SWAT studies is some level of hydraulic calibration and/or validation. Many of the studies also feature additional pollutant transport calibration and/or validation, depending on the purpose of the study and the availability of upland or in-stream pollutant monitoring data. The calibration and validation methods used in these studies reflect a variety of approaches ranging from what could be termed “ad hoc” to relatively rigorous and systematic methods.

An important subcategory of this existing literature is the application of SWAT for “huge systems”; e.g., major river basin-, national-, or continental-scale systems. Examples include the 492,000 km² Upper Mississippi River basin in the north central U.S. [19,20], the 7.5 million km² Missouri River basin in the west central U.S. [21,22], the 630,000 km² Lower Mekong River basin in southeast Asia [23], 18 major water resource regions (MWRRs) which comprise the conterminous U.S. (48 states) using the Hydrologic Unit Modeling System of the U.S. (HUMUS; [24], the nation of Iran [25], the 4 million km² West African subcontinent [26,27], and the entire continent of Africa [28].

A present example of a huge system SWAT application is the adaptation of HUMUS for the conterminous U.S. within the Conservation Effects Assessment Project (CEAP), which is a U.S. Department of Agriculture (USDA)-Natural Resource Conservation Service (NRCS) initiative that includes both the national study as well as specific watershed studies. The current national HUMUS approach within CEAP features the interface of the Agricultural Policy/Environmental eXtender (APEX) model [29,30] with SWAT. APEX is used to simulate the cropland areas within the 18 U.S. Major Water Resource Regions (MWRRs), which are also known as 2-digit hydrological unit codes (HUCS) or 2-digit watersheds which are part of hierarchical HUC classification system for the U.S. [31]. SWAT is used to simulate the remaining land areas within each MWRR and is also used to simulate the streamflow and pollutant routing within and between subwatersheds [32]. The goal of the national CEAP study is to assess how effective existing conservation practices have been in reducing sediment and nutrient pollution within each MWRR; analyses have been completed for several MWRRs including the Upper Mississippi River Basin [33], the Chesapeake Bay Region [34], the Great Lakes Region [35], and the Ohio-Tennessee River Basin [36]. Calibration and validation of the integrated APEX-SWAT modeling system using in-stream monitoring data has been a key step of the national CEAP study, which has included addressing a range of spatial and temporal calibration problems such as those addressed by Santhi et al. [37] and developing automated calibration procedures that require less computation time for the large-scale MWRR systems [38].

Several previous studies have presented protocols for conducting hydrologic and water quality model applications and/or suggest guidelines for judging model output based on various statistical, graphical, or other criteria. For example, [2] outlined a series of protocol steps for conceptualizing, parameterizing, calibrating, validating, and post-auditing hydrologic model output, which included a case study. The goal of this paper is to outline some of the challenges a model user encounters in calibrating a hydrologic model for a large-scale study. This is carried out based on the results obtained from the CEAP study using SWAT model for the entire Mississippi and Atchafalaya River basin (MARB). The specific objectives of this paper are to outline the challenges in (a) obtaining the data required for model calibration-validation; (b) setting up the model for the study area; (c) simulating the physical processes operating on the study area; (d) accounting the model parameter uncertainties; (e) technical limitations imposed by the model for application to large-scale studies.

2. Materials and Methods

2.1. Model Setup

The Hydrologic Unit Modeling of the United States (HUMUS) modeling system [39] was designed to quantify the environmental benefits of conservation practices and programs implemented in cultivated cropland throughout the United States. More details on the model setup are provided in this Sections 2.1.1–2.1.7). The computer model used to driving HUMUS was SWAT [4].

The SWAT model was developed to quantify the impact of land management practices on surface water quality in large, complex catchments [4,6,40,41]. It provides a continuous simulation of hydrological processes (evapotranspiration, surface runoff, percolation, return flow, groundwater flow, channel transmission losses, pond and reservoir storage, channel routing, and field drainage), crop growth and material transfers (soil erosion, nutrient and organic chemical transport and fate). The model is usually executed on a daily time step, although sub-daily climate data can also be used to drive SWAT. It incorporates the combined and interacting effects of weather and land management (e.g., irrigation, planting and harvesting operations and the application of fertilizers, pesticides or other inputs). Users build a SWAT simulation by dividing the watershed into sub-watersheds using topography or burned in stream networks. Each subwatershed is then typically further divided into hydrological response units (HRUs), which are unique homogeneous combinations of soil, topography and land cover. Although individual HRU's are simulated independently from one another, predicted water and material flows are routed within the channel network for a given subwatershed, which allows for large catchments with hundreds or even thousands of HRUs to be simulated. The subwatersheds used for this study were coincident with the boundaries of what are commonly called 8-digit hydrologic unit codes (HUCs) or 8-digit watersheds, which are also part of hierarchal HUC classification system for the U.S. [31].

2.1.1. Databases

The Hydrologic Modeling of the United States (HUMUS)/Soil and Water Assessment Tool (SWAT) system requires several databases such as land cover, soils, management practices and weather. For the present study, recently available data are processed to update the original HUMUS/SWAT databases and prepare the SWAT input files for the MARB [42].

2.1.2. Land cover

The model setup used the most recent version of the land cover data available during the period of study. The source of land cover data is the United States Geological Survey (USGS)—National Land Cover Data (NLCD) available at 30 m resolution for the United States [43]. The land cover categories include agriculture, urban, pasture, range, forest, wetland, barren, and water.

2.1.3. Soils

Each land cover category within an 8-digit watershed is associated with the most recent version of soil data available. Soil data required for SWAT were processed from the STaTeSoil GeOgraphic (STATSGO) database [44]. Each STATSGO polygon contains multiple soil series and the areal percentage of each soil series. Within a STATSGO polygon, the soil series with the largest area was identified and the associated physical properties for that soil series were extracted for SWAT. This procedure was followed for all of the 8-digit watersheds [42].

2.1.4. Topography

Topographic information on accumulated drainage area, overland field slope, overland field length, channel dimensions and channel slope were derived from the elevation data (DEM) of the previous HUMUS project [39]. The resolution of DEM used is approximately 90 m which would be adequate to capture the topography for large-scale studies like this. However, the resolution may not be sufficient to capture the elevation differences in steep mountainous areas.

2.1.5. Management

Management operations such as tillage, planting, harvesting, and applications of fertilizers, manure, pesticides and irrigation water along with timings or potential heat units are specified for various land uses in the management files. Management operations/inputs vary across regions. These data are gathered from various sources such as Agricultural Census Data and USDA—National Agricultural Statistics Service (NASS)'s agricultural chemical use data [42].

2.1.6. Weather

Measured daily precipitation and maximum and minimum temperature data sets from 1960 to 2006 were used in this study. The precipitation and temperature data sets were created from a combination of point measurements of daily precipitation, and maximum and minimum temperature [45] and Parameter-elevation Regressions on Independent Slopes Model (PRISM; [46,47]). The point measurements compose a serially complete (without missing values) data set processed from the National Climatic Data Center (NCDC) station records. PRISM is an analytical model that uses point data and a digital elevation model (DEM) to generate gridded estimates of monthly climatic parameters. PRISM data are distributed at a resolution of approximately 4 km². An approach has been developed to combine the point measurements and the monthly PRISM grids to develop the distribution of the daily records with orographic adjustments over each of the 8-digit watersheds [48]. Other data such as solar radiation, wind speed and relative humidity are simulated using the weather generator [49,50] available within SWAT.

2.1.7. Reservoirs

Many rivers in the country are impounded for various purposes such as flood control, irrigation, power generation, or recreation. Accounting for these impoundments while modeling stream flow is important to realistically capture the stream flow in different rivers. Therefore, reservoir input files were generated in the model setup and flow was routed through reservoirs where appropriate. Information on reservoirs such as drainage area, storage volumes, and surface areas at the principal and emergency spillway elevations, were derived from the U.S. Army Corps of Engineers National Inventory of Dams [51] and the USDA–NRCS. The version of SWAT model used in the present study allowed inclusion of information for one reservoir per 8-digit HUC simulated although the current version of the model is capable of simulating individual ponds and reservoirs. Therefore, the reservoir information is lumped to account for multiple reservoirs present in an 8-digit HUC. Only major reservoirs (in most cases multi-purpose reservoirs) that involve storing of a significant quantity of water is simulated. It does not include structures such as lock and dams typical to the MARB system.

Any monthly reservoir outflow data available from reservoirs is included as an input to the model to avoid some of the lumping effects of reservoirs in the simulation results. When the outflow data from reservoirs are not available, the SWAT model simulations use a simple target release-storage approach. The model accumulates flows during flood season and release the flow during non-flood season. Rules for storage and release can be defined with the help of input parameters. More details on this approach are available in the SWAT theoretical documentation [52].

2.2. Model Calibration and Validation

The study was designed to perform a suite of scenarios focused on pollutant reduction at various spatial scales using this model setup. Since the study is for the cultivated area in conterminous United States, flow and pollutant results from the model need to be calibrated and validated in many river basins throughout the MARB. This requires robust calibration procedures that are applicable in different river basins with varying hydrologic conditions. This approach is applicable for the entire MARB, which is comprised of the Tennessee, Ohio, Missouri, Upper Mississippi, Arkansas-White-Red, and Lower Mississippi MWRRs (Figure 1).

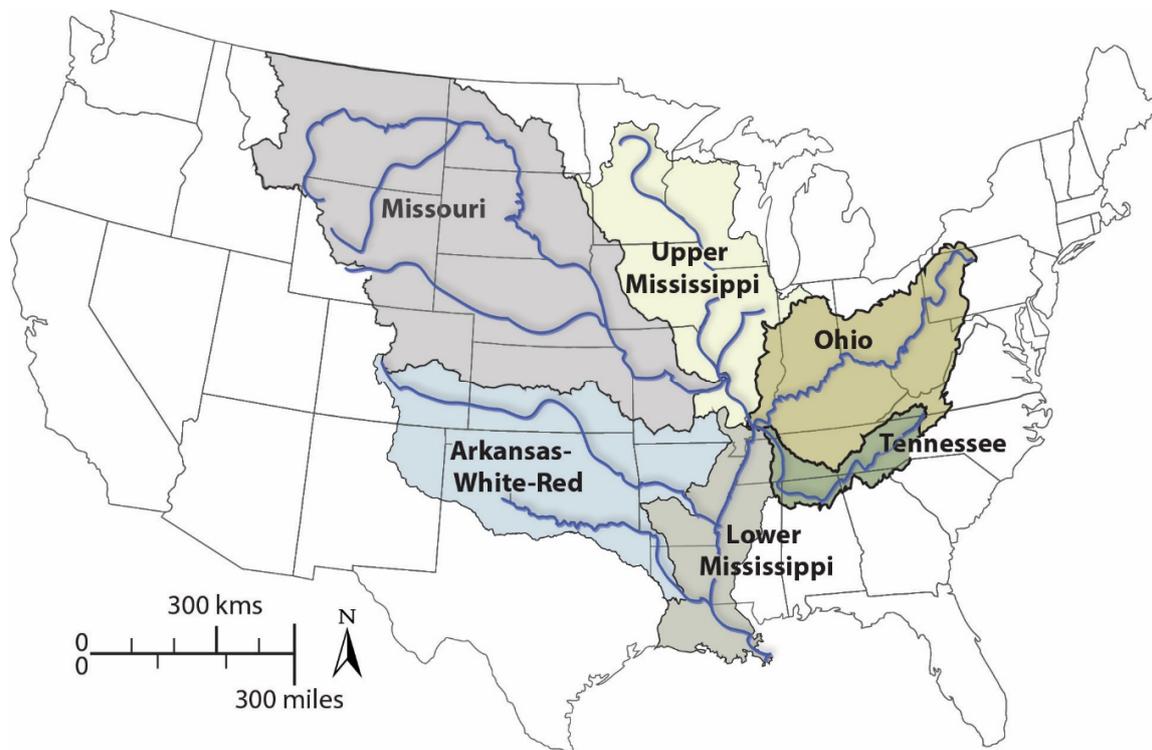


Figure 1. Location of the study area.

To perform pollutant modeling and calibrate the results from the model, reasonably accurate estimates of water runoff from upland (spatial patterns of runoff) and stream flow in rivers are required. In addition to matching predicted versus observed runoff and stream flow, it is essential to partition simulated runoff correctly into different hydrological pathways such as surface runoff and subsurface flow, or base flow. This requires a robust procedure to calibrate runoff/water yield as well as partition runoff into surface runoff and subsurface flow.

Calibration and validation need to be carried out using many years of data. Data from 1960–2006 was used for modeling. Data from 1961–1990 (30 years) was used for calibration and the remaining for validation. Average annual runoff from each MARB 8-digit watershed was used for spatial calibration. Annual average stream flows at selected gauging stations along major rivers and tributaries were used for temporal validation. Calibration of average annual runoff helps ensure local water balances were

accurate at the 8-digit watershed level. The temporal validation is performed to ensure whether the model is capturing annual variability in stream flows similar to that of observations.

Model calibration was carried out separately for each of the six MWRRs (Figure 1). This was done to (a) minimize calibration time, and (b) look at results closely. The model calibration goal was set to match predicted values from HUMUS to the estimated annual average water yield (estimated by the USGS) for each 8-digit watershed in each MWRR. This ensures good agreement on contribution of annual runoff spatially across all of the 8-digit watersheds in the MARB.

The annual average estimated water yield values (target values for spatial runoff calibration) are based on runoff contours for the nation prepared by Gebert et al. [53]. The source of information for the runoff contours was stream flows recorded from 5951 USGS gauging stations during 1951–1980 with an area not exceeding an 8-digit watershed. The runoff contour data reflects the runoff of tributary streams so that small-scale variations in runoff are represented with reasonable accuracy. Annual average water yield by HUC or 8-digit watershed is obtained by overlaying interpolated runoff contours representing average annual runoff (in inches) for the conterminous United States with the HUC map. The resulting annual average runoff values are used as target values for calibrating the predicted annual average water yield from SWAT.

Arnold et al. [54] developed a digital filter technique to partition the stream flow between surface runoff and base flow. In this technique, the base flow ratio is the ratio of subsurface flow to total flow. To estimate subsurface flow, the ratio is multiplied by the observed water yield. Santhi et al. [55] have estimated the base flow (subsurface flow) ratio for all of the 8-digit watersheds in the United States using the digital filter technique. To obtain subsurface flow for an 8-digit watershed in a river basin, the base flow ratio was multiplied with the corresponding water yield for the 8-digit watershed. The difference between water yield and subsurface flow is used as surface runoff. The data obtained this way are used as target values to calibrate runoff/water yield, subsurface flow, and surface runoff.

A fully automated calibration procedure was developed to carry out calibration. The procedure uses nine parameters to calibrate average annual water yield or total runoff, surface runoff, and subsurface flow. The procedure uses the linear interpolation method to obtain a better value of each model parameter. The calibration process was carried out in three major steps with adjustments of: (1) water yield, (2) surface runoff, and (3) subsurface flow. The major part of the flow calibration was achieved by the automated procedure. Some manual adjustment of parameters was made in addition to the automated procedure where needed. However, no adjustment of parameters was made to calibrate predicted stream flow results to match that of observations at selected gauging stations along major rivers and their tributaries. This is based on the assumption that adequate capturing of the spatial variation of runoff in the different 8-digit watersheds will also capture the time series of stream flow at different river reaches. Additional calibration may be needed for daily flow.

2.2.1. Details on Automated Flow Calibration Procedure

The first calibration step requires the arrangement of all the input files necessary for running the model and the calibration procedure. Initially, the model will be executed without any parameter adjustment. The results at this stage will be analyzed to check the quality of existing results with respect to calibration target values. The next step is to decide what parameters should be adjusted for each 8-digit watershed. The automated calibration procedure spatially calibrates the following SWAT model parameters so that the simulated average annual water yield, sub-surface flow and surface runoff match the corresponding target values for each 8-digit watershed in each 2-digit MWRR. The calibration goals are to maintain the differences between the simulated and target values within 10 percent for surface runoff, 10 percent for subsurface flow, and 20 percent for water yield.

- HARG_PETCO, a coefficient used to adjust potential evapotranspiration (PET) estimated by the Hargreaves method [56,57] and calibrate the runoff/water yield in each 8-digit watershed. In the Hargreaves method, PET is a function of temperature and terrestrial radiation. This coefficient is

related to radiation and can be varied to account for the differences in PET in different parts of the river basin depending on weather conditions [57].

- Soil water depletion coefficient (CN_COEF), a coefficient used in the curve number method to adjust the antecedent moisture conditions on surface runoff generation.
- Curve Number (CN), which is used to adjust surface runoff and relates to soil, land use, and hydrologic condition at the HRU level.
- Groundwater re-evaporation coefficient (GWREVAP) controls the upward movement of water from the shallow aquifer to the root zone due to water deficiencies in proportion to potential evapotranspiration. This parameter can be varied depending on the land use/crop. The re-evaporation process is significant in areas where deep-rooted plants are growing and affects the groundwater and the water balance.
- GWQMN—Minimum threshold depth of water in the shallow aquifer to be maintained for groundwater flow to occur to the main channel.
- Soil available water-holding capacity (AWC), which varies by soil at HRU level. It is the capacity factor of soil signifying the extent of water storage in a given depth of soil (fraction)
- Soil evaporation compensation factor (ESCO), which controls the depth distribution of water in soil layers to meet soil evaporative demand. This parameter varies by soil at the HRU level.
- Plant evaporation compensation factor (EPCO), which allows water from lower soil layers to meet the potential water uptake in upper soil layers and varies by soil at the HRU level.

Parameters CN_COEF, and AWC are adjusted when the water yield meets the calibration goal but the surface runoff and sub-surface flow do not. Any change in parameters CN, ESCO, and EPCO affects surface runoff. Similarly changes to the GWQMN, GWREVAP, and REVAPMN parameters affects base flow. To achieve the calibration goals, the above input parameters were adjusted within literature reported ranges for the SWAT model [37,40], as well as drawing on “soft data” sources such as expert knowledge and previous experiences [58].

Statistical measures such as mean, standard deviation, coefficient of determination (R^2), and Nash–Sutcliffe prediction efficiency (NSE) [59] were used to evaluate the model performance of predicting annual flows. If the R^2 and NSE values were less than or very close to zero, the model prediction is considered unacceptable or poor. If the values are 1.0, then the model prediction is perfect. Values greater than 0.6 for R^2 and greater than 0.5 for NSE were considered acceptable [37,60,61]. More details on the flow calibration procedure and validation of flow results are available in Kannan et al. [38] and Santhi et al. [55] respectively.

2.2.2. Annual Flow Validation at Stream Gages

For validation of stream flow, 32 USGS stream gages were selected along major rivers and tributaries throughout the MARB. For each USGS 2-digit HUC a minimum of five gauges (except Tennessee River basin) were chosen to have brevity in presentation of calibration/validation results. The gauges were chosen to satisfy the following conditions as much as possible: (a) cover the entire river basin by accounting for the major rivers and tributaries, (b) continuous data availability for the period of interest, (c) availability of quality controlled data, and (d) unaffected by large abstractions/impoundments.

The annual average stream flow for the years 1961–2006 were included for the analysis. It should be noted that the calibration was carried out for annual runoff at the 8-digit watershed level and predicted stream flow was not calibrated to match observations at gauging stations. However, the results obtained were summarized for the calibration and validation periods separately, to maintain similar analysis periods per the annual runoff calibration at the 8-digit watershed level.

3. Results

There is a close correlation between model predicted annual water yield and the target values. This is evident from the model performance evaluation statistics (Table 1). When analyzed spatially throughout the MARB, there is a close correspondence between predicted and target annual mean runoff (Figure 2 and Table 1). In summary, the model estimated annual mean runoff values appear satisfactory and acceptable. Similar results were obtained for both the surface runoff and base flow (Figures 3 and 4). However, the modeled base flow results appear inferior relative to the predicted surface runoff and water yield. There are several reasons that can be attributed for these results. First, the non-linear nature of some of the model parameters controlling base flow (e.g., GWQMN); when such non-linear parameters were used with a simple linear interpolation method to obtain a better value to take care of an under or over-estimation situation, the parameter was not adequately adjusted (it is either under or over-parameterized), Second, bias in the target values used for calibration of base flow, which could be a result of: (1) non-availability of data, (2) base flow separation method used to obtain base flow, and (3) the interpolation method used to obtain gridded base flow values from point estimates [55]. More details on results by individual river basins and the challenges encountered in performing flow calibration and validation are discussed in the discussion section.

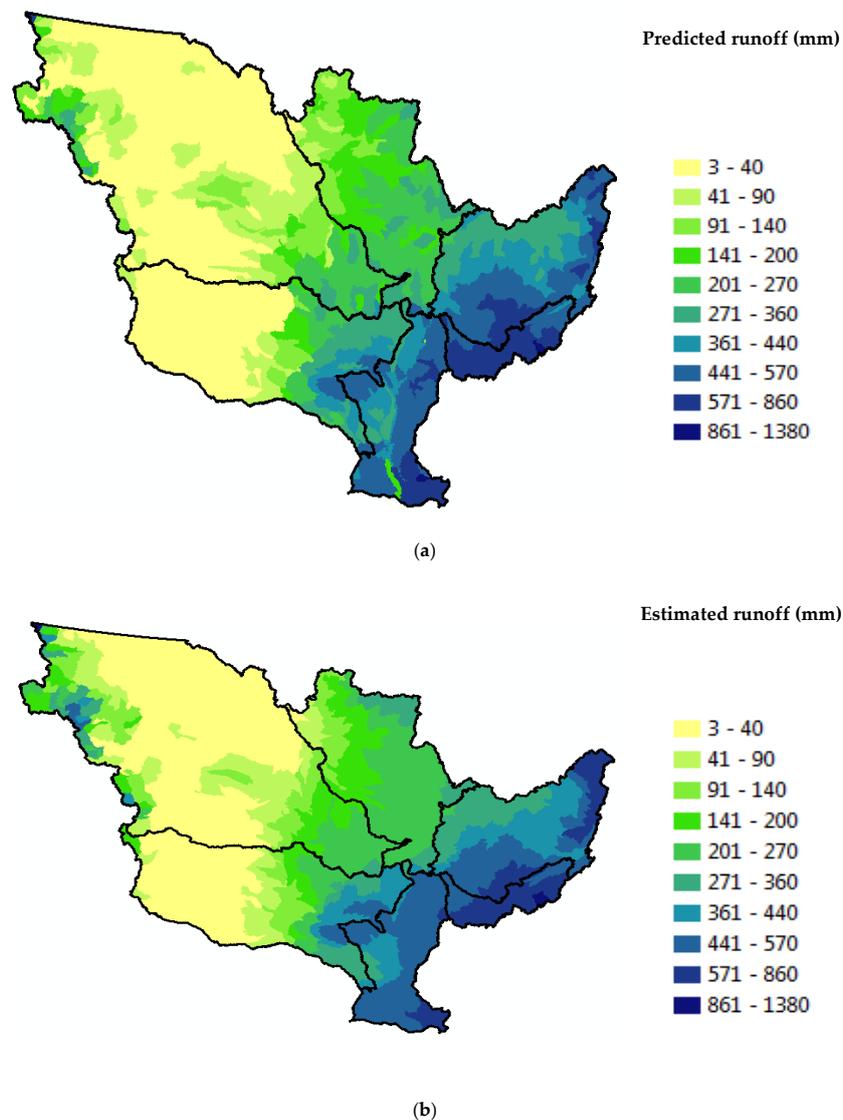
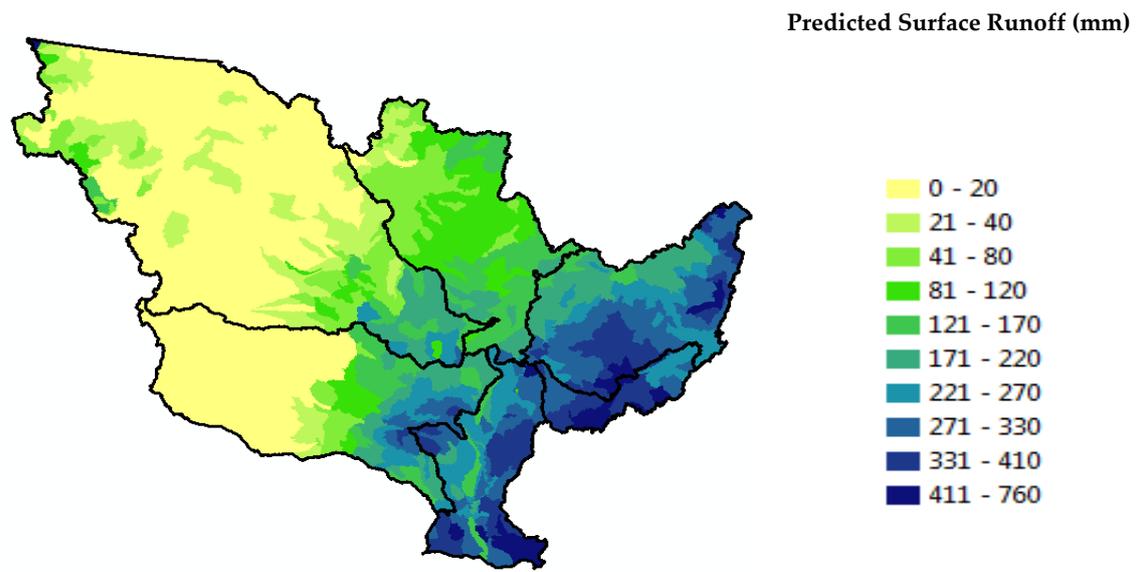
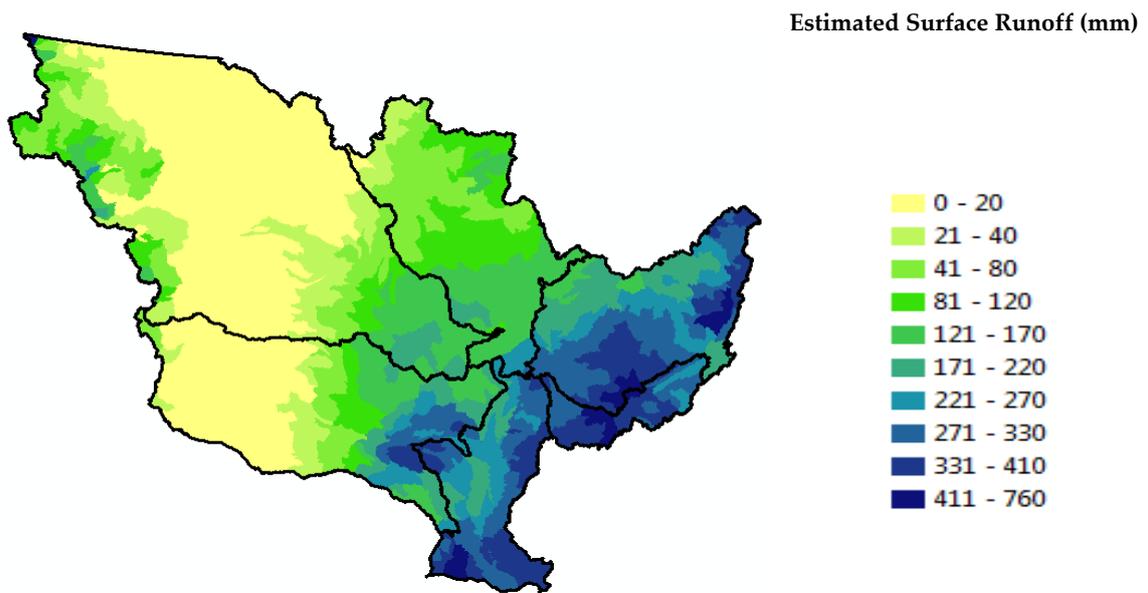


Figure 2. Comparison of total runoff: (a) Model predictions; (b) Calibration target.



(a)



(b)

Figure 3. Comparison of surface runoff: (a) Model predictions; (b) Calibration target.

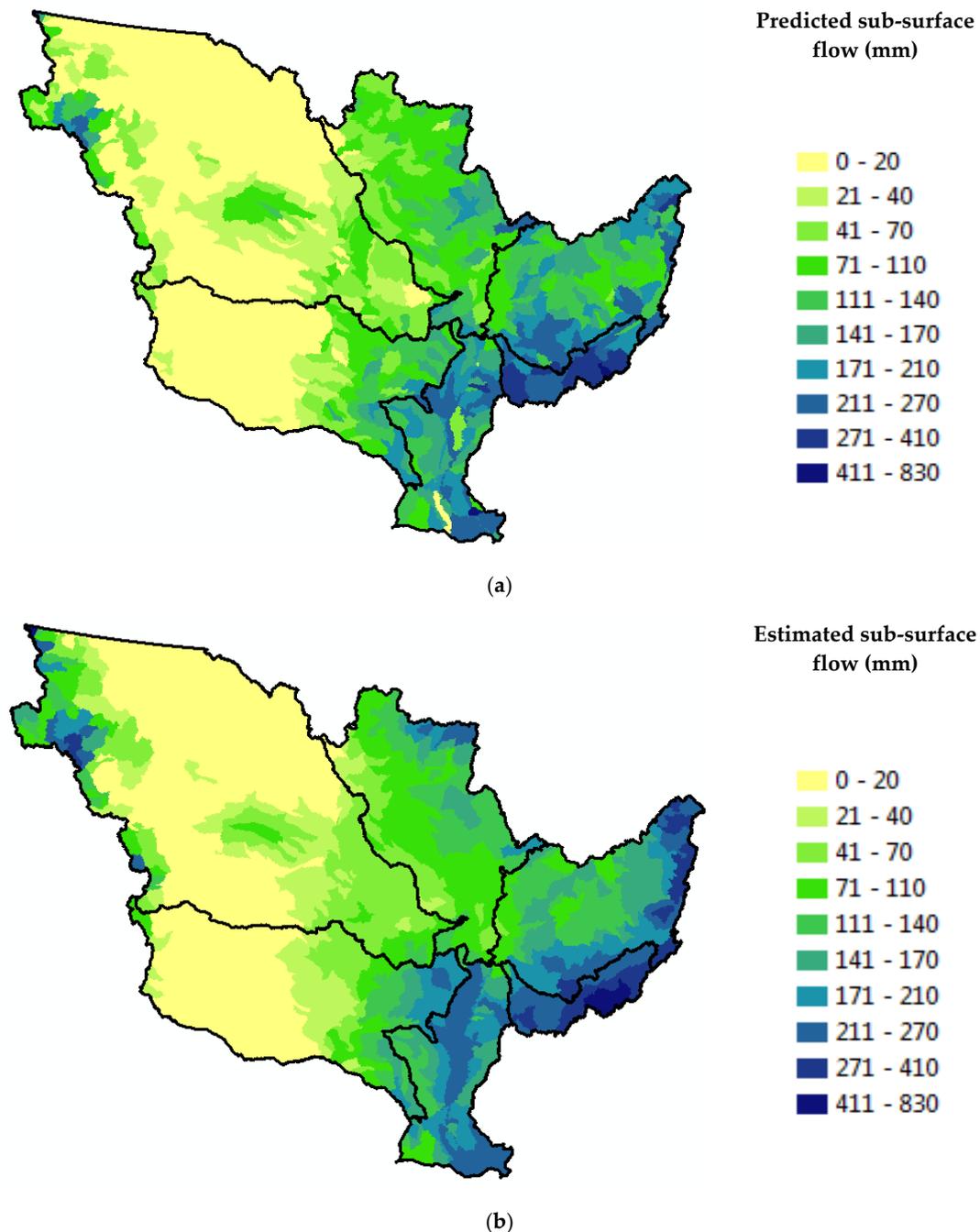


Figure 4. Comparison of sub-surface flow: (a) Model predictions; (b) Calibration target.

Average annual stream flow results from the model are completely satisfactory in terms of all the model performance evaluation statistics for 22 out of 32 gauging stations analyzed (Table 2 and Table 4). This was true for both calibration and validation periods. For 7 of the gauging stations (1. Ohio river at Sewickley, PA, 2. Kanawha river at Charleston, WV, 3. Minnesota river at Jordan, MN, 4. White river at Devalls Bluff, AR, 5. Yellowstone River at Sidney, MT, 6. Missouri river at Omaha, NE and 7. Missouri river at Hermann, MO) the results are acceptable with some exceptions. For these gauges, either the calibration or the validation period shows acceptable results. For the remaining three stations (1. Atchafalaya River at Melville, LA, 2. Platte River at Louisville, NE and 3. Arkansas River at Arkansas City, KS) the results were not acceptable. The probable reasons for the inferior stream flow results for some gauging stations are provided in the discussion section.

Table 1. Annual average runoff of 8-digit hydrological unit codes (HUCS) in Mississippi River Basin.

Parameter	Mean (mm)		Standard Deviation (mm)	
	Predicted	Observed	Predicted	Observed
Surface runoff	130.9	133.2	132.1	124.2
Base flow	85.8	95.9	82.0	97.1
Water yield	216.7	229.1	204.2	209.1
		R²	Nash and Sutcliffe Efficiency	
Surface runoff		0.94	0.94	
Base flow		0.76	0.75	
Water yield		0.93	0.93	

Table 2. Model performance evaluation for estimation of annual average flow results.

Gauge Details			Calibration Period (1961–1990)		Validation Period (1991–2006)	
River	Location	Area (km ²)	R ²	NSE	R ²	NSE
Allegheny	Natrona, PA	29,551.8	0.97	0.87	0.96	0.94
Ohio	Sewickley, PA	50,504.8	0.92	0.47	0.97	0.79
Kanawha	Charleston, WV	27,060.2	0.91	0.69	0.89	0.41
Ohio	Greenup, KY	160,579.3	0.93	0.78	0.93	0.86
Ohio	Metropolis, IL	525,767.6	0.96	0.90	0.95	0.88
Tennessee	Whitesburg, AL	66,329.6	0.97	0.76	0.94	0.93
Tennessee	Savannah, TN	85,832.0	0.97	0.97	0.99	0.99
Minnesota	Jordan, MN	41,957.8				
Mississippi	Clinton, IA	221,704.0	0.84	0.81	0.88	0.49
Iowa	Wapello, IA	32,374.9	0.83	0.83	0.76	0.64
Illinois	Valley City, IL	69,264.1	0.92	0.87	0.96	0.76
Mississippi	Grafton, IL	444,183.0	0.94	0.59	0.98	0.81
Mississippi	Thebes, IL	1,847,179.4	0.94	0.91	0.93	0.78
White	Devalls Bluff, AR	60,605.7	0.36	0.28	0.76	0.73
Ouachita	Camden, AR	13,882.3	0.96	0.95	0.96	0.90
Mississippi	Vicksburg, MS	2,964,241.3	0.94	0.99	0.88	0.99
Mississippi	St Francisville, LA	2,914,513.5	0.88	0.86	0.91	0.71
Atchafalaya	Melville, LA	241,646.0	0.93	−2.04	0.75	−3.00
Yellowstone	Sidney, MT	177,134.5	0.65	−0.07	0.71	0.53
Missouri	Culbertson, MT	237,131.5	0.99	0.77	0.97	0.92
Missouri	Bismark, ND	482,773.8	0.79	0.79	0.99	0.97
Missouri	Yankston, SD	723,901.6	0.97	0.78	1.00	0.99
Missouri	Omaha, NE	836,048.1	0.75	0.29	0.98	0.96
Platte	Louisville, NE	221,107.3	0.74	−6.05	0.63	−3.75
Osage	St Thomas, MO	37,554.8	0.95	0.95	0.95	0.94
Missouri	Hermann, MO	1,353,268.7	0.92	0.32	0.95	0.80
White	Calico Rock, AR	25,848.1	0.78	0.58	0.82	0.74
Arkansas	Arkansas City, KS	113,216.1	0.38	−1.23	0.77	0.38
Canadian	Calvin, OK	72,395.3	0.82	0.75	0.75	0.71
Arkansas	Murray Dam, AR	409,575.5	0.95	0.92	0.93	0.89
Red	Gainsville, TX	79,725.0	0.80	0.69	0.85	0.82
Red	Alexandria, LA	174,824.2	0.88	0.87	0.88	0.61

4. Discussion

4.1. Discussion on Average Annual Water Yield/Total Runoff of 8-Digit Watersheds in MARB

Among the six MARB MWRRS, inferior performance occurred in matching the annual runoff with that of the calibration targets for the Missouri and Arkansas-White-Red Rivers. Although the

overall runoff match at the basin outlets are satisfactory, about 50% of the total sub-basins did not meet the calibration goal for annual runoff/water yield for these two MWRRs. The mismatch is more pronounced in some places within the river basins (Figures 5 and 6). A detailed analysis revealed some reasons for this behavior.

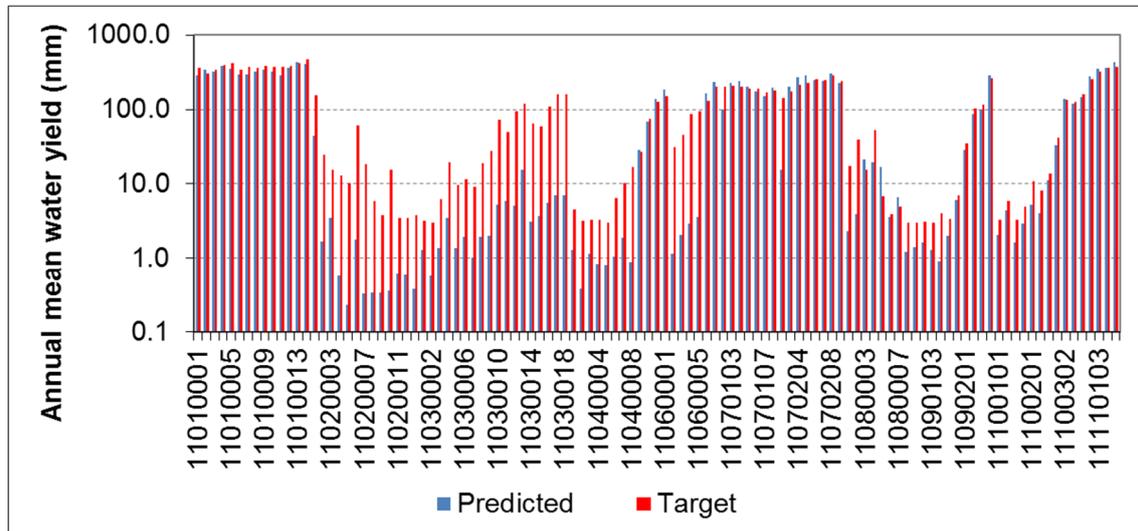


Figure 5. Annual water yield for Arkansas-White-Red river basin.

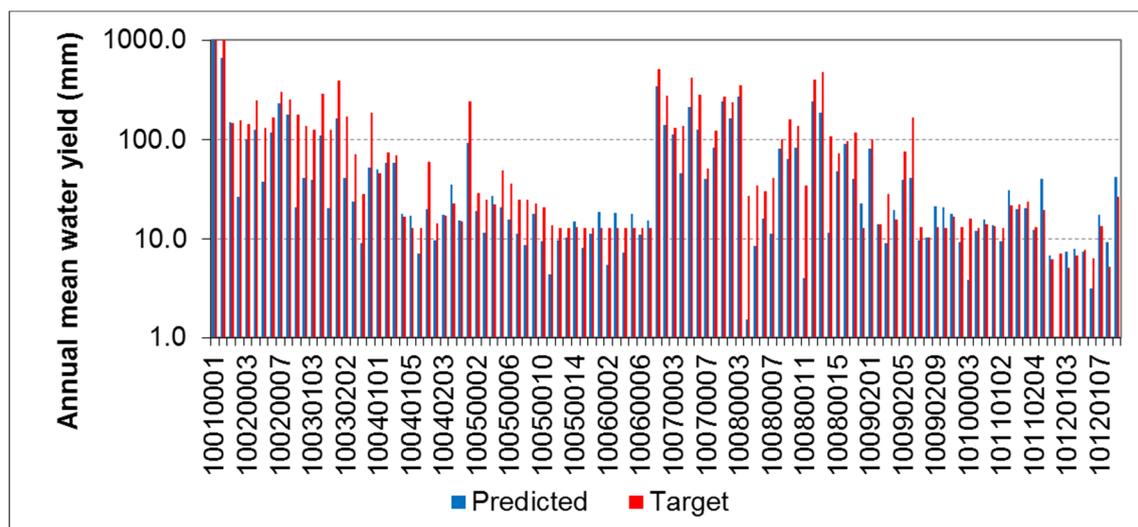


Figure 6. Annual water yield for Missouri river basin.

The Missouri River Basin contains 310 sub-basins (8-digit watersheds) with a wide variation of topography and climate patterns. Mean annual precipitation ranges from 215 mm to 1640 mm. A similar range of variation exists in annual runoff calibration targets (3 mm to 1370 mm). As well, a similar trend was observed in Arkansas-White-Red river basin, where the mean annual precipitation varies from 285 mm to 1400 mm and runoff varies from 3 mm to 545 mm. Capturing these high hydrologic variations in the river basin with a limited parameter set was indeed a daunting task for the simple flow calibration procedure.

It should be noted that the flow calibration procedure identifies the subbasins to be calibrated based on the percentage difference between predicted and calibration target values. This is problematic for subbasins with relatively small annual runoff values (e.g., <10 mm). If the predicted runoff is 6 mm and the calibration target is only 3 mm, the percentage difference between predictions and target

is 100. Although the predicted and calibration target for that particular sub-basin are close enough, statistically the subbasin is un-calibrated which is reflected in the model performance.

One of the mismatches occurs in the upper reaches of the Missouri river (Table 3) including the Milk River and Marias River tributaries. The Milk River originates in the Rocky Mountains and passes through the Province of Alberta, Canada and the State of Montana in the USA. In contrast, the Marias River originates in the Blackfeet Indian Reservation in Glacier County, which is located in northwestern Montana. Elevations in the reservation range from 1000 m (3400 feet) to 2763 m (9066 feet). The eastern part of the reservation is mostly open hills of grassland, while a narrow strip along the western edge is covered by forests of fir and spruce. Two possible reasons could be attributed to the under-estimation of surface runoff, base flow, and water yield for these two rivers. First, both of the rivers passes through a highly sloping terrain, where the predicted SWAT stream flows resulted in an under estimation of surface runoff and sub-surface flow, and therefore water yield. An examination of parameters governing the snowfall-snow melt mechanisms in the model, and the snow fall (estimated based on temperature and precipitation data) snow melt results from the model suggest snow fall and snow melt mechanisms were captured adequately in both of these river basins. Inadequate capturing of slope could have under-estimated surface runoff, over-estimated soil water storage, and therefore under-estimated the total runoff. It is understandable given, the scale and scope of the project. A better representation of topography for the upper reaches of Missouri and its tributaries would have improved the runoff estimates from the model. Second, the Milk river drains some areas in the Canadian provinces of Alberta and Saskatchewan which are not included in the model setup; i.e., there is some un-accounting of drainage areas in the model setup.

Table 3. Major mismatches between the Soil and Water Assessment Tool (SWAT) predictions and calibration targets.

River Basin	Location	Model Estimation Problem			Annual Precipitation (mm)	Target Water Yield (mm)
		Surface Runoff	Base Flow	Water Yield		
Missouri headwaters	Montana-Wyoming	Over	Under	Under	500	194
Missouri-Marias	Montana	Under	Under	Under	430	171
Milk	Montana	Under	Under	Under	312	36
Upper Yellowstone	Montana-Wyoming	Under	Under	Under	541	241
Big Horn	Wyoming	Under	Under	Under	384	161
Kansas	Nebraska-Missouri	Over	Over	Over	789	115
Middle Arkansas	Colorado, Kansas	Under	Under	Under	637	56
Upper Cimarron	CO, KS, NM, OK	Under	Under	Under	445	6
Arkansas-Keystone	Kansas, Oklahoma	Under	Under	Under	754	90

Another high mismatch between the SWAT-predicted versus observed stream flows occurred for the Upper Yellowstone River, including the Boulder River and Clarks Fork River (or Clarks Fork Yellowstone river) tributaries. The Boulder River rises in the Gallatin National Forest in the Absaroka Range in southern Park County, Montana. It flows north through mountainous canyons, a cataract, and northwest in a widening valley past McLeod, Montana. It joins the Yellowstone River at Big Timber, Montana. The Clarks Fork River begins northeast of Cooke City in southern Montana, within the Gallatin National Forest and the Beartooth Mountains (southwest of Granite Peak). It flows southeast into the Shoshone National Forest in Northwest Wyoming, east of Yellowstone National Park, then northeast back into Montana. It joins the Yellowstone southeast of Laurel. Both the rivers pass through a highly sloping terrain. Therefore, a better representation of topography for the Upper Yellowstone and its tributaries would have improved the runoff estimates from the model.

The Kansas River begins at the confluence of the Republican and Smoky Hill rivers. It flows generally eastward to join the Missouri River at Kaw Point in Kansas City. Much of the drainage of the river lies within the Great Plains. During its course, the river falls less than 38 cm/km (2 feet per mile). The river valley averages 4.2 km (2.6 miles) in width. The river is shallow, wide, with a gentle

channel slope. As well the drainage is slow and generally free flowing. However, its tributaries are dammed for flood control. There are several possible reasons for over-estimation of runoff in Kansas River. One possible problem could be over-estimation of the carrying capacity of the river channel because of inadequate representation of the river channel width, depth and slope. This in turn could be due to the quality of Digital Elevation Model (low resolution) used to obtain the model input files defining river channel and tributaries. Another possible issue is inadequate representation of the flood control reservoirs present in the Kansas River system. In the present model setup, reservoirs are modeled on the basis of only one input file per 8-digit watershed, which represents all the reservoirs present within the 8-digit watershed. This simplification may be inadequate to capture the storage and release of water in the Kansas River system. Moreover, reservoir release rates were not available for each impoundment.

The Arkansas river flows through the states of Colorado, Kansas, Oklahoma, and Arkansas. The mismatch between runoff predictions and the calibration target occur in the 8-digit watersheds covering the middle stretches of the river which are located entirely within Kansas. At this stretch the river behaves like a typical Great Plains river with a wide, shallow channel that is subject to seasonal flooding. Groundwater is pumped for irrigating corn and wheat in the western parts of Kansas. The irrigation results in return flows to stream. This is captured in the model setup through base flow values from APEX on per ha basis, for each 8-digit watershed. However, this could not capture the irrigation and subsequent return flows on a HRU basis, which will result in base flow within the 8-digit watershed. The same problem could occur for other intensively irrigated areas in the Mississippi watershed system.

Flow under-estimation occurs in the upper reaches of the Cimarron River. The headwaters of Cimarron River flow from the Johnson Mesa west of Folsom in northeastern New Mexico. The river enters the Oklahoma Panhandle near Kenton, crosses the southeastern corner of Colorado into Kansas through the Cimarron National Grassland, re-enters the Oklahoma Panhandle, re-enters Kansas, and finally returns to Oklahoma where it joins the Arkansas River at the Keystone Reservoir west of Tulsa, Oklahoma. In New Mexico and the western most part of Oklahoma the river is known as the Dry Cimarron River. The Dry Cimarron River is not completely dry, but sometimes the water disappears entirely under the sand in the river bed.

The Cimarron National Grassland is a National Grassland located in Morton County, Kansas, United States. The terrain is mostly flat, sloping downward west to east. Vegetation is mostly shortgrass prairie grassland, dominated by sand sagebrush in salty soils. Groves of Cottonwood and other trees are found near the river. The climate of the National Grassland is semi-arid, receiving about 18 inches (460 mm) of precipitation annually, mostly in summer. High winds are common and further desiccate the soil. The area has hot summers and cold winters. The climate, topography, soils, and land cover suggest the extent of dryness of the Upper Cimarron River. A verification of model input files revealed adequate capturing of climate and topography. Therefore, the flow under-estimation could be a result of over-estimation of ET. The over estimation of ET could have come from over-representation of soil water storage (and therefore more evaporation) through soil properties or inadequate representation of the vegetative cover present in the 8-digit watersheds covering Upper Cimarron river.

4.2. Discussion on Average Annual Streamflow at Selected Gauging Stations in MARB

The R^2 values during calibration and validation periods are acceptable for the Ohio River at Sewickley, PA. However, the NSE during calibration is 0.47 (<0.5) and therefore could not be accepted. An investigation of the means between the predicted and observed stream flows revealed under-estimation of flow (Table 3). From the analysis of monthly flow results (not shown here), it appears that the monthly peaks were under-estimated for this station. In addition, it also appears that the under-estimation of stream flow is a result of consistent under-estimation (13%) of runoff at the 8-digit watershed level (Predicted: 469 mm versus calibration target 539 mm). It should be noted that the predicted runoff of the 8-digit watersheds contributing flow to this station were still

within 20% of the calibration target. For the Kanawha River at Charleston, WV, the R^2 values during the calibration and validation periods are acceptable. However, the NSE during validation is 0.41 (<0.5) and thus should again be considered inadequate per the suggested criteria of Moriasi et al. (2007; 2015). A comparison of the means between the predicted and observed stream flows revealed under-estimation of flow (Table 3). It appears that the under-estimation of stream flow is a result of consistent under-estimation of runoff (13%) at the 8-digit watershed level (Predicted: 468 mm versus calibration target 538 mm). However, the predicted runoff of the 8-digit watershed that contributed flow to the Kanawha River were again still within 20% of the calibration target.

For the Minnesota River at Jordan, MN the R^2 values during calibration and validation periods are acceptable. However, the NSE during validation is not acceptable (0.49, which is less than 0.5). Overestimation and underestimation of stream flows were seen as trends during calibration and validation periods respectively (Table 3). There are 12 8-digit watersheds contributing flow to this station. Annual runoff from these 8-digit watersheds shows over-estimation and therefore results in over estimation of stream flows during calibration period. Therefore, it appears that the problem in modeled stream flow during validation period is due to in-stream issues. There are five major dams in Minnesota River basin mainly built for wildlife management; recreation and flood control [62]. Three major flood events were recorded in the river basin during the validation period in 1993, 1997, and 2001, especially around Granite Falls dam [62]. Inadequate accounting of major dams in the river basin and inadequate reproduction of the above-mentioned flood events are probably responsible for the poor model validation statistics for the Minnesota River.

Both the R^2 and NSE values during the validation period were acceptable for the White River at Devalls Bluff, AR. However, the statistical values are below the acceptable standards during the calibration period (Table 4). It should be noted that this station belongs to lower reaches of the White River within the Lower Mississippi region. From Table 3 it appears that the annual stream flow patterns and magnitude are well reproduced by the model for the upstream reaches of the same river. Therefore, inferior model performance at this gauging station is probably due to the contributing drainage area within the lower reaches. There are four 8-digit watersheds contributing drainage to the lower reaches of the river. No major dams are present in these four reaches. Therefore, in-stream flow modification could not have caused poor model performance. Adjustment of in-stream parameters could not have helped either. These issues point out the likely problems with annual runoff for the four contributing 8-digit watersheds. Although the predicted and calibration target annual runoff are close for the combined four 8-digit watersheds (440 mm predicted versus 427 calibration target), large differences occurred between the predicted runoff and calibration target values for two of the 8-digit watersheds. Under-estimation (12%) occurred for one of the 8-digit watersheds and over-estimation (21%) for the other 8-digit watershed. Mismatches in the magnitude of annual runoff from these two HUCs appears to be the problem behind the poor R^2 and NSE during calibration period.

The stream flow patterns were reproduced well by the model for the Yellowstone River at Sidney, MT, (acceptable R^2 values—Table 4). Although, the magnitudes are consistently under-estimated for the calibration and validation periods, the unacceptable NSE values occurred only during calibration. The main reason for the stream flow under-estimation during calibration is under-estimation of monthly flow peaks (around 20 mm; not shown here), especially during 1961–1980. There seems to be a steady decrease in the monthly flow peaks of the river after 1980. This might be due to increased water abstractions from the river. The model does not pick up this unique trend during 1961–1980 of the calibration period, which resulted in poor NSE values. Another reason that could be attributed to under-estimation of stream flow is under-estimation of annual runoff in the Upper Yellowstone River covering about eight 8-digit watersheds as described above.

Table 4. Comparison of predicted and observed annual average flow results.

Gauge Details		Annual Mean Flow (mm)			
		Calibration Period (1961–1990)		Validation Period (1991–2006)	
River	Location	Predicted	Observed	Predicted	Observed
Allegheny	Natrona, PA	555.0	589.1	562.8	583.6
Ohio	Sewickley, PA	528.9	598.9	549.6	611.9
Kanawha	Charleston, WV	478.6	436.4	486.4	423.5
Ohio	Greenup, KY	467.7	504.5	568.1	598.1
Ohio	Metropolis, IL	466.9	490.6	486.2	511.6
Tennessee	Whitesburg, AL	583.1	591.7	576.4	552.8
Tennessee	Savannah, TN	582.9	585.6	610.1	600.2
Minnesota	Jordan, MN	108.8	98.9	122.3	161.2
Mississippi	Clinton, IA	175.7	196.8	191.2	237.6
Iowa	Wapello, IA	253.3	233.3	269.6	287.9
Illinois	Valley City, IL	270.8	302.8	263.9	314.0
Mississippi	Grafton, IL	221.0	227.6	229.6	253.3
Mississippi	Thebes, IL	111.8	103.1	114.5	111.6
White	Devalls Bluff, AR	333.4	338.4	366.1	371.1
Ouachita	Camden, AR	512.1	502.8	492.3	467.6
Mississippi	Vicksburg, MS	194.7	185.6	201.0	190.9
Mississippi	St Francisville, LA	142.6	148.1	147.1	161.7
Atchafalaya	Melville, LA	1107.7	850.1	1101.1	856.2
Yellowstone	Sidney, MT	69.2	90.4	64.4	75.7
Missouri	Culbertson, MT	46.3	41.6	34.8	32.9
Missouri	Bismark, ND	45.7	45.2	39.0	37.8
Missouri	Yankston, SD	32.9	34.8	23.6	24.5
Missouri	Omaha, NE	41.0	35.7	39.3	37.5
Platte	Louisville, NE	56.4	28.0	54.1	31.2
Osage	St Thomas, MO	265.9	263.3	363.9	373.5
Missouri	Hermann, MO	68.8	56.0	68.8	59.6
White	Calico Rock, AR	280.1	342.9	301.3	340.6
Arkansas	Arkansas City, KS	22.1	15.4	23.8	16.3
Canadian	Calvin, OK	15.0	18.2	19.9	23.4
Arkansas	Murray Dam, AR	97.6	102.7	97.9	102.4
Red	Gainsville, TX	26.5	33.4	40.1	39.0
Red	Alexandria, LA	184.6	178.7	180.5	186.8

The stream flow patterns and magnitudes were reproduced well by the model for the Missouri River at Omaha, NE, except during calibration period, which resulted in an inferior NSE value. A further examination of the annual flow patterns revealed the mismatches in magnitude were mostly seen during 1961–1968. In addition, there is occasional over-estimation of monthly flow peaks (not shown) especially during calibration period. This could be due to an inadequate representation of reservoirs by the model. It should be noted that there are multiple reservoirs within the river basin constructed in different years. Sometimes more than one reservoir exists within an 8-digit watershed. In the model setup, only one reservoir could be defined for each 8-digit watershed. Therefore, simulation problems can result when multiple reservoirs built in different years have to be aggregated to a single reservoir. This kind of problem could have caused flow mismatches during a specific period in the hydrograph. For the Missouri River at Hermann, MO, the poor NSE during the calibration period was due to consistent over-estimation of stream flow, which included consistent over-estimation of peaks. It is probable that there are multiple reasons for this over estimation because Hermann, MO is the outlet of the entire river basin. One obvious reason for flow over-estimation at this gauging station could be the over-estimation of monthly flow peaks at Omaha, NE (which is upstream of Hermann, MO) that likely would have carried over to the Hermann gauging station.

The Atchafalaya River is a distributary of the Red and Mississippi Rivers. Although the river has its own basin and delta, the flow through the river is very much controlled. It carries the discharge from Red river as well as about 30% of the flow of Mississippi river [63]. Three 8-digit watersheds form the drainage basin for this river. It is not a surprise that the HUMUS model setup did not capture the flow patterns and magnitudes for the Atchafalaya River because the flow control rules and rates of water release to Atchafalaya from the Red and Mississippi Rivers were not included in the present model setup.

In the case of the Platte River, the R^2 values are in acceptable range suggesting the model is able to reproduce the annual stream flow patterns. However, the NSE values are not acceptable which suggests problems for the model in reproducing the flow magnitudes. An examination of the magnitudes reveals consistent over-estimation of flows for both the calibration and validation periods. The Platte River is a large system, drained by approximately 52 8-digit watersheds. The river is formed by the confluence of the North Platte River (drains 14 HUCs) and South Platte River (drains 18 HUCs). The important tributaries include the Loup River (drains 10 8-digit HUCs) and Elkhorn River (drains 4 HUCs) [64]. The North Platte River is dammed for water storage and irrigation. Large quantities of water are taken from the river for irrigating cultivated areas in Wyoming and Nebraska. The South Platte River is extensively dammed for water storage, public supply, and irrigation in Colorado. Given the information on the course of the river, water impoundments and abstractions, there are several reasons that could be considered for the mismatch in flow magnitude between predictions and observations for the Platte River. One issue is the fact that the cultivated area and irrigation applications were modeled with APEX. The irrigation return flows area were accounted for in the base flow and then read into HUMUS-SWAT. However, the total quantity of irrigation flows were not discounted from the stream flow in the model setup. Another problem was inadequate accounting of reservoirs in the model setup. The number of reservoirs represented in the model and the modeling approach used to simulate those reservoirs were not adequate for capturing the flow routing mechanisms occurring in Platte Reservoir systems, especially for the South Platte River. Third, the over-estimation of annual runoff from Loup River is key, because it is an important tributary of the Platte River. It should be noted here that annual runoff estimated by the model for the 8-digit watersheds draining the North and South Platte Rivers were not over-estimated.

None of the model performance evaluation measures are acceptable for the Arkansas River, except the R^2 values during the validation period (Table 4). The model over-estimated annual stream flow consistently during both the calibration and validation periods. The drainage for the gauging station at Arkansas City, KS comes from the Upper and Middle Arkansas River reaches, which cover nearly 31 8-digit watersheds. Although the annual runoff from these contributing watersheds is under-estimated, the stream flow is over-estimated at Arkansas City, KS. This suggests in-stream issues associated with the flow over-estimation problem. An examination of channel routing files revealed high channel slope values for some reaches of the Upper Arkansas River. However, this itself could not have caused such high consistent over-estimation. Existing literature on this river suggests a couple of other likely reasons for the flow over-estimation problem. The first, issue pertains to reduced base flow contributions from the alluvial aquifers. The Arkansas River overlays the Ogallala Aquifer, where water is withdrawn for irrigation at a much higher rate than recharge. This causes decreased or no base-flow contribution to the Arkansas River except during times of recharge to bank storage after high flow events resulting from intense precipitation associated with thunderstorms. The second issue is due to documented loss of all or a majority of flow of the river to infiltration near the Dodge City, KS area. Inadequate or total lack of representation of the above two mechanisms in the model setup could have caused over-estimation of stream flows for the Arkansas River.

4.3. Summary of Challenges in Hydrologic Model Calibration for Large-Scale Studies

Major challenges are encountered with configuring SWAT for large systems, including limitations in input data availability and/or resolution, problems with configuring reservoirs and other

impoundments, which can greatly affect streamflow and pollutant transport, and limited measured data, that is necessary for testing simulated crop growth, hydrologic, and pollutant transport outputs. These limitations can prove to be particularly acute when calibrating and validating SWAT for huge systems, especially for lesser developed regions that typically have sparser temporal and spatial datasets such as the African conditions described by Schuol and Abbaspour [26] and Schuol et al. [27,28].

SWAT model has many parameters to calibrate which might complicate the calibration process for large-scale studies [9,65]. The version of SWAT used in the present study is a semi-distributed model, with no interaction between HRUs' limits the routing and nutrient transport within the sub-watersheds [66,67]. Additionally, assumptions such as constant channel roughness and hydraulic conductivity might affect the flow and pollutant transport processes in complex landscape [68–70].

The current structure of SWAT represents a fusion of both physically-based processes and empirical functions. Examples of two well-known empirical functions used in the model include the: (1) Runoff Curve Number (RCN) approach [71], which is used for estimating partitioning of precipitation between surface runoff and infiltration, and (2) Modified Universal Soil Loss Equation (MUSLE), which is described by Williams and Berndt [72] and is used to estimate sediment yield and delivery in SWAT. Beyond the use of such established empirical relationships, there are other ways in which processes are represented in a simplified manner in various components of the model such as the use of a generic crop/plant growth submodel and various transformations that are accounted for in the nutrient cycling component. In general, this approach has provided a robust platform for representing a wide range of cropping systems, land use, management, climate conditions and other factors that are required for evaluating water resource issues. However, the use of empirical and other simplified relationships in SWAT poses constraints, regarding how accurate the processes that are represented by those functions can be simulated.

At present, physically-based alternative approaches that could provide improved results with absolute certainty are not necessarily readily available. For example, the physically-based Green and Ampt (G&A) equation [73] is an alternative option provided in SWAT for estimating surface runoff and infiltration from precipitation inputs. However, results of a small subset of studies to date suggest that clear advantages of using the G&A approach are limited relative to the RCN method [74–76]. Meanwhile, research continues to introduce improved representation of processes in SWAT such as the recent study that describes the development of a module based on the Richards Equation [77]. Such efforts are consistent with the ongoing developmental history of SWAT, which has essentially continued unabated for almost three decades (e.g., [9,29,67]).

5. Conclusions

This study has identified multiple issues as some of the challenges to large-scale hydrologic model calibration. These include the availability of good quality data, accounting of multiple reservoirs within a sub-watershed, inadequate accounting of elevation and slopes in mountainous regions, representation of carrying capacity of channels (channel width, depth, and slope), omission of a part of the drainage area because of administrative boundary issues (e.g., between countries), inadequate replication of irrigation return flows, inadequate representation of vegetative cover and soils because of the coarse scale of the respective maps, poor representation of water abstractions (both surface and groundwater), and under-accounting of controlled releases of water between rivers.

This project was the first large scale study to utilize soft calibration techniques (Arnold et al. 2015). Average annual surface runoff, baseflow and total runoff were used as soft data for calibration and stream gage data (hard data) was analyzed without further calibration. An important finding of this paper is that good statistical agreement was found in hard data (annual gaged flows) when only calibrating the soft data (regionalized average annual surface runoff, baseflow, and total runoff). Additional calibration may be required when calibrating monthly and daily flows.

The lessons learned in this study are being utilized in a current national agroecosystem model that is replacing the HUMUS system [78]. The new modeling system used refined data including SSURGO soils data and 30 m DEM. The system also downscales processes to individual fields and first order streams which eliminates many of the current issues. In addition, individual ponds and reservoirs will be simulated, eliminating the problems currently caused by lumping impoundments within each 8-digit subbasin.

Author Contributions: N.K. conceptualized the calibration methodology, and designed it with the help of J.G.A., C.S., along with N.K. and M.J.W. developed the model input files. Validation of model results were carried out by J.G.A., along with C.S., N.K. and M.J.W. The original draft of the manuscript was prepared by N.K. Review of literature and editing were carried out by N.K., S.M. and P.W.G. Project administration and supervision were carried out by J.G.A. The software used for the model calibration was designed by N.K.

Funding: The authors thank the USDA for providing funding for this research as a part of Conservation Effects Assessment Project (CEAP).

Acknowledgments: The authors acknowledge the USGS, USDA, other state agencies and universities for providing access to the data required for this study.

Conflicts of Interest: The authors declare no conflict of interest.

References

- James, L.D.; Burges, S.J. Selection, calibration, and testing of hydrologic models. In *Hydrologic Modeling of Small Watersheds*; Haan, C.T., Johnson, H.P., Brakensiek, D.L., Eds.; ASAE Monograph: St. Joseph, MI, USA, 1982; pp. 437–472.
- Refsgaard, J.C. Parameterisation, calibration and validation of distributed hydrological models. *J. Hydrol.* **1997**, *198*, 69–97. [[CrossRef](#)]
- Daggupati, P.; Pai, N.; Ale, S.; Douglas-Mankin, K.R.; Zeckoski, R.W.; Jeong, J.; Parajuli, P.B.; Saraswat, D.; Youssef, M.A. A recommended calibration and validation strategy for hydrologic and water quality models. *Trans. ASABE* **2015**, *58*, 1705–1719.
- Arnold, J.G.; Srinivasan, R.; Muttiah, R.S.; Williams, J.R. Large area hydrologic modeling and assessment part I: Model development. *J. Am. Water Resour. Assoc.* **1998**, *34*, 73–89. [[CrossRef](#)]
- Arnold, J.G.; Fohrer, N. SWAT2000: Current Capabilities and Research Opportunities in Applied Watershed Modelling. *Hydrol. Processes* **2005**, *19*, 563–572. [[CrossRef](#)]
- Gassman, P.W.; Reyes, M.R.; Green, C.H.; Arnold, J.G. The Soil and Water Assessment Tool: Historical development, applications, and future research directions. *Trans. ASABE* **2007**, *50*, 1211–1250. [[CrossRef](#)]
- Williams, J.R.; Izaurralde, R.C.; Steglich, E.M. *Agricultural Policy/Environmental Extender Model: Theoretical Documentation*; Version 0604. BREC Report 2008-17; Texas A&M University, Texas AgriLIFE Research, Blackland Research and Extension Center: Temple, TX, USA, 2008.
- Abbaspour, K. SWAT-CUP 2012: SWAT Calibration and Uncertainty Programs-A User Manual. Eawag: Dübendorf, Switzerland. Available online: https://swat.tamu.edu/media/114860/usermanual_swatcup.pdf (accessed on 9 February 2019).
- Arnold, J.G.; Moriasi, D.N.; Gassman, P.W.; Abbaspour, K.C.; White, M.J.; Srinivasan, R.; Santhi, C.; Harmel, R.D.; Van Griensven, A.; Van Liew, M.W. SWAT: Model use, calibration, and validation. *Trans. ASABE* **2012**, *55*, 1491–1508. [[CrossRef](#)]
- Daggupati, P.; Yen, H.; White, M.J.; Srinivasan, R.; Arnold, J.G.; Keitzer, C.S.; Sowa, S.P. Impact of model development, calibration and validation decisions on hydrological simulations in West Lake Erie Basin. *Hydrol. Processes* **2015**, *29*, 5307–5320. [[CrossRef](#)]
- Wang, S.; Zhang, Z.; Sun, G.; Strauss, P.; Guo, J.; Tang, Y.; Yao, A. Multi-site calibration, validation, and sensitivity analysis of the MIKE SHE Model for a large watershed in northern China. *Hydrol. Earth System Sci.* **2012**, *16*, 4621–4632. [[CrossRef](#)]
- Cho, J.; Bosch, D.; Vellidis, G.; Lowrance, R.; Strickland, T. Multi-site evaluation of hydrology component of SWAT in the coastal plain of southwest Georgia. *Hydrol. Processes* **2013**, *27*, 1691–1700. [[CrossRef](#)]
- CARD. *SWAT Literature Database for Peer-Reviewed Journal Articles*; Center for Agricultural and Rural Development, Iowa State University: Ames, IA, USA, 2019. Available online: https://www.card.iastate.edu/swat_articles/ (accessed on 9 February 2019).

14. Gassman, P.W.; Sadeghi, A.M.; Srinivasan, R. Applications of the SWAT Model Special Section: Overview and insights. *J. Environ. Qual.* **2014**, *43*, 1–8. [[CrossRef](#)] [[PubMed](#)]
15. Douglas-Mankin, K.R.; Srinivasan, R.; Arnold, J.G. Soil and Water Assessment Tool (SWAT) model: Current developments and applications. *Trans. ASABE* **2010**, *53*, 1423–1431. [[CrossRef](#)]
16. Tuppad, P.; Douglas-Mankin, K.R.; Lee, T.; Srinivasan, R.; Arnold, J.G. Soil and Water Assessment Tool (SWAT) hydrologic/water quality model: Extended capability and wider adoption. *Trans. ASABE* **2011**, *54*, 1677–1684. [[CrossRef](#)]
17. Bressiani, D.d.A.; Gassman, P.W.; Fernandes, J.G.; Garbossa, L.H.P.; Srinivasan, R.; Bonumá, N.B.; Mendiondo, E.M. A review of Soil and Water Assessment Tool (SWAT) applications in Brazil: Challenges and prospects. *Int. J. Agric. Biol. Eng.* **2015**, *8*, 9–35. [[CrossRef](#)]
18. Krysanova, V.; White, M. Advances in water resources assessment with SWAT—An overview. *Hydrol. Sci. J.* **2015**, *60*, 771–783. [[CrossRef](#)]
19. Jha, M.K.; Gassman, P.W.; Secchi, S.; Arnold, J.G. Upper Mississippi River Basin modeling system part 2: Baseline simulation results. In *Coastal Hydrology and Processes*; Singh, V.P., Xu, Y.J., Eds.; Water Resources Publications: Highland Ranch, CO, USA, 2006; ISBN 1-887201-46-7.
20. Srinivasan, R.; Zhang, X.; Arnold, J. SWAT Ungauged: Hydrological Budget and Crop Yield Predictions in the Upper Mississippi River Basin. *Trans. ASABE* **2010**, *53*, 1533–1546. [[CrossRef](#)]
21. Mehta, V.M.; Rosenberg, N.J.; Mendoza, K. Simulated Impacts of Three Decadal Climate Variability Phenomena on Water Yields in the Missouri River Basin. *J. Am. Water Resour. Assoc.* **2011**, *47*, 126–135. [[CrossRef](#)]
22. Daggupati, P.; Deb, D.; Srinivasan, R.; Yeganantham, D.; Mehta, V.M.; Rosenberg, N.J. Large-scale fine-resolution hydrological modeling using parameter regionalization in the Missouri River basin. *J. Am. Water Resour. Assoc.* **2016**, *52*, 648–666. [[CrossRef](#)]
23. Rossi, C.G.; Srinivasan, R.; Jirayoot, K.; Le Duc, T.; Souvannabouth, P.; Binh, N.D.; Gassman, P.W. Hydrologic evaluation of the lower Mekong River basin with the Soil and Water Assessment Tool model. *Int. Agric. Engin. J.* **2009**, *18*, 1–13.
24. Arnold, J.G.; Allen, P.M. Automated methods for estimating baseflow and ground water recharge from streamflow records. *J. Am. Water Resour. Assoc.* **1999**, *35*, 411–424. [[CrossRef](#)]
25. Faramarzi, M.; Abbaspour, K.C.; Schulin, R.; Yang, H. Modelling blue and green water resources availability in Iran. *Hydrol. Processes* **2009**, *23*, 486–501. [[CrossRef](#)]
26. Schuol, J.; Abbaspour, K.C. Calibration and uncertainty issues of a hydrological model (SWAT) applied to West Africa. *Adv. Geosci.* **2006**, *9*, 137–143. [[CrossRef](#)]
27. Schuol, J.; Abbaspour, K.C.; Sarinivasan, R.; Yang, H. Estimation of freshwater availability in the West African Sub-continent using the SWAT hydrologic model. *J. Hydrol.* **2008**, *352*, 30–42. [[CrossRef](#)]
28. Schuol, J.; Abbaspour, K.C.; Yang, H.; Srinivasan, R.; Zehnder, A.J.B. Modelling blue and green water availability in Africa. *Water Resour. Res.* **2008**, *44*, 18. [[CrossRef](#)]
29. Williams, J.R.; Arnold, J.G.; Kiniry, J.R.; Gassman, P.W.; Green, C.H. History of model development at Temple, Texas. *Hydrol. Sci.* **2008**, *53*, 948–960. [[CrossRef](#)]
30. Gassman, P.W.; Arnold, J.G.; Srinivasan, R.; Reyes, M. The worldwide use of the SWAT model: Technological drivers, networking impacts, and simulation trends. In *Proceedings of the 21st Century Watershed Technology: Improving Water Quality and Environment*, Concepcion, Chile, 29 March–3 April 2008; ASABE Publication: St. Joseph, MI, USA, 2010.
31. USGS. *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*, 4th ed.; U.S. Department of the Interior: Washington, DC, USA; U.S. Geological Survey: Reston, VA, USA; U.S. Department of Agriculture: Washington, DC, USA; Natural Resources Conservation Service: Washington, DC, USA, 2013; p. 63. Available online: <http://pubs.usgs.gov/tm/11/a3/> (accessed on 9 February 2019).
32. Wang, X.; Kannan, N.; Santhi, C.; Potter, S.R.; Williams, J.R.; Arnold, J.G. Integrating APEX output for cultivated cropland with SWAT simulation for regional modeling. *Trans. ASABE* **2011**, *54*, 1281–1298. [[CrossRef](#)]

33. USDA–NRCS. *Assessment of The Effects of Conservation Practices on Cultivated Cropland in the Upper Mississippi River Basin*; USDA Natural Resources Conservation Service: Washington, DC, USA, 2012. Available online: https://www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/stelprdb1042093.pdf (accessed on 15 February 2019).
34. USDA–NRCS. *Assessment of the Effects of Conservation Practices on Cultivated Cropland in the Ohio–Tennessee River Basin*; USDA Natural Resources Conservation Service: Washington, DC, USA, 2011a. Available online: <https://www.nrcs.usda.gov/wps/portal/nrcs/detail/national/technical/nra/ceap/pub/?cid=stelprdb1046185> (accessed on 15 February 2019).
35. USDA–NRCS. *Assessment of the Effects of Conservation Practices on Conservation Effects Assessment Project (CEAP) Cultivated Cropland in the Great Lakes Region*; USDA Natural Resources Conservation Service: Washington, DC, USA, 2011b. Available online: <https://www.nrcs.usda.gov/Internet/FSE.../stelprdb1045481.pdf> (accessed on 15 February 2019).
36. USDA–NRCS. *Assessment of the Effects of Conservation Practices on Conservation Effects Assessment Project (CEAP) Cultivated Cropland in the Chesapeake Bay Region*; USDA Natural Resources Conservation Service: Washington, DC, USA, 2011c. Available online: https://www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/nrcsdev11_023934.pdf (accessed on 15 February 2019).
37. Santhi, C.; Arnold, J.G.; Williams, J.R.; Dugas, W.A.; Hauck, L. Validation of the SWAT model on a large river basin with point and nonpoint sources. *J. Am. Water Resour. Assoc.* **2001**, *37*, 1169–1188. [[CrossRef](#)]
38. Kannan, N.; Santhi, C.; Arnold, J.G. Development of an automated procedure for estimation of the spatial variation of runoff in large river basins. *J. Hydrol.* **2008**, *359*, 1–15. [[CrossRef](#)]
39. Srinivasan, R.; Arnold, J.G.; Jones, C.A. Hydrologic modeling of the United States with the Soil and Water Assessment Tool. *Int. Water Resour. Dev.* **1998**, *14*, 315–325. [[CrossRef](#)]
40. Neitsch, S.L.; Arnold, J.G.; Kiniry, J.R.; Williams, J.R.; King, K.W. *Soil and Water Assessment Tool Theoretical Documentation: Version 2000*; GSWRL Report 02-01, BRC Report 02-05, TR-191; Texas Water Resources Institute: College Station, TX, USA, 2002.
41. Blackland Research and Extension Center. Available online: <http://www.brc.tamus.edu/swat/index.html> (accessed on 31 January 2019).
42. Santhi, C.; Kannan, N.; Di Luzio, M.; Potter, S.R.; Arnold, J.G.; Atwood, J.D.; Kellog, R.L. An approach for estimating water quality benefits of conservation practices at the national level. In Proceedings of the American Society of Agricultural and Biological Engineers (ASABE), Annual International Meeting, Tampa, FL, USA, 17–20 July 2005; p. 052043. [[CrossRef](#)]
43. Vogelmann, J.E.; Howard, S.M.; Yang, L.; Larson, C.R.; Wylie, B.K.; Van Driel, J.N. Completion of the 1990’s National Land Cover Dataset for the conterminous United States. *Photogramm. Eng. Remote Sens.* **2001**, *67*, 650–662.
44. USDA–NRCS. *State Soil Geographic (STATSGO) Database*; USDA Natural Resources Conservation Service: Washington, DC, USA, 1994. Available online: https://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/survey/geo/?cid=nrcs142p2_053629 (accessed on 9 February 2019).
45. Eischeid, J.K.; Pasteris, P.A.; Diaz, H.F.; Plantico, M.S.; Lott, N.J. Creating a serially complete, national daily time series of temperature and precipitation for the Western United States. *J. Appl. Meteorol.* **2000**, *39*, 1580–1591. [[CrossRef](#)]
46. Daly, C.R.; Neilson, P.; Phillips, D.L. A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *J. Appl. Meteorol.* **1994**, *33*, 140–158. [[CrossRef](#)]
47. Daly, C.; Gibson, W.P.; Taylor, G.H.; Johnson, G.L.; Pasteris, P. A knowledge-based approach to the statistical mapping of climate. *Climate Res.* **2002**, *22*, 99–113. [[CrossRef](#)]
48. Di Luzio, M.; Johnson, G.L.; Daly, C.; Eischeid, J.; Arnold, J.G. Constructing Retrospective gridded daily precipitation and temperature datasets for the conterminous United States. *J. Appl. Meteorol. Climatol.* **2008**, *47*, 475–497. [[CrossRef](#)]
49. Nicks, A.D. Stochastic generation of the occurrence, pattern and location of maximum amount of rainfall. *Misc. Publ.* **1974**, *1275*, 154–171.
50. Sharpley, A.N.; Williams, J.R. EPIC Erosion/Productivity Impact Calculator: 1. Model Documentation. *USDA Tech. Bull.* **1990**, *1759*, 235.
51. CorpsMap, National Inventory of Dams. Available online: <http://nid.usace.army.mil> (accessed on 31 January 2019).

52. Neitsch, S.L.; Arnold, J.G.; Kiniry, J.R.; Williams, J.R. *Soil and Water Assessment Tool Theoretical Documentation Version 2009*; TR-406; Texas Water Resources Institute: College Station, TX, USA, 2011.
53. Gebert, W.A.; Graczyk, D.J.; Krug, W.R. Annual Average Runoff in the United States, 1987, 1951–1980: US Geological Survey Hydrologic Investigations Atlas HA-710, 1 sheet, scale 1:7,500,000. Available online: <https://doi.org/10.3133/ha710> (accessed on 28 December 2018).
54. Arnold, J.G.; Allen, P.M.; Muttiah, R.S.; Bernhardt, G. Automated base flow separation and recession analysis techniques. *Groundwater* **1995**, *33*, 1010–1018. [[CrossRef](#)]
55. Hargreaves, G.H.; Samani, Z.A. Reference crop evapotranspiration from temperature. *Appl. Eng. Agric.* **1985**, *1*, 96–99. [[CrossRef](#)]
56. Hargreaves, G.; Allen, R. History and Evaluation of Hargreaves Evapotranspiration Equation. *J. Irrig. Drain. Eng.* **2003**, *129*, 53–63. [[CrossRef](#)]
57. Santhi, C.; Kannan, N.; Arnold, J.G.; Luzio, M.D. Spatial calibration and temporal validation of flow for regional scale hydrologic modeling. *J. Am. Water Resour. Assoc.* **2008**, *44*, 829–846. [[CrossRef](#)]
58. Arnold, J.G.; Youssef, M.A.; Yen, H.; White, M.J.; Sheshukov, A.Y.; Sadeghi, A.M.; Moriasi, D.N.; Steiner, J.L.; Amatya, D.M.; Skaggs, R.W.; et al. Hydrological processes and model representation: Impact of soft data on calibration. *Trans. ASABE* **2015**, *58*, 1637–1660. [[CrossRef](#)]
59. Nash, J.E.; Sutcliffe, J.V. River flow forecasting through conceptual models Part 1-Discussion of principles. *J. Hydrol.* **1970**, *10*, 282–290. [[CrossRef](#)]
60. Moriasi, D.N.; Arnold, J.G.; Van Liew, M.W.; Bingner, R.L.; Harmel, R.D.; Veith, T. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASABE* **2007**, *50*, 885–900. [[CrossRef](#)]
61. Moriasi, D.N.; Gitau, M.W.; Pai, N.; Daggupati, P. Hydrologic and water quality models: Performance measures and evaluation criteria. *Trans. ASABE* **2015**, *58*, 1763–1785. [[CrossRef](#)]
62. Musser, K.; Kudelka, S.; Moore, R. Minnesota River Basin Trends. Mankato State University: Mankato, MN, USA, 2009; p. 64. Available online: https://mrbdc.mnsu.edu/sites/mrbdc.mnsu.edu/files/public/mnbasin/trends/pdfs/trends_full.pdf (accessed on 9 February 2019).
63. Controlling the Mississippi River for National Commerce. Available online: http://www.americaswetlandresources.com/background_facts/detailedstory/LouisianaRiverControl.html (accessed on 31 January 2019).
64. United States Geological Survey, Boundary Descriptions and Names of Regions, Subregions, Accounting Units and Cataloging Units. Available online: http://water.usgs.gov/GIS/huc_name.html (accessed on 31 January 2019).
65. Mehan, S.; Kannan, N.; Neupane, R.P.; McDaniel, R.; Kumar, S. Climate change impacts on the hydrological processes of a small agricultural watershed. *Climate* **2016**, *4*, 56. [[CrossRef](#)]
66. Arnold, J.G.; Allen, P.M.; Volk, M.; Williams, J.R.; Bosch, D.D. Assessment of different representations of spatial variability on SWAT model performance. *Trans. ASABE* **2010**, *53*, 1433–1443. [[CrossRef](#)]
67. Bieger, K.; Arnold, J.G.; Rathjens, H.; White, M.J.; Bosch, D.D.; Allen, P.M.; Volk, M.; Srinivasan, R. Introduction to SWAT+, a completely restructured version of the Soil and Water Assessment Tool. *J. Am. Water Resour. Assoc.* **2017**, *53*, 115–130. [[CrossRef](#)]
68. Sophocleous, M.A.; Koelliker, J.K.; Govindaraju, R.S.; Birdie, T.; Ramireddygar, S.R.; Perkins, S.P. Integrated numerical modeling for basin-wide water management: The case of the Rattlesnake Creek basin in south-central Kansas. *J. Hydrol.* **1999**, *214*, 179–196. [[CrossRef](#)]
69. Sophocleous, M.; Perkins, S.P. Methodology and application of combined watershed and groundwater models in Kansas. *Hydrology* **2000**, *236*, 185–201. [[CrossRef](#)]
70. FitzHugh, T.W.; Mackay, D.S. Impacts of input parameter spatial aggregation on an agricultural nonpoint source pollution model. *J. Hydrol.* **2000**, *236*, 35–53. [[CrossRef](#)]
71. USDA-FSA. *Conservation Reserve Program Overview. CRP: Planting for the Future*; U.S. Department of Agriculture, Farm Service Agency: Washington, DC, USA, 2004. Available online: <https://www.fsa.usda.gov/programs-and-services/conservation-programs/> (accessed on 9 February 2019).
72. Williams, J.R.; Berndt, H.D. Sediment yield prediction based on watershed hydrology. *Trans. ASAE* **1977**, *20*, 1100–1104. [[CrossRef](#)]
73. Green, W.H.; Ampt, G.A. Studies on soil physics, 1. The flow of air and water through soils. *J. Agric. Sci.* **1911**, *4*, 11–24.

74. King, K.W.; Arnold, J.G.; Bingner, R.L. Comparison of Green-Ampt and curve number methods on Goodwin creek watershed using SWAT. *Trans. ASAE* **1999**, *42*, 919–926. [[CrossRef](#)]
75. Kannan, N.; White, S.M.; Worrall, F.; Whelan, M.J. Sensitivity analysis and identification of the best evapotranspiration and runoff options for hydrological modeling in SWAT-2000. *J. Hydrol.* **2007**, *332*, 456–466. [[CrossRef](#)]
76. Ficklin, D.L.; Zhang, M. A comparison of the curve number and green-ampt models in an agricultural watershed. *Trans. ASABE* **2013**, *56*, 61–69. [[CrossRef](#)]
77. Qi, C.; Grunwald, S. GIS-based hydrologic modeling in the Sandusky watershed using SWAT. *Trans. ASAE* **2005**, *48*, 169–180. [[CrossRef](#)]
78. White, M.J.; Bieger, K.; Gambone, M.; Haney, L.; Arnold, J.G.; Gao, J. Development of a hydrologic connectivity dataset for SWAT assessments in the U.S. *Water* **2017**, *9*, 892. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).