Article

Development of an Evapotranspiration Data Assimilation Technique for Streamflow Estimates: A Case Study in a Semi-Arid Region

Ying Zhang [1,2] , Ling Zhang [1,2], Jinliang Hou [1], Juan Gu [3] and Chunlin Huang [1,*]

[1] Key Laboratory of Remote Sensing of Gansu Province, Heihe Remote Sensing Experimental Research Station, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou 730000, China; hexiayouxi@126.com (Y.Z.); zhanglingky@lzb.ac.cn (L.Z.); jlhours@lzb.ac.cn (J.H.)

[2] University of Chinese Academy of Sciences, Beijing 100000, China

[3] College of Earth Environmental Sciences, Lanzhou University, Lanzhou 730000, China; gujuan@lzu.edu.cn

[*] Correspondence: huangcl@lzb.ac.cn; Tel.: +86-931-496-7934

Abstract: Streamflow estimates are substantially important as fresh water shortages increase in arid and semi-arid regions where evapotranspiration (ET) is a significant contribution to the water balance. In this regard, evapotranspiration data can be assimilated into a distributed hydrological model (SWAT, Soil and Water Assessment Tool) for improving streamflow estimates. The SWAT model has been widely used for streamflow estimations, but the applications combining SWAT and ET products were rare. Thus, this study aims to develop a SWAT-based evapotranspiration data assimilation system. In particular, SWAT is gridded at Hydrologic Response Unit (HRU) level to incorporate gridded ET products acquired from the remote sensing-based ETMonitor model. In the modeling case, Gridded SWAT (GSWAT) shows a good agreement of streamflow modeling with the original SWAT. Such a scant margin between them is due to the modeling domain mismatch caused by different HRU delineations. In the ET assimilation case, we carry out a synthetic data experiment to illustrate the state augmentation Direct Insertion (DI) method and a real data experiment for the upper Heihe River Basin. The results demonstrate the benefits of the ET assimilation for improving hydrologic processes representations. In the future, more remotely sensed data can be assimilated into the data assimilation system to provide more reliable hydrological predictions.

Keywords: evapotranspiration; streamflow; data assimilation; SWAT; semi-arid

1. Introduction

Highly accurate forecasting of streamflow is critical for water resource management as fresh water shortages increase in arid and semi-arid regions [1]. As an important component of the water budget, it is also critical for sustainable crop production, food security and social sustainability [2–4]. With the data assimilation method and remotely sensed data, streamflow estimates can be improved a lot [5–9]. However, most of these studies used remotely sensed soil moisture or snow; little attention has been focused on assimilating observed evapotranspiration (ET) data. In particular, evapotranspiration is the largest mode of water loss in arid and semi-arid regions [2]. Streamflow estimates will benefit a lot from taking into account observed evapotranspiration from remote sensing in these areas.

Remote sensing-derived ET products have been used to evaluate or calibrate the hydrological model and improve streamflow forecasting [10–13]. However, such calibrations need a lot of time and labor or computing resources. Additionally, calibration is insufficient for streamflow forecasting when forcing input (i.e., precipitation) is not correct. Data assimilation as a complementary tool can be used to fill up the calibration deficiency in decreasing the discrepancy between streamflow

prediction and observation. Pan et al. [14] assimilated Moderate Resolution Imaging Spectroradiometer (MODIS)-derived ET into Variable Infiltration Capacity (VIC) and Surface Energy Balance System (SEBS) models to estimate water budget terms. However, ET was only a non-state variable in the model, which could not feed the assimilation effect back to the model and update the model's state variables, so the other hydrological variables were not optimized as a whole and this method is not suitable for improving streamflow prediction.

ET-related data assimilation researches are still in the formative stage for most hydrological models [15,16]. For model state updating purposes, ET observation should be associated with the model state variable; considering ET as a state variable in a model (e.g., [15]) or associating it with other state variables are two practical methods. Remotely sensed evapotranspiration can be used for hydrological model parameter calibration or estimation, which can affect the surface runoff generation [10,12,13,16,17]. Parr et al. [18] found that hydrological predictions may be subject to the bias in past ET modeling, leading to substantial uncertainties in the modeled hydrological trend. They replaced the simulated ET components (canopy evaporation, transpiration, bare ground evaporation, canopy sublimation, and surface sublimation) of the VIC model with the bias-corrected ET calculated from a simple bias-correction algorithm and historic remotely sensed ET. The streamflow estimation was improved due to the corrected ET components. This procedure can still be considered as a model bias correction method. Understanding historic evapotranspiration trends and high-quality forcing data are prerequisites, because the mass balance is the main constrain for ET components and runoff updating in VIC [18]. Generally, monthly time scale water balance, including ET and streamflow, can be simulated well by well calibrated hydrological models. They are particularly valuable for applications where one is primarily interested in monthly, seasonal and annual streamflow volumes [19]. While the flood risk warning and management need hydrological data at high spatiotemporal resolution, utilizing fresh incoming high-resolution ET data for hydrological forecasting simultaneously is still a big challenge.

It is a major challenge to relate surface runoff modeling to the ET modeling. The reports and literatures on the relationship among precipitation, ET and surface runoff will be useful for associating daily ET and surface runoff [20]. Rawlins et al. [21] found that simulations of arctic evapotranspiration and runoff are both strongly dependent on the quality of time series data used to drive the model. The uncertainty of model input (i.e., precipitation) can be the main factor to influence the simulation accuracy of both ET and surface runoff [22]. The uncertainty of precipitation can be defined by a widely used multiplicative error model [23,24]. An estimated precipitation error can be obtained by analyzing observed and simulated ET. The error for simulated ET can be transformed to precipitation error based on the ET model. So using remotely sensed evapotranspiration to reduce input uncertainties and further to improve the hydrologic predictions is practical.

In this paper, we assess the performance of streamflow predictions with remote sensing-derived ET and state augmentation direct insertion method for a distributed hydrological model, SWAT. We have to solve two problems: (1) matching ET data with the modeling units efficiently; and (2) assimilating the ET into the model to update the streamflow estimates. In this way, a modified SWAT model and a state augmentation direct insertion assimilation method are designed to assimilate ET data.

The remainder of the paper is organized as follows. Sections 2 and 3 provide background on hydrological models. Section 4 describes the direct insertion data assimilation method. We describe case studies including the design of the synthetic and real-word data assimilation experiments in Section 5. Section 6 presents the results from two designed cases: (1) hydrological modeling scheme; and (2) evapotranspiration data assimilation scheme. Finally, conclusions and outlooks are summarized in Section 7.

2. Soil and Water Assessment Tool (SWAT)

The SWAT model is a semi-distributed hydrological model that includes detailed representations of hydrological processes at the watershed scale. The model calculates vertical and horizontal water

movement through three spatial levels (HRU level, subbasin level and basin level). There are two hydrologic cycle modeling phases (land phase and routing phase) in SWAT. In Land phase, snow melt is calculated based on a linear function of the difference between the average snow pack maximum air temperature and the base or threshold temperature for snow melt. Daily surface runoff is calculated based on Soil Conservation Service curve number (SCS-CN) method. Vegetation growth and canopy intercept is calculated based on plant growth function and LAI. Potential and actual evapotranspiration (evaporation from ground and vegetation, transpiration from vegetation) can be calculated based on Penman-Monteith method (used in this study), Priestley-Taylor method or Hargreaves method. Water percolates through several soil layers and ultimately becomes aquifer recharge. Some of it moves laterally in the soil profile and contribute to streamflow. Water leaves groundwater storage primarily by discharge into rivers or lakes. In the routing phase, water routing from reaches and channels to basin outlet is calculated using Muskingum or the variable storage coefficient routing method. See Arnold, et al. [25] for a detailed description of the SWAT model. SWAT has been widely used for regional water budget estimations around the word for decades [26,27].

3. Gridded SWAT (GSWAT)

The modeling units in SWAT are Hydrologic Response Unit (HRU), subbasin and basin from fine to large. HRU is the combination of unique soil, land cover and topography characters. HRUs may have various sizes and shapes and no geospatial locations. The HRU attributes will cost users a lot of work to process the grid format remote sensing data to match their spatial location and irregular size. We need to match HRU with grid data efficiently. More user-friendly modeling units will save SWAT users a lot of time and labor for incorporating grid format data. We gridded SWAT through two main procedures: data gridded and model gridded. The comparison between SWAT and GSWAT is shown in Table 1. The software of GSWAT is available at http://hexiayouxi.tk.

3.1. Data Gridded

We gridded HRU in SWAT, which can be called GSWAT. Each grid cell is an individual HRU in GSWAT. Forcings and parameters for HRUs are the same for the corresponding grid cells. An HRU parameters database is generated from the SWAT database (e.g., SWAT.mdb). HRU identification number raster (hruid.txt) and subbasin identification raster (subid.txt) are generated based on the HRU and subbasin maps. Only the area and shape of HRU are changed to match with grid format data. The model will link the HRU index with the corresponding HRU parameters when reading input data. Therefore, there are no original HRU level input files (e.g., *.hru). The number of HRUs is determined by the number of cells within the watershed. The grid spatial dimension and resolution of the watershed is defined in the control file (file.cio). Forcing data are derived from the weather stations from SWAT. More details of the modification of SWAT can be seen in [28,29]. We can match gridded ET data with gridded HRUs spatially. Spatial extent and location between gridded ET and gridded HRU should be consistent.

3.2. Model Gridded

Vertical and horizontal water balance equations in SWAT are retained (Equation (1)) for GSWAT. There are two source soil water balance equation which separates evaporation and transpiration in surface and root zone [30]. The hydrologic process as simulated by SWAT and GSWAT at each HRU is based on the single source water balance equation [31]:

$$SW_t = SW_o + \sum_{j=1}^{t} \left(P_j - SURQ_j - ET_{a,j} - S_j - G_j \right) \tag{1}$$

where SW_t is the soil water content above the wilting point in soil profile at the end of the day t, SW_o is the initial soil water content. P_j is the amount of the precipitation including rainfall and snowfall on

day j, $SURQ_j$ is the amount of surface runoff on day j, $ET_{a,j}$ is the amount of actual evapotranspiration on day j, S_j is the amount of water entering the vadose zone from the soil profile and G_j is the amount of return flow on day j. All units are mm H_2O. Soil water content can be expressed as volume fraction based on the thickness of each layer.

The conventional SWAT model code is modified to incorporate the gridded forcing data and parameter data. The HRU related parameter variables are redefined to include the grid location dimension, for example, soil water available variable (sol_awc), whose HRU number dimension size is modified to ncolums × nrows. Ncolums and nrows are grid cells' column and row respectively. HRU id (hruid) variable is added to imply variable's spatial location. At this point, the original jth HRU's sol_awc(:,j) variable is modified to sol_awc(:,hruid(j)). Grid data reading/writing subroutines (e.g., readsub1.f, readhru1.f, pmeas2.f) are added into the code to read/write ASCII/NETCDF format data (e.g., hruid.txt).

Table 1. Comparison between SWAT and GSWAT.

	Name	SWAT	GSWAT
	Forcing input	Weather station data	Gridded data
	Parameters		
	Basin level	basin.bsn	basin.bsn
Data	Subbasin level	*.sub, *.rte	*.sub, *.rte
	HRU level	*.hru, *.mgt, *.swq, *.gw	Grid files (hruid.txt, subid.txt) and database file (hru.txt)
	Output	output.*	Daily and yearly variables grid files (*.asc)
	Unit	HRU	Grid cell
Model	Domain	DEM-based watershed delineation	Grid scheme-based watershed delineation
	Water balance equation	Equation (1)	Equation (1)

*: subbasin or HRU name.

4. Implementation of Data Assimilation Technique

4.1. State Augmentation Direct Insertion Data Assimilation

According to the water balance equation (Equation (1)), the water balance components are correlated with each other. During both ET and surface runoff simulations, the precipitation uncertainty plays a leading role. When rainfall occurs, water is partitioned into three main components: surface runoff, evapotranspiration and infiltration. Surface runoff is calculated based on SCS-CN conceptual method. The physically surface runoff hydrological process is not represented well [32].

Based on the data assimilation concepts [33], the observed and modeled hydrological variables relationships calculation equations can be written as follows:

$$\varepsilon_p = \lambda \cdot P_o \tag{2}$$

$$ET = ET_o - \delta_e = ET_m - \varepsilon_e \tag{3}$$

$$SURQ = SURQ_o - \delta_q = SURQ_m - \varepsilon_q \tag{4}$$

where P_o is model input (observed) precipitation, ε_p is the precipitation error, λ is the multiplicative error coefficient, ET_o is the observed ET, ET_m is the modeled ET. ε_e is the model error for ET process, δ_e is the ET measurement error, $SURQ_o$ is updated surface runoff (corresponds to observed surface

runoff), $SURQ_m$ is modeled surface runoff. ε_q is the model error in surface runoff process, δ_q is the surface runoff measurement error.

Due to the precipitation having multiplicative error most possibly [23], we define the precipitation error as a multiplicative error (see Equation (2)). The multiplicative error model extracts the systematic errors more cleanly, which is more consistent with the large variability of precipitation measurements. When the observed data has low bias and deserves to be trust ($\delta_e = 0$, $\delta_q = 0$), we can assume that the discrepancy between observed data and simulated data is due to the model uncertainty (see Equations (3) and (4)). When the model has been calibrated and validated well against long term observations, model system errors could be neglected, therefore we can assume that the model uncertainties for evapotranspiration and surface runoff are dominated by the input (precipitation) uncertainty. Additionally, precipitation, evapotranspiration and surface runoff are highly correlated with each other [34]. The error coefficient for simulated ET and surface runoff can be linearly correlated. We can also assume the model error is partially represented by simulated ET and surface runoff and precipitation error:

$$\varepsilon_e = \lambda_{et} \cdot ET_m = \alpha_{et} \cdot \varepsilon_p \tag{5}$$

$$\varepsilon_q = \lambda_{surq} \cdot SURQ_m = \alpha_q \cdot \varepsilon_p \tag{6}$$

where λ_{et} and λ_{surq} is the multiplicative error coefficient for simulated ET and surface runoff, α_{et} and α_q is the multiplicative error coefficient for precipitation error, representing the precipitation error contribution to the simulated ET and surface runoff. Here, we simply assume the errors sourced from precipitation error contribution equally ($\alpha_{et} = \alpha_q$). Along with observed ET, surface runoff can be estimated with the error coefficient calculated from simulated ET.

We can redeploy the water balance components based on the proportional discrepancy between observed and simulated water balance components. With observed and modeled ET data, we can define the uncertainty and calculate the correct coefficient. We already assumed that the proportional discrepancy between observed ET and simulated ET is also suitable for surface runoff variable. Then we multiply modeled surface runoff variable with the correct coefficient to acquire unbiased surface runoff. The basin states will be updated along with the hydrologic processes subsequently. This data assimilation method can be termed as state augmentation direct insertion (DI) method.

4.2. Implementation Procedures

The flowchart of the implementation procedures of GSWAT coupled with DI is shown in Figure 1. As shown in Figure 1, in each time step gridded HRU simulation, when GSWAT model detects available ET data, it will replace modeled ET value with incoming ET data at the end of the ET calculation. PET and surface runoff will be updated according to the correct coefficient derived from observed ET (ET_o) and modeled ET ET_m (see Equations (3) and (4)). Finally, the updated HRU runoff will be accumulated to the subbasin level and enters routing phase. Streamflow will be simultaneously updated by the updated water storage in each reach.

5. Case Study

5.1. Study Area and Data

The Heihe River Basin is the second largest inland river basin in China (Figure 2). The upper Heihe River Basin is around 10,000 km^2 and located above 1700 m elevation. This watershed is located in the semi-arid area. The upper Heihe River provides 80.2% of the total river flow to the middle and lower reach. Water from glacial melt only accounted for 3.6% of the river flow. Precipitation is the main water sources in this region. The annual precipitation varies in the range 200–500 mm and 60% of the total precipitation occurred in the summer. The annual potential evapotranspiration is over 600 mm. Runoff from the upper Heihe River maintains industrial and agricultural production in the midstream and ecological system heath of the downstream [35]. There are two major tributaries originate from

the Qilian Mountain in east and the west. Yingluoxia hydrological station is the outlet of the upstream watershed. The dominant land cover types are mountain forest and alpine meadow.
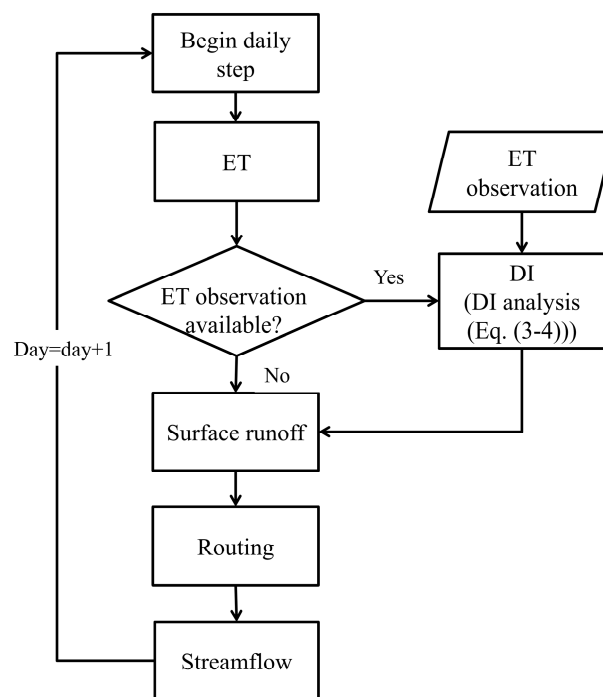
Figure 1. The flowchart for state augmentation direct insertion data assimilation.

All input data including forcing data, parameter data for SWAT and GSWAT are all acquired from WestDC (http://westdc.westgis.ac.cn). They are resampled to 1 km for both models. Streamflow observations obtained from the Yingluoxia station are used to calibrate SWAT model parameters (e.g., CN2). Details of calibration can be seen in Zhang et al. [36]. The meteorological data (air temperature, precipitation, relative humidity, wind speed, and solar radiation) and calibrated parameters (e.g., soil texture and conductivity) are both used for SWAT and GSWAT. The SWAT model grid size is 30m based on the DEM resolution. The grid resolution for GSWAT is 1 km, which will be consistent with the following evapotranspiration data.

Evapotranspiration data was produced based on remote sensing-based ETMonitor model [37] and WRF (Weather Research and Forecasting) model data by Cui and Jia [38]. With a variety of parameterization methods of the aerodynamic and bulk surface resistances, the estimation of soil evaporation and vegetation transpiration was based on the Shuttleworth-Wallace Dual-Source model. The model contains a variant form of Penman-Monteith equation that combines energy balance with plant physiology and soil water diffusion based on the assumption of aerodynamic mixing occurring at a mean canopy source height within the canopy. The rainfall interception was estimated using a reformulated remote sensing-based Gash model, which is forced by remote sensing observations of canopy structure (e.g., LAI and vegetation storage capacity) and rainfall intensity. The interception of vegetation was used to replace the interception of canopy, and a Poisson distribution function was applied to the LAI value of each pixel. A recommended method proposed by the World Meteorological Organization (WMO) was used to estimate the sublimation of snow and ice using the formula developed by Kuzmin [39]. The evaporation from water surface was estimated using the classical Penman equation [40].
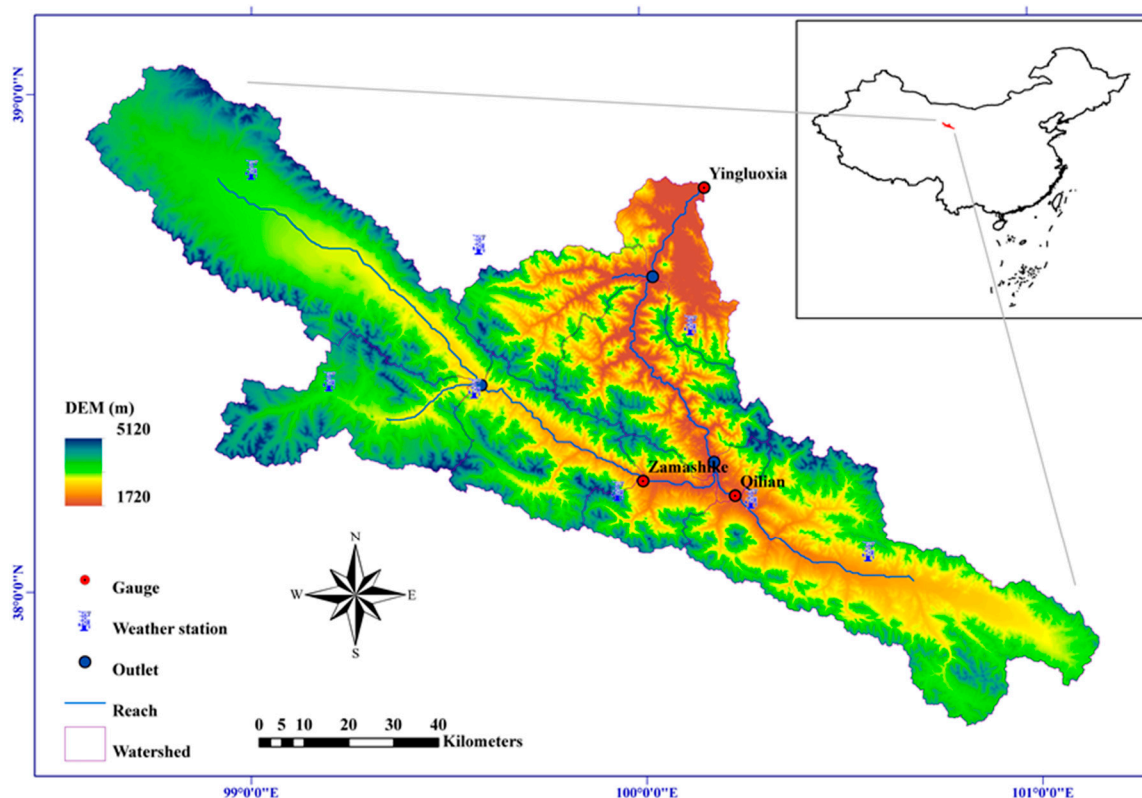
Figure 2. Location of the upper Heihe River Basin in Northwest China.

The ETMonitor model combines three different actual ET parameterizations for water body, snow/ice surface and soil-vegetation canopy. It utilizes microwave and optical remote sensing observations including MODIS albedo product (MCD43), MODIS vegetation index product (MOD13A2), soil moisture (AMSR-E), MODIS land cover type product (MCD12Q1), snow cover (MOD10A2) and precipitation (Tropical Rainfall Measuring Mission (TRMM)) as input data to estimate the daily ET_p for all sky conditions (see Equation (7)). The saturated and residual soil water contents of soil surface layer for the ETMonitor were provided by the China dataset of soil hydraulic parameters [41], which are derived from multiple Pedotranfer Functions, including the parameters in the Clapp and Hornberger Functions (FCH) and in the van Genuchten and Mualem Functions (FGM), respectively.

$$ET_P = M \begin{Bmatrix} TRMM \\ WRF\ data \\ AMSR\text{-}E \\ MCD12Q1 \\ MOD10A2 \\ MCD43 \\ MOD13A2 \end{Bmatrix} \tag{7}$$

where ET_p is the evapotranspiration product. M is ETMonitor model.

The daily actual evapotranspiration of the Heihe River Basin for the years 2009–2011 was estimated with the ETMonitor at a spatial resolution of 1 km. This product was evaluated using eddy covariance (EC) flux observations (The data were collected from the Watershed Allied Telemetry Experimental Research (WATER) and provided by the Cold and Arid Regions Sciences Data Center at Lanzhou (available at http://westdc.westgis.ac.cn/) at field scale and compared with other ET product (MOD16A2) at regional scale with smaller RMSE (low to 0.39 mm/day) and higher R^2 (up to

0.96). This ET product presents spatial and temporal consistent with other product. Merging such lots of remote sensing data and models provides us a reliable ET observation.

5.2. Hydrological Modeling Scheme

To evaluate the performance of hydrological modeling, we carried out Yingluoxia watershed experiment with SWAT and GSWAT. SWAT was used to simulate daily streamflow within the upper Heihe River Basin between 1990 and 2008. The previous SWAT calibration and validation work is based on 1990–1999 and 2000–2009 Yingluoxia hydrological station observation data [36]. Twenty-two sensitive parameters were selected to be calibrated against the streamflow observations based on the previous studies. The basin and subbasin level parameters for SWAT were automatically calibrated using Particle Swarm Optimization method with 100 iterations. We used the *hydroPSO* R package to conduct the automatic model calibration which is proposed by Zambrano-Bigiarini and Rojas [42], and HRU level parameters were manually calibrated (see Table 2).

Table 2. Description and range of the calibrated model parameters.

Parameter Name	Description	Range	
		Minimum	Maximum
ALPHA_BF	Base flow alpha factor (days)	0.01	0.99
CH_K2	Effective hydraulic conductivity in main channel alluvium (mm/h)	0	200
CN2	Initial SCS CN II value	10	95
SOL_AWC	Available water capacity (mm H_2O/mm soil)	0.01	0.035
SOL_K	Saturated hydraulic conductivity (mm/h)	0.001	1000
ESCO	Soil evaporation compensation factor	0.01	1
CH_N2	Manning's "n" value for the main channel	0.016	0.15
SFTMP	Snowfall temperature (°C)	−5	5
SURLAG	Surface runoff lag time (days)	1	12
SMTMP	Snowmelt base temperature (°C)	−5	5
SMFMN	Minimum melt factor for snow (°C)	1.4	6.9
TIMP	Snowpack temperature lag factor	0.01	0.6
SMFMX	Maximum melt factor for snow (°C)	1.4	4
OV_N	Manning's "n" value for overland flow	0.008	0.6
RCHRG_DP	Deep aquifer percolation factor	0	1
GWQMN	Threshold water depth in the shallow aquifer for flow (mm)	0	5000
GW_REVAP	Groundwater "revap" coefficient	0	0.2
GW_DELAY	Groundwater delay time (days)	1	100
SOL_ALB	Moist soil albedo	0	0.1
REVAVMN	Threshold water depth in the shallow aquifer for "revap" (mm)	1	500
EPCO	Plant uptake compensation factor	0.01	1
CANMX	Maximum canopy storage (mm H2O)	0	10

To exclude the influence of parameters source, we used the already calibrated parameters from SWAT for GSWAT. During the SWAT calibration period, NSE (Nash Sutcliffe Efficiency) and R^2 were 0.86 and 0.92 respectively for the simulated monthly streamflow. During the SWAT validation period, NSE and R^2 were 0.88 and 0.86 respectively. Then, the GSWAT was driven by the same forcings derived from SWAT to simulate upper Heihe River Basin variables from 1990 to 2008.

5.3. ET Data Assimilation Scheme

5.3.1. Synthetic Experiment Design

Our primary goal of synthetic data experiment is to test the data assimilation method. We designed a synthetic "twin" experiment and specifications for different cases are shown in Table 3. In the TRUE case, we first ran GSWAT to simulate upper Heihe river basin throughout the year 2000 and considered the results as true states. We conducted many years (1991–2009) simulations and just selected the year 2000 as a representation. As a synthetic experiment, we want to set up an experiment to include various types of errors. Based on the model error (i.e., precipitation error), we found that the daily precipitation and streamflow in the year 2000 are very different from those in the year 2009. So we selected the year 2000 as our synthetic experiment simulation period to test whether the ET data

assimilation method can really diminish model errors under different forcing patterns in the real data experiment. The ET result was selected as ET observation. The precipitation was selected as true precipitation. For data assimilation and open loop simulation, the precipitation uncertainty was defined as a multiplicative log normally distributed error (defined in Equations (8)–(11)) [43]; In the DA and NODA cases, GSWAT was driven by the perturbed precipitation ($P_{perturbed}$) to simulate basin states and fluxes with and without ET assimilated throughout year 2000. Finally, we analyzed streamflow results derived from this experiment.

$$e \sim N(1, 0.25) \tag{8}$$

$$\mu = ln\left(\frac{P^2}{\sqrt{P^2 + (s \times P)^2}}\right) \tag{9}$$

$$\sigma = \sqrt{ln\left(\frac{(s \cdot P)^2}{P^2} + 1\right)} \tag{10}$$

$$P_{perturbed} = exp(\mu + e \times \sigma) \tag{11}$$

where e is normally distributed random variable with 1 mean and standard deviation of 0.25. P is precipitation, and s is the variance scaling factor of precipitation with a set value of 0.5 in this study.

Table 3. Test plan for synthetic "twin" experiment.

CASE	Model	Forcing	Measurement
TRUE	GSWAT	Weather station data (P)	
NODA	GSWAT	$P_{perturbed}$ (Equation (11))	
DA	GSWAT	$P_{perturbed}$	ET

5.3.2. Real Data Experiment Design

The real data experiment design is similar to the synthetic experiment except the observation data. The observation data in the real data experiment is the ET product (see Section 5.1). The parameters derived from SWAT may be not appropriate for GSWAT, the streamflow forecasting can be further improved by an additional calibration. So we added extra calibration procedures for GSWAT modeling and ET data assimilation cases. More details of the test plan for real data experiment are shown in Table 4.

Table 4. Test plan for real data experiment.

CASE	Model		Method	
	SWAT	GSWAT	Calibration	ET assimilation
SWAT	√		√	
GSWAT		√		
GSWATET		√		√
GSWATCALI		√	√	
GSWATCALIET		√	√	√

√: used in the case.

6. Results and Discussion

As noted above, our work is centered on assimilating remote sensing-based ET data into the modified SWAT model. In particular, we first assess the modified model by comparing it with the original model; then we assess the performance of ET data assimilation with the modified SWAT model through synthetic and real data experiments. All cases are tested in the upper Heihe River Basin located in northwest China.

6.1. Hydrological Modeling Experiment

Streamflow, especially flood peak rate estimations from GSWAT, is a little higher than SWAT under the same forcing conditions (Figure 3). That is due to some parameters, e.g., time for flow from farthest point in subbasin to enter a channel whose calculation is based on HRU and subbasin area, derived from SWAT are not most suitable for GSWAT. The area of watershed delineated by GSWAT and SWAT are 9674 km^2 and 9759 km^2 respectively. Area mismatch between GSWAT and SWAT is over 85 km^2. Another reason is that when irregular HRUs (Figure 4a) are outputted to grid cells (Figure 4b), scale issue will cause parameter mismatch [28,44]. The parameter linkages across scales from original HRU to gridded HRU will produce bias [44]. The discrepancy between GSWAT and SWAT is mainly caused by system bias (HRU delineation) [45]. It can be diminished through model calibration. Monthly and daily streamflow trend estimations from GSWAT and SWAT agree well with each other ($R^2 = 0.99$ and $R^2 = 0.96$) (Figure 3). The overestimation of peak value can be corrected with proper parameters for GSWAT at daily time scale. As high correlation exists between GSWAT and SWAT, parameters for GSWAT can be calibrated as well as SWAT against streamflow observation. Then the GSWAT model will have the ability to simulate watershed streamflow as well as SWAT.

Figure 5 plots averaged annual basin hydrological variables obtained from SWAT and GSWAT from 1991 to 2008. Precipitations are the same for both models. Snow water equivalents are sublimated in both cases. GSWAT presents overestimations of surface runoff, groundwater flow and recharge, percolation and Potential ET (PET), underestimations of Lateral flow and evapotranspiration (ET) compared with SWAT. As noted above, that is due to parameter mismatching after grid HRU delineation. In the end, GSWAT produces more water yield than SWAT. For streamflow predictions, unsuitable parameters (e.g., slope length and CN2) related to HRU area and shape attributes for GSWAT can be calibrated against observations individually instead of using parameters calibrated based on SWAT.
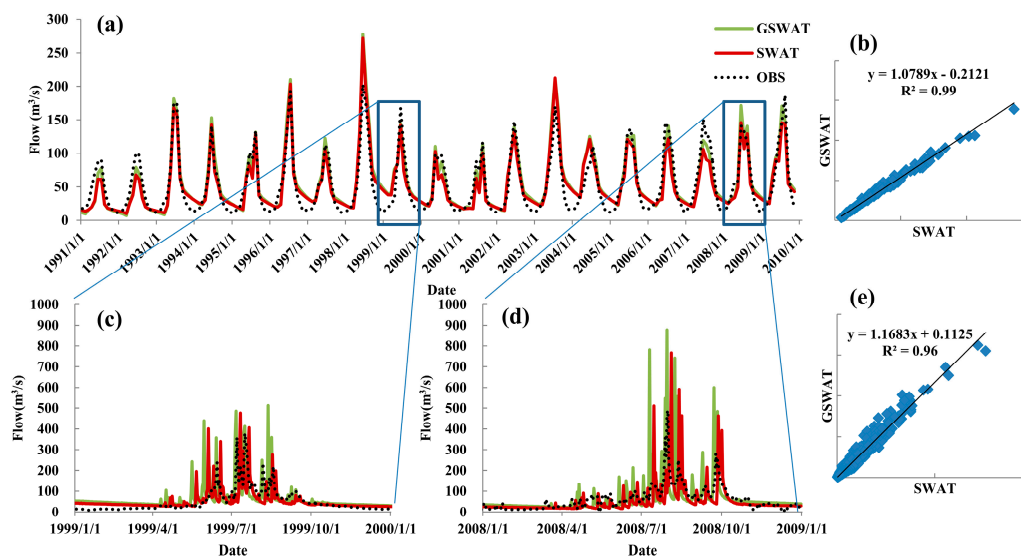

Figure 3. Monthly (**a**) and daily (**c**,**d**) streamflow estimations and linear correlation (**b**,**e**) from GSWAT and SWAT at the outlet of the upper Heihe River Basin.

Previously, the parameters for GSWAT are derived from calibrated SWAT. The parameters may not be appropriate for GSWAT. In this case, HRU level parameters (i.e., CN2) for GSWAT were calibrated against streamflow observations at Yingluoxia station from 1991 to 2008 individually. We calibrated the HRU level parameters by percentage manually. In validation period (2009), we got Root Mean Square Error (RMSE) of 63 m^3/s (79 m^3/s), correlation coefficient (R) of 0.57 (0.49) and Mean Absolute Error (MAE) of 32 m^3/s (35 m^3/s) for GSWAT (SWAT) respectively (see Table 5). With the same

simple calibration procedure, calibrated GSWAT (GSWATCALI) can present a better performance than calibrated SWAT for streamflow prediction.
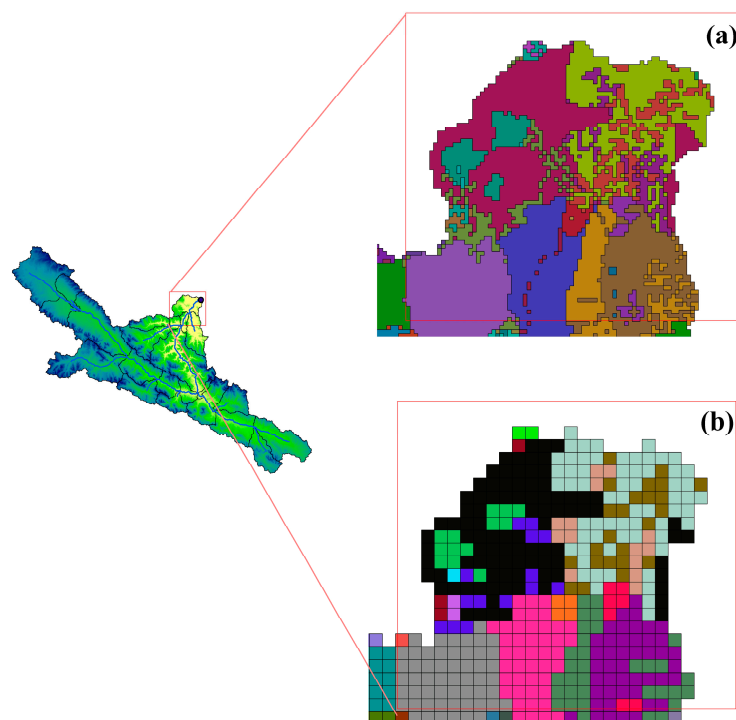
Figure 4. HRU delineations from SWAT ((**a**), HRU is the same color area) and GSWAT ((**b**), HRU is grid cell) in the upper Heihe River Basin. The same color represents the same parameters.
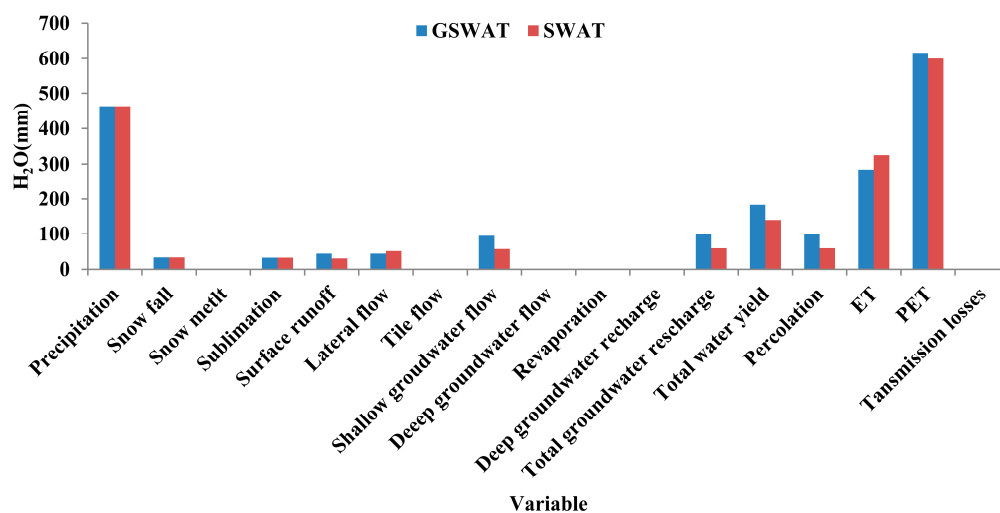
Figure 5. Averaged annual basin variables estimations from GSWAT and SWAT in the upper Heihe River Basin simulation within 1991–2008.

We have tested another grid version of SWAT (SWATgrid) and found that there is little difference (The correlation coefficient between the two data sets is 0.998) between SWATgrid and SWAT [46]. So we just compared our model (GSWAT) with original SWAT. Additionally, SWATgrid is an interface to prepare the input files (e.g., *sub, *.hru) for running SWAT in a grid subbasin-based scheme. There is a large amount of input files (the input files are determined by the number of grid cells), which will affect the application of SWATgrid in large-area and high-resolution modeling and data assimilation,

which is our main focus in our future studies. So we developed our own grid-based SWAT model and also incorporated a parallel technique (OpenMP) into it to speedup our model [29].

Table 5. Statistics of daily streamflow estimations for SWAT, GSWAT, GSWATCALI, GSWATET and GSWATCALIET against observations at the outlet of the upper Heihe River Basin in 2009.

|	**SWAT**	**GSWAT**	**GSWATCALI**	**GSWATET**	**GSWATCALIET**
RMSE (m^3/s)	78.8	97.4	63.4	78.5	56.8
R^2	0.49	0.45	0.57	0.49	0.60
MAE (m^3/s)	34.5	37.8	31.9	33.9	30.8
NSE	0.17	0.19	0.23	0.17	0.17

6.2. Evapotranspiration Data Assimilation Experiments

6.2.1. Synthetic Experiment

The result from GSWAT with ET assimilated (DA) is closer to true streamflow than GSWAT without data assimilation (NODA) (Figure 6). Figure 6 shows that DA decreased RMSE from 83.57 m^3/s to 28.13 m^3/s and MAE from 36.05 m^3/s to 10.40 m^3/s and increased NSE from 0.41 to 0.77 and R^2 from 0.98 to 0.99. In streamflow estimation, DA performs better than NODA with lower RMSE, MAE and higher NSE, R^2, especially for peak value (flood) estimation. It is essential to consider these forecasting accuracy when developing and implementing water quantity deployment actions in a watershed. Streamflow estimation error in DA is lower than that in NODA. The uncertainty in surface runoff modeling was successfully decreased by ET observation integration. The ET observation can correct not only ET modeling uncertainty but also surface runoff modeling uncertainty. The state augmentation direct insertion ET assimilation integration proved to be a successful method to improve the streamflow prediction for GSWAT. ET can also be used to constrain SWAT calibration [47]. The effectiveness of decreasing model errors with ET data assimilation method is similar to calibrating hydrological model with ET measurement. Both of them use additional measurement to correct water cycle components of the model based on their own accuracies. In synthetic calibration experiment, the calibration with accurate ET measurements is more resistant to errors and a robust way to produce accurate streamflow predictions [12].

From Figures 7 and 8, we can see that precipitation bias (Pbias equal to perturbed precipitation minus true precipitation), ET bias (ETbias) and surface runoff bias (SQbias) are spatially heterogeneous. The spatial patterns among ET correct coefficient, ET observation, ET bias and precipitation bias are related to each other. Surface runoff bias is decreased by ET correct coefficient especially in high bias region (Figures 7f and 8f). The most improvement for surface runoff estimation corresponds to the overlay of high ETbias and high SQbias (Figure 8b,c,f). In this synthetic data experiment case, precipitation uncertainty dominates the uncertainty in ET modeling and surface runoff modeling. ET observation is used to calculate the ET modeling bias correction coefficient, which is used to correct the bias in surface runoff generation modeling. The correction coefficient for surface runoff modeling is equal for ET modeling based on our assumption. This kind of assumption is the simplification of the relationship between uncertainties in ET and surface runoff modeling. However, the precipitation uncertainty for the ET modeling and surface runoff modeling being the same is not proper. This simplified assumption is suitable for modeling bias correction when the uncertainty relationship between ET and surface runoff modeling is within the range of linear uncertainty distribution. More work on the uncertainty relationship among precipitation, ET and surface runoff are needed in the future. Historic ET data can be used to correct the ET components estimations, which affect surface runoff generation in VIC [18]. The ET components estimations in GSWAT model have little effect on surface runoff generation. We cannot use the corrected ET components to update surface runoff generation in GSWAT. So in this study, we correct surface runoff

modeling bias in GSWAT model by using state augmentation assimilation method to integrate ET observation directly. The correction coefficient calculated from modeled and observed ET for ET modeling was augmented for modeled surface runoff. The simultaneous updating of surface runoff with the method we provided and real-time ET observation is worth looking forward to. Besides the rule-based insertion method, the ensemble-based data assimilation method can also be used to deal with the correlated model errors. For example, Hartanto et al. [48] assimilated satellite-based actual evapotranspiration in a distributed hydrological model to improve the streamflow predictions based on a particle filter data assimilation framework; the hydrological states from the model runs (particles) that simulate actual ET close to the remotely sensed actual ET map are then chosen as initial conditions for the next simulation. Based on a hydrological model which expresses ET as state variables, Zou et al. [49] assimilate ET into Distributed Time Variant Gain Model (DTVGM) using ensemble Kalman filter (EnKF) to obtain more accurate soil moisture and streamflow.

	NODA	DA
RMSE	84	28
NSE	0.41	0.77
R^2	0.98	0.99
MAE	36	10
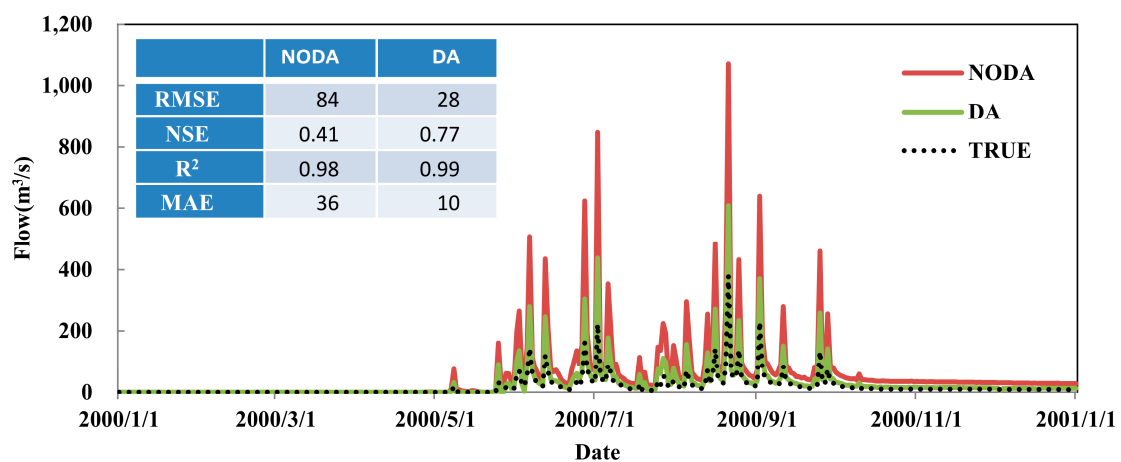
Figure 6. Streamflow observation (TRUE) and estimations from GSWAT without ET assimilated (NODA) and with ET assimilated (DA) at the Yingluoxia outlet of the upper Heihe River Basin.
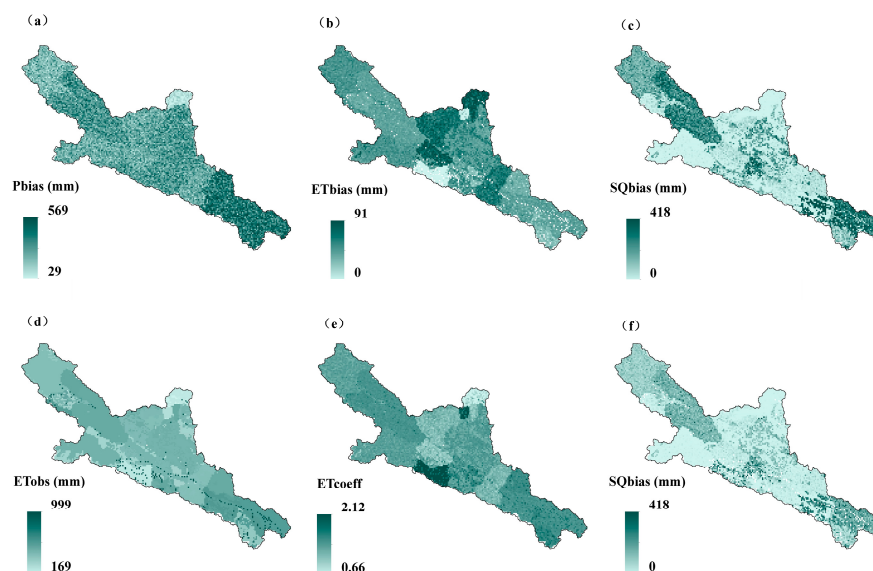
Figure 7. Annual H$_2$O volume (mm) bias in precipitation (**a**), evapotranspiration (**b**), surface runoff (**c**) from GSWAT without data assimilation and evapotranspiration observation (**d**), evapotranspiration correct coefficient (**e**) and surface runoff bias (mm) generated by GSWAT with data assimilation (**f**) in the upper Heihe River Basin.
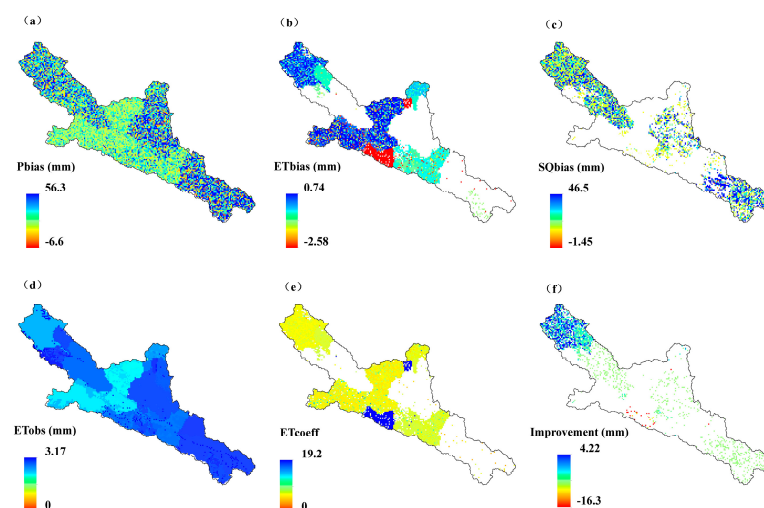
Figure 8. Bias in precipitation (Pbias) (**a**), evapotranspiration (ETbias) (**b**) and surface runoff (SQbias) (**c**) for GSWAT without data assimilation (NODA); evapotranspiration observation (ETobs) (**d**), evapotranspiration correct coefficient (ETcoeff) (**e**) and surface runoff estimation improvement (SQbias derived from NODA minus SQbias derived from DA) (**f**) generated by GSWAT with data assimilation (DA) at June 24 in the upper Heihe River Basin.

### 6.2.2. Real Data Experiment

Figure 9 plots daily streamflow observations at the Yingluoxia station and predictions obtained from GSWAT, GSWAT with ET assimilated (GSWATET), calibrated SWAT (SWAT), calibrated GSWAT (GSWATCALI) and calibrated GSWAT with ET assimilated (GSWATCALIET). The calibration is used to diminish some system biases (parameters biases) in models. Table 5 summarizes the statistical fit criteria for daily streamflow predictions from these cases. We get the following results and deductions:

(1) Calibration integration: GSWATCALI performs better than GSWAT and GSWATCALIET performs better than GSWATET. For example, the RMSE for GSWATCALI is less than for the GSWAT (82 versus 97 m^3/s). The calibrated SWAT parameters may not be suitable for GSWAT daily streamflow estimations. Daily scale streamflow estimation can be further improved by calibrating model parameters based on daily streamflow observation. With or without ET assimilated, calibration procedure can still further diminish parameter (system) errors in models.

(2) ET assimilation integration: GSWATET shows clearly improved streamflow peak rates estimations compared with GSWAT; ET assimilation decreases RMSE from 97 m^3/s to 78 m^3/s. After being calibrated, GSWATCALIET also presents a better streamflow prediction than GSWATCALI; ET assimilation decreased RMSE from 63 m^3/s to 57 m^3/s. Model system bias has little effect on ET assimilation. With or without calibration procedure, ET assimilation can still further improve streamflow prediction for GSWAT.

Streamflow estimation from GSWATCALIET is closest to streamflow observation. GSWATCALIET is better than both GSWATCALI and GSWATET. Calibration and ET assimilation are equally important in streamflow prediction in this study. Calibration can diminish model system bias (parameter bias) [50] and data assimilation can correct both model input bias and model parameter bias [51]. Observation and model uncertainties cause discrepancy between streamflow prediction and observation. There is a very high peak streamflow prediction from all models and low peak streamflow observation, for example at 18 August. These are several types of errors in streamflow observation, such as instrumental error, representative error and artificial error (i.e., miscalculation and parallax). The errors in streamflow observation are mostly temporally consistent and cannot suddenly increase so much. So we deduce that the most possibility of causing such large discrepancy is that precipitation data

has a large bias during those days. Calibration cannot correct this kind of bias by adjusting model parameters based on the observed streamflow, but ET assimilation can diminish the effects of these kinds of errors on streamflow estimation in GSWAT.
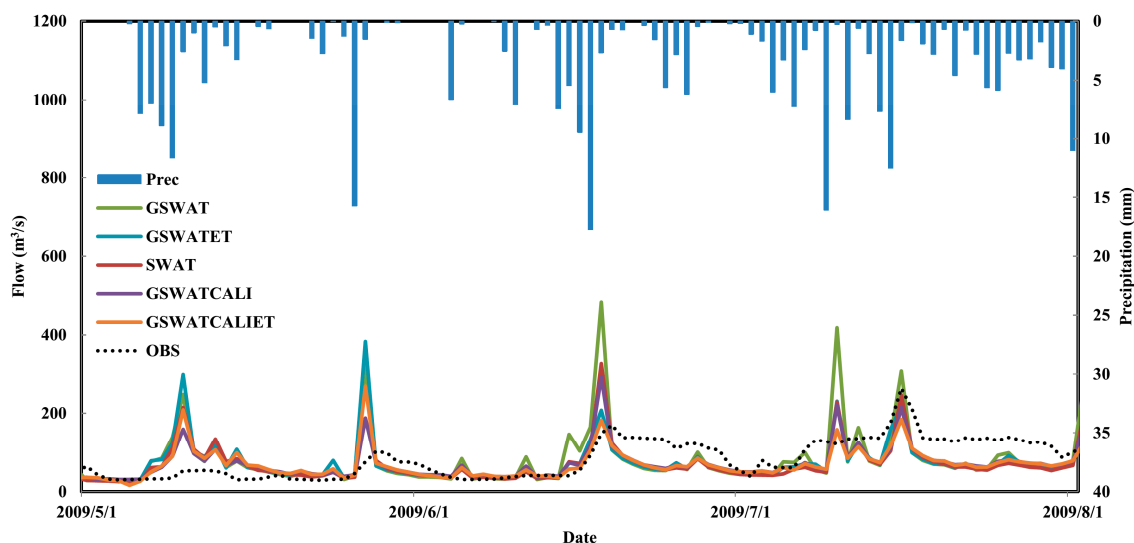

Figure 9. Daily streamflow observation (OBS) and estimations from GSWAT, GSWAT with ET assimilated (GSWATET), SWAT, calibrated GSWAT (GSWATCALI) and calibrated GSWAT with ET assimilated (GSWATCALIET) at the Yingluoxia outlet of the upper Heihe River Basin.

However, there are still some discrepancies between rainfall events and streamflow observations. For SWAT and GSWAT, the daily streamflow simulation is good (NSE > 0.50) for most years. The daily streamflow simulation results from SWAT and GSWAT in 2009 are bad (NSE < 0.50) due to that the weather stations dramatically overestimate the sub-watershed rainfall during 2009 (MAE > 30 m^3/s). For example, some of the streamflow peaks are not corresponding well to the rainfall events (see Figure 9), because the rainfall only occurred near the weather station and no rain occurred in most areas of the sub-watershed. The rainfall data representative error is too large in 2009; the NSE value will be not satisfied whenever how hard we try to calibrate the model. This kind of errors is obvious in the original SWAT, which uses one station's data to represent the whole sub-watershed when data are scarce. GSWAT model is originally developed to represent HRU level weather conditions; we can integrate more spatial representative rainfall data (i.e., radar data [52,53]) into the model to decrease this kind of errors. In this paper, we just showed the effectiveness of the ET assimilation method in decreasing rainfall volume error, which cannot handle the spatial representative error of weather stations. Precipitation input affects discharge estimates more than calibrating model parameters and assimilating remote sensing soil moisture and ET [54].

Figure 10 shows the monthly and annual hydrological variables (precipitation (PREC), surface runoff (SURQ), lateral flow (LATQ), groundwater flow (GWQ), percolation (PERC), soil water [55], evapotranspiration (ET), potential evapotranspiration (PET) and water yield (YIELD)) estimations from GSWAT and GSWATET. GSWATET shows underestimations of ET (observed), PET, SURQ and YIELD compared with GSWAT. As observed ET data is lower than GSWAT modeled, surface runoff in GSWATET is corrected to be lower than GSWAT modeled. Therefore, the water yield in GSWATET is also lower than that in GSWAT and closer to observed streamflow. The large corrected water volume is in summer where high rainfall rate and peak streamflow occurred (Figure 10). In this study, the evapotranspiration is overestimated in GSWAT. GSWATET corrects this kind of error by simply replacing it with observed data and updates surface runoff with the correct coefficient calculated from modeled and observed ET. Through the routing phase modeling, the model summarizes the updated states (water storage in reach) and fluxes (evapotranspiration and surface runoff) to the outlet

of the basin. Finally, the updated streamflow can be closer to the observed streamflow at the outlet of the basin.

ET assimilation updates the basin variable ET and SURQ, removes some estimation errors in GSWAT ET and SURQ modeling. In GSWAT, ET includes evaporation from rivers and lakes, bare soil, and vegetative surfaces; evaporation from within the leaves of plants (transpiration); and sublimation from snow and ice surfaces. The potential evapotranspiration is calculated based on the Penman-Monteith equation in this study. Potential soil water evaporation is estimated as a function of potential evapotranspiration and Leaf Area Index (LAI). Actual soil water evaporation is estimated by using exponential functions of soil depth and water content. There are large uncertainties (such as LAI and soil water content errors) in evapotranspiration modeling in GSWAT. The bias in modeled ET can be attributed to imperfect modeling processes for various land cover types, parameterizations and forcing input uncertainties etc. The flood peak is impacted by the magnitude and distribution of precipitation, antecedent soil moisture condition and surface runoff generation in the basin. In addition to long-term RMSE, R, MAE, streamflow results should be evaluated based on their ability to capture peak streamflow values. In 2009, observed peak streamflow does not correspond well to precipitation that indicates that the forcing data has large bias. The high-quality ET product is derived from merged multiple remote sensing data including TRMM precipitation data, WRF data and comprehensive ET model. Merged forcing data has lower bias than that from a single source [56,57]. This kind of ET product integrating low bias forcing data and comprehensive ET modeling process is better than observed or modeled ET alone [37]. With this kind of ET data, GSWATET can correct errors in ET modeling and other related variables modeling. In the end, GSWATET presents a superior streamflow prediction than GSWAT. Besides calibration, assimilating a reliable ET product into gridded SWAT improves the streamflow prediction in the upper Heihe River Basin.
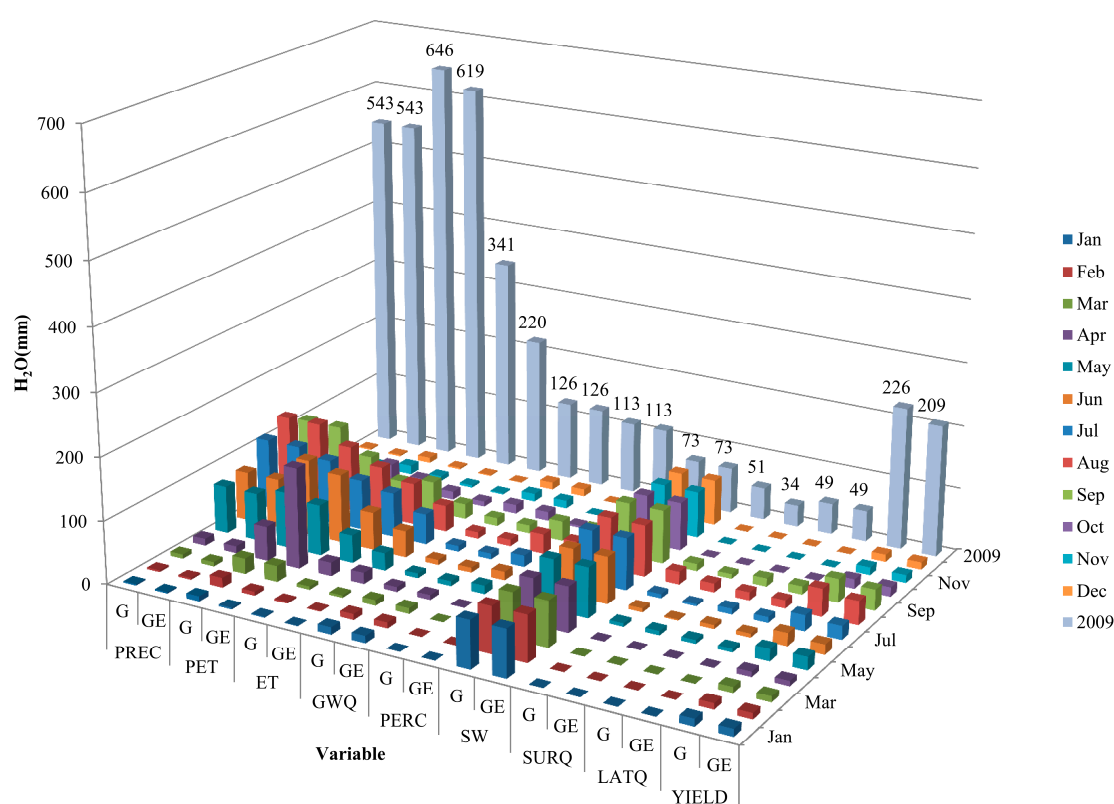
Figure 10. Monthly and annual hydrological variables estimations from GSWAT (G) and GSWATET (GE) in 2009 at the outlet of the upper Heihe River Basin.

7. Summary and Outlooks

Streamflow is an important water resource for sustainable watershed management. To promote the accuracy of the streamflow estimation in a semi-arid basin, the evapotranspiration product was assimilated into the gridded SWAT (GSWAT) model using the direct insertion method. We first modified SWAT to incorporate grid format data and assessed its capability in hydrological modeling. Parameters were further calibrated to improve daily streamflow prediction for SWAT and GSWAT. We related the ET modeling uncertainty (mainly precipitation uncertainty) to surface runoff modeling and updated it with incoming ET observation at daily time scale. The observed ET along with modeled ET updated surface runoff by using correct coefficient. Correct coefficient was calculated based on the assumption of the precipitation uncertainty for surface runoff and evapotranspiration modeling is correlated. This kind of assumption can be further improved with deeper investigation. Updated fluxes enter into the reach and update the water storage, which will route to the outlet of the basin finally. A synthetic data experiment and a real data experiment were carried out for a semi-arid basin to validate the efficacy of the state augmentation direct insertion method. The results show that the state augmentation direct insertion (DI) assimilation method can achieve an acceptable streamflow estimation improvement with a reliable ET observation, which derived from the combination of multisource remote sensing data and ET model with multiple parameterizations [37].

Parameters for models (SWAT, GSWAT and GSWATET) are calibrated against daily and monthly streamflow observations from 1990 to 1999 and validated from 2000 to 2009. Spatial domain and parameter mismatch between gridded and original HRU dominate the small discrepancies between basin variables simulated by GSWAT and SWAT. Spatial resolution of GSWAT can be matched with SWAT but will cost a lot of computational time. The detailed comparison between GSWAT and SWAT has been studied and will not be discussed in greater detail here [28]. Toward perfect streamflow prediction, calibration and data assimilation are all necessary. When the calibration is not sufficient, the data assimilation method can be used as a complement tool to diminish forecasting errors sourced from model uncertainties. In this study, parameter optimization and data assimilation procedures are separate. The simultaneous optimization and data assimilation (SODA) method can be used for improving hydrological forecasting [58,59].

Not only streamflow but also other hydrological variables (e.g., soil moisture and snow water equivalent) can be updated by assimilating ET product into the model. We did not include this kind of study in this article to briefly highlight the possibility of improving streamflow estimation with ET assimilated. Different ET products generated from different methods (e.g., [60]) can also be assimilated into the model to evaluate their effects on regional water balance simulation. Besides ET products, GSWAT can also assimilate other remote sensing data such as snow state products and soil moisture products from MODIS, AMSR-E, SMOS (Soil Moisture and Ocean Salinity) and SMAP (Soil Moisture Active and Passive) etc. Remotely sensed precipitation data can also be assimilated into GSWAT to decrease the forcing input uncertainty. The use of remote sensing data for hydrological model data assimilation has important implications to the catchments with sparse or no gauging. With more and more available remote sensing data, data assimilation technology will be a promising method to merge these data to provide more accurate land surface states and fluxes products for assisting watershed management and flood/draught prediction.

Author Contributions: Ying Zhang, Jinliang Hou, Juan Gu and Chunlin Huang conceived and designed the experiments; Ying Zhang performed the experiments and analyzed the data; Ying Zhang contributed models; Ling Zhang contributed parameter data; Ying Zhang wrote the paper.

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

Abbreviations

The following abbreviations are used in this manuscript:

ET	Evapotranspiration
SWAT	Soil and Water assessment Tool
HRU	Hydrologic Response Unit
GSWAT	Gridded Soil and Water Assessment Toll
DI	Direct Insertion
MODIS	Moderate Resolution Imaging Spectroradiometer
VIC	Variable Infiltration Capacity
AMSR-E	Advanced Microwave Scanning Radiometer for Earth Observing System
SWE	Snow Water Equivalent
LAI	Leaf Area Index
TRMM	Tropical Rainfall Measuring Mission
PET	Potential Evapotranspiration
CN2	Curve Number II
SWATCALI	Calibrated SWAT model
GSWATCALI	Calibrated Gridded SWAT model
GSWATCALIET	Calibrated Gridded SWAT model with evapotranspiration assimilated

References

1. Vorosmarty, C.J.; Green, P.; Salisbury, J.; Lammers, R.B. Global water resources: Vulnerability from climate change and population growth. *Science* **2000**, *289*, 284–288. [CrossRef] [PubMed]
2. Moiwo, J.P.; Tao, F.L. Contributions of precipitation, irrigation and soil water to evapotranspiration in (semi)-arid regions. *Int. J. Climatol.* **2015**, *35*, 1079–1089. [CrossRef]
3. Tu, M.; Wang, F.; Zhou, Y.; Wang, S. Gridded water resource distribution simulation for China based on third-order basin data from 2002. *Sustainability* **2016**, *8*, 1309. [CrossRef]
4. Fernald, A.; Tidwell, V.; Rivera, J.; Rodríguez, S.; Guldan, S.; Steele, C.; Ochoa, C.; Hurd, B.; Ortiz, M.; Boykin, K. Modeling sustainability of water, environment, livelihood, and culture in traditional irrigation communities and their linked watersheds. *Sustainability* **2012**, *4*, 2998–3022. [CrossRef]
5. Aubert, D.; Loumagne, C.; Oudin, L. Sequential assimilation of soil moisture and streamflow data in a conceptual rainfall-runoff model. *J. Hydrol.* **2003**, *280*, 145–161. [CrossRef]
6. Lievens, H.; Tomer, S.K.; Al Bitar, A.; De Lannoy, G.J.M.; Drusch, M.; Dumedah, G.; Franssen, H.J.H.; Kerr, Y.H.; Martens, B.; Pan, M.; et al. Smos soil moisture assimilation for improved hydrologic simulation in the murray darling basin, australia. *Remote Sens. Environ.* **2015**, *168*, 146–162. [CrossRef]
7. Montero, R.A.; Schwanenberg, D.; Krahe, P.; Lisniak, D.; Sensoy, A.; Sorman, A.A.; Akkol, B. Moving horizon estimation for assimilating h-saf remote sensing data into the hbv hydrological model. *Adv. Water Resour.* **2016**, *92*, 248–257. [CrossRef]
8. Lu, H.S.; Crow, W.T.; Zhu, Y.H.; Ouyang, F.; Su, J.B. Improving streamflow prediction using remotely-sensed soil moisture and snow depth. *Remote Sens.* **2016**, *8*, 503. [CrossRef]
9. Lopez, P.L.; Wanders, N.; Schellekens, J.; Renzullo, L.J.; Sutanudjaja, E.H.; Bierkens, M.F.P. Improved large-scale hydrological modelling through the assimilation of streamflow and downscaled satellite soil moisture observations. *Hydrol. Earth Syst. Sci.* **2016**, *20*, 3059–3076. [CrossRef]
10. Immerzeel, W.W.; Droogers, P. Calibration of a distributed hydrological model based on satellite evapotranspiration. *J. Hydrol.* **2008**, *349*, 411–424. [CrossRef]
11. Huang, C.; Li, Y.; Gu, J.; Lu, L.; Li, X. Improving estimation of evapotranspiration under water-limited conditions based on sebs and modis data in arid regions. *Remote Sens.* **2015**, *7*, 16795–16814. [CrossRef]
12. Kunnath-Poovakka, A.; Ryu, D.; Renzullo, L.J.; George, B. The efficacy of calibrating hydrologic model using remotely sensed evapotranspiration and soil moisture for streamflow prediction. *J. Hydrol.* **2016**, *535*, 509–524. [CrossRef]
13. Vervoort, R.W.; Miechels, S.F.; van Ogtrop, F.F.; Guillaume, J.H.A. Remotely sensed evapotranspiration to calibrate a lumped conceptual model: Pitfalls and opportunities. *J. Hydrol.* **2014**, *519*, 3223–3236. [CrossRef]

14. Pan, M.; Wood, E.F.; Wojcik, R.; McCabe, M.F. Estimation of regional terrestrial water cycle using multi-sensor remote sensing observations and data assimilation. *Remote Sens. Environ.* **2008**, *112*, 1282–1294. [CrossRef]

15. Yin, J.; Zhan, C.; Ye, W. An experimental study on evapotranspiration data assimilation based on the hydrological model. *Water Resour. Manag.* **2016**, *30*, 5263–5279. [CrossRef]

16. Irmak, A.; Kamble, B. Evapotranspiration data assimilation with genetic algorithms and swap model for on-demand irrigation. *Irrig. Sci.* **2009**, *28*, 101–112. [CrossRef]

17. Rientjes, T.H.M.; Muthuwatta, L.P.; Bos, M.G.; Booij, M.J.; Bhatti, H.A. Multi-variable calibration of a semi-distributed hydrological model using streamflow data and satellite-based evapotranspiration. *J. Hydrol.* **2013**, *505*, 276–290. [CrossRef]

18. Parr, D.; Wang, G.L.; Bjerklie, D. Integrating remote sensing data on evapotranspiration and leaf area index with hydrological modeling: Impacts on model performance and future predictions. *J. Hydrometeorol.* **2015**, *16*, 2086–2100. [CrossRef]

19. Wang, Q.J.; Pagano, T.C.; Zhou, S.L.; Hapuarachchi, H.A.P.; Zhang, L.; Robertson, D.E. Monthly versus daily water balance models in simulating monthly runoff. *J. Hydrol.* **2011**, *404*, 166–175. [CrossRef]

20. Robles-Morua, A.; Vivoni, E.R.; Mayer, A.S. Distributed hydrologic modeling in northwest mexico reveals the links between runoff mechanisms and evapotranspiration. *J. Hydrometeorol.* **2012**, *13*, 785–807. [CrossRef]

21. Rawlins, M.A.; Frolking, S.; Lammers, R.B.; Vorosmarty, C.J. Effects of uncertainty in climate inputs on simulated evapotranspiration and runoff in the western arctic. *Earth Interact.* **2006**, *10*. [CrossRef]

22. Vrugt, J.A.; ter Braak, C.J.F.; Clark, M.P.; Hyman, J.M.; Robinson, B.A. Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with markov chain monte carlo simulation. *Water Resour. Res.* **2008**, *44*. [CrossRef]

23. Tian, Y.D.; Huffman, G.J.; Adler, R.F.; Tang, L.; Sapiano, M.; Maggioni, V.; Wu, H. Modeling errors in daily precipitation measurements: Additive or multiplicative? *Geophys. Res. Lett.* **2013**, *40*, 2060–2065. [CrossRef]

24. Shao, Q.X.; Lerat, J.; Brink, H.; Tomkins, K.; Yang, A.; Peeters, L.; Li, M.; Zhang, L.; Podger, G.; Renzullo, L.J. Gauge based precipitation estimation and associated model and product uncertainties. *J. Hydrol.* **2012**, *444*, 100–112. [CrossRef]

25. Arnold, J.G.; Srinivasan, R.; Muttiah, R.S.; Williams, J.R. Large area hydrologic modeling and assessment—Part 1: Model development. *J. Am. Water Resour. Assoc.* **1998**, *34*, 73–89. [CrossRef]

26. Lee, G.; Shin, Y.; Jung, Y. Development of web-based recess model for estimating baseflow using swat. *Sustainability* **2014**, *6*, 2357–2378. [CrossRef]

27. Douglas-Mankin, K.R.; Srinivasan, R.; Arnold, J.G. Soil and water assessment tool (swat) model: Current developments and applications. *ASABE* **2010**, *53*, 1423–1431. [CrossRef]

28. Zhang, Y.; Hou, J.L.; Cao, Y.P.; Huang, C.L. Development and evaluation of a hru-based gridded approach in swat model for watershed-scale hydrological modelling. *Environ. Model. Softw.* **2016**, under review.

29. Zhang, Y.; Hou, J.; Cao, Y.; Gu, J.; Huang, C. Openmp parallelization of a gridded swat (swatg). *Comput. Geosci.* **2017**. [CrossRef]

30. Dhungel, R.; Allen, R.G.; Trezza, R.; Robison, C.W. Evapotranspiration between satellite overpasses: Methodology and case study in agricultural dominant semi-arid areas. *Meteorol. Appl.* **2016**, *23*, 714–730. [CrossRef]

31. Neitsch, S.L.; Williams, J.R.; Arnold, J.; Kiniry, J. *Soil and Water Assessment Tool Theoretical Documentation Version 2009*; Texas Water Resources Institute: College Station, TX, USA, 2011.

32. Easton, Z.M.; Fuka, D.R.; Walter, M.T.; Cowan, D.M.; Schneiderman, E.M.; Steenhuis, T.S. Re-conceptualizing the soil and water assessment tool (swat) model to predict runoff from variable source areas. *J. Hydrol.* **2008**, *348*, 279–291. [CrossRef]

33. Li, X.; Bai, Y. Assimilating remote sensing data into land surface models: Theory and methods. In *Land Surface Observation, Modeling and Data Assimilation*; World Scientific: Singapore, 2013; pp. 143–170.

34. Solomon, S. Relationship between precipitation evaporation and runoff in tropical-equatorial regions. *Water Resour. Res.* **1967**, *3*, 163–172. [CrossRef]

35. Cheng, G.; Xin, L.; Zhao, W.; Xu, Z.; Qi, F.; Xiao, S.; Xiao, H. Integrated study of the water-ecosystem-economy in the heihe river basin. *Natl. Sci. Rev.* **2014**, *1*, 413–428. [CrossRef]

36. Zhang, L.; Nan, Z.T.; Yu, W.J.; Ge, Y.C. Modeling land-use and land-cover change and hydrological responses under consistent climate change scenarios in the heihe river basin, china. *Water Resour. Manag.* **2015**, *29*, 4701–4717. [CrossRef]

37. Hu, G.C.; Jia, L. Monitoring of evapotranspiration in a semi-arid inland river basin by combining microwave and optical remote sensing observations. *Remote Sens.* **2015**, *7*, 3056–3087. [CrossRef]

38. Cui, Y.K.; Jia, L. A modified gash model for estimating rainfall interception loss of forest using remote sensing observations at regional scale. *Water* **2014**, *6*, 993–1012. [CrossRef]

39. Kuzmin, P.P. On method for investigations of evaporation from the snow cover. *Trans. State Hydrol. Inst.* **1953**, *41*, 34–52. (In Russian)

40. Penman, H.L. Natural evaporation from open water, bare soil and grass. *Proc. R. Soc. Lond.* **1948**, *193*, 120. [CrossRef]

41. Dai, Y.; Shangguan, W.; Duan, Q.; Liu, B.; Fu, S.; Niu, G. Development of a china dataset of soil hydraulic parameters using pedotransfer functions for land surface modeling. *J. Hydrometeorol.* **2013**, *14*, 869–887. [CrossRef]

42. Zambrano-Bigiarini, M.; Rojas, R. A model-independent particle swarm optimisation software for model calibration. *Environ. Model. Softw.* **2013**, *43*, 5–25. [CrossRef]

43. Moradkhani, H.; Meskele, T.T. *Probabilistic Assessment of the Satellite Rainfall Retrieval Error Translation to Hydrologic Response*; Springer: Dordrecht, The Netherlands, 2010; pp. 229–242.

44. Bloschl, G.; Sivapalan, M. Scale issues in hydrological modeling—A review. *Hydrol. Process.* **1995**, *9*, 251–290. [CrossRef]

45. 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. *ASABE* **2010**, *53*, 1433–1443. [CrossRef]

46. Rathjens, H.; Oppelt, N. Swatgrid: An interface for setting up swat in a grid-based discretization scheme. *Comput. Geosci.* **2012**, *45*, 161–167. [CrossRef]

47. Tobin, K.J.; Bennett, M.E. Constraining swat calibration with remotely sensed evapotranspiration data. *J. Am. Water Resour. Assoc.* **2017**, *53*, 593–604. [CrossRef]

48. Hartanto, I.M.; van der Kwast, J.; Alexandridis, T.K.; Almeida, W.; Song, Y.; van Andel, S.J.; Solomatine, D.P. Data assimilation of satellite-based actual evapotranspiration in a distributed hydrological model of a controlled water system. *Int. J. Appl. Earth Obs.* **2017**, *57*, 123–135. [CrossRef]

49. Zou, L.; Zhan, C.; Xia, J.; Wang, T.; Gippel, C.J. Implementation of evapotranspiration data assimilation with catchment scale distributed hydrological model via an ensemble kalman filter. *J. Hydrol.* **2017**, *549*, 685–702. [CrossRef]

50. 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.; et al. Swat: Model use, calibration, and validation. *ASABE* **2012**, *55*, 1491–1508. [CrossRef]

51. Huang, C.L.; Hu, B.X.; Li, X.; Ye, M. Using data assimilation method to calibrate a heterogeneous conductivity field and improve solute transport prediction with an unknown contamination source. *Stoch. Env Res. Risk A* **2009**, *23*, 1155–1167. [CrossRef]

52. Moon, J.; Srinivasan, R.; Jacobs, J.H. Stream flow estimation using spatially distributed rainfall in the trinity River Basin, Texas. *Trans ASAE* **2004**, *47*, 1445–1451. [CrossRef]

53. Tobin, K.J.; Bennett, M.E. Using swat to model streamflow in two river basins with ground and satellite precipitation data. *Jawra J. Am. Water Resour. Assoc.* **2010**, *45*, 253–271. [CrossRef]

54. Lopez, P.L.; Sutanudjaja, E.H.; Schellekens, J.; Sterk, G.; Bierkens, M.F.P. Calibration of a large-scale hydrological model using satellite-based soil moisture and evapotranspiration products. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 3125–3144. [CrossRef]

55. Hassan, S.M.T.; Lubczynski, M.W.; Niswonger, R.G.; Su, Z.B. Surface-groundwater interactions in hard rocks in Sardon catchment of western Spain: An integrated modeling approach. *J. Hydrol.* **2014**, *517*, 390–410. [CrossRef]

56. Pan, X.D.; Li, X.; Yang, K.; He, J.; Zhang, Y.L.; Han, X.J. Comparison of downscaled precipitation data over a mountainous watershed: A case study in the heihe river basin. *J. Hydrometeorol.* **2014**, *15*, 1560–1574. [CrossRef]

57. Pan, X.D.; Li, X.; Cheng, G.D.; Li, H.Y.; He, X.B. Development and evaluation of a river-basin-scale high spatio-temporal precipitation data set using the wrf model: A case study of the heihe river basin. *Remote Sens.* **2015**, *7*, 9230–9252. [CrossRef]

58. Vrugt, J.A.; Diks, C.G.H.; Gupta, H.V.; Bouten, W.; Verstraten, J.M. Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation. *Water Resour. Res.* **2005**, *41*. [CrossRef]

59. Chu, N.; Huang, C.L.; Li, X.; Du, P.J. Simultaneous estimation of surface soil moisture and soil properties with a dual ensemble kalman smoother. *Sci. China Earth Sci.* **2015**, *58*, 2327–2339. [CrossRef]

60. Dhungel, R.; Allen, R.G.; Trezza, R.; Robison, C.W. Comparison of latent heat flux using aerodynamic methods and using the penman-monteith method with satellite-based surface energy balance. *Remote Sens.* **2014**, *6*, 8844–8877. [CrossRef]