


## Article

# On the Evaluation of Both Spatial and Temporal Performance of Distributed Hydrological Models Using Remote Sensing Products

Tam V. Nguyen <sup>1,\*</sup> , Bhumika Uniyal <sup>2</sup>, Dang An Tran <sup>3</sup> and Thi Bich Thuc Pham <sup>4</sup>

<sup>1</sup> Department of Hydrogeology, Helmholtz Centre for Environmental Research (UFZ), 04318 Leipzig, Germany

<sup>2</sup> Professorship of Ecological Services, Bayreuth Center of Ecology and Environmental Research, Bayreuth University, 95447 Bayreuth, Germany; bhumika.uniyal@uni-bayreuth.de

<sup>3</sup> Faculty of Water Resources Engineering, Thuyloi University, 175 Tay Son, Dong Da, Hanoi 100000, Vietnam; antd@tlu.edu.vn

<sup>4</sup> Institute of Applied Mechanics and Informatics, Vietnam Academy of Science and Technology, 291 Dien Bien Phu Street, Ward 7, District 3, Ho Chi Minh City 70000, Vietnam; ptbthuc@iami.vast.vn

\* Correspondence: tam.nguyen@ufz.de

**Abstract:** Evaluating the spatial and temporal model performance of distributed hydrological models is necessary to ensure that the simulated spatial and temporal patterns are meaningful. In recent years, spatial and temporal remote sensing data have been increasingly used for model performance evaluation. Previous studies, however, have focused on either the temporal or spatial model performance evaluation. In addition, temporal (or spatial) model performance evaluation is often conducted in a spatially (or temporally) lumped approach. Here, we evaluated (1) the temporal model performance evaluation in a spatially distributed approach (spatiotemporal) and (2) the spatial model performance in a temporally distributed approach (temporospatial). We further demonstrated that both spatiotemporal and temporospatial model performance evaluations are necessary since they provide different aspects of the model performance. For this, a case study was developed using the Soil and Water Assessment Tool (SWAT) for the Upper Baitarani catchment in India, and the spatiotemporal and temporospatial model performance was evaluated against three different remotely based actual evapotranspiration (ETa) products (MOD16 A2, SSEBop, and TerraClimate). The results showed that an increase in the spatiotemporal model performance would not necessarily lead to an increase in the temporospatial model performance and vice versa, depending on the evaluation statistics. Overall, this study has highlighted the necessity of a joint spatiotemporal and temporospatial model performance evaluation to understand/improve spatial and temporal model behavior/performance.

**Keywords:** distributed hydrological models; model performance evaluation; spatiotemporal; temporospatial; remote-sensing evapotranspiration



**Citation:** Nguyen, T.V.; Uniyal, B.; Tran, D.A.; Pham, T.B.T. On the Evaluation of Both Spatial and Temporal Performance of Distributed Hydrological Models Using Remote Sensing Products. *Remote Sens.* **2022**, *14*, 1959. <https://doi.org/10.3390/rs14091959>

Academic Editors: Yang Hong, Jinsong Deng and Salah Elsayed

Received: 29 March 2022

Accepted: 18 April 2022

Published: 19 April 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 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 (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Spatial heterogeneity in catchment characteristics (land use, soil type, and slope) and spatial and temporal variations in hydrological forcing (rainfall, wind speed, and solar radiation) are often observed in nature [1]. These factors cause spatial and temporal variations of hydrological processes/responses, e.g., soil moisture [2,3], evapotranspiration [4], and groundwater recharge [5]. Spatially and temporally distributed hydrological models (herein referred to as distributed models) have been proven as effective tools for modeling spatial and temporal hydrological processes, e.g., the Soil and Water Assessment Tool (SWAT) [6,7], the TOPography-based hydrological MODEL (TOPMODEL) [8], and the Variable Infiltration Capacity (VIC) model [9], and the mesoscale Hydrologic Model (mHM) [10,11].

In order to ensure that the distributed models provide meaningful spatial and temporal patterns of the interested processes, the simulated spatial and temporal patterns should be

evaluated against the reference spatiotemporal data. In recent years, remote-sensing data of high spatial and temporal resolution have been increasingly used for model evaluation, especially in data-scarce regions [12–23]. However, evaluating both the spatial and temporal performance of distributed models using remote sensing data has not been well addressed in the literature. The preceding studies have focused on evaluating either the temporal or spatial model performance. The majority of the conducted studies have evaluated the temporal model performance while the spatial evaluation is often performed using a spatially lumped approach, e.g., at the basin or subbasin level [20,21,24]. Similarly, the spatial model evaluation is often conducted using a temporally lumped approach, e.g., monthly, yearly, or the entire simulation time (e.g., Koch et al. [17]). Up to now, no study has evaluated both (1) the temporal model performance in a spatially distributed approach (hereinafter referred to as the spatiotemporal model performance evaluation) and (2) spatial model performance in a temporally distributed approach (hereinafter referred to as the temporospatial model performance evaluation).

Evaluating the spatiotemporal and temporospatial model performance are the two different aspects of the model evaluation. In this study, we focus on the spatiotemporal model performance evaluation approach which quantifies the temporal pattern matching between the simulated and reference time-series data at each spatial unit (e.g., pixel) of the model domain. This can be undertaken by using different temporal performance indices, e.g., the Nash–Sutcliffe efficiency (NSE) [25], the Kling–Gupta efficiency (KGE) [26], percentage bias (PBIAS), the root mean square error (RMSE), and the ratio of the RMSE to the standard deviation of measured data (RSR). Spatiotemporal model performance evaluation, therefore, provides information about the locations where the temporal model performance is not good (spatial model issue).

In contrast, this study also assessed the temporospatial model performance evaluation that quantifies the matching between the simulated and reference spatial patterns at each model time step. It can be undertaken using different spatial performance indices. Some of the temporal performance metrics could be used (e.g., RSR, BIAS). Some performance metrics were specially developed for evaluating the spatial pattern matching, e.g., the SPAtial Efficiency metric (SPAEF; Koch et al. [17]), fractions skill score [27], and others [28]. Among these spatial performance matrices, the SPAEF metric, which consists of three equally weighted components (correlation, coefficient of variation, and histogram overlap) was demonstrated as a robust statistical index for evaluating spatial pattern matching [17]. Temporospatial model performance evaluation provides information about the timing when the spatial model performance is not good (temporal model issue). This topic has been less explored in depth so far.

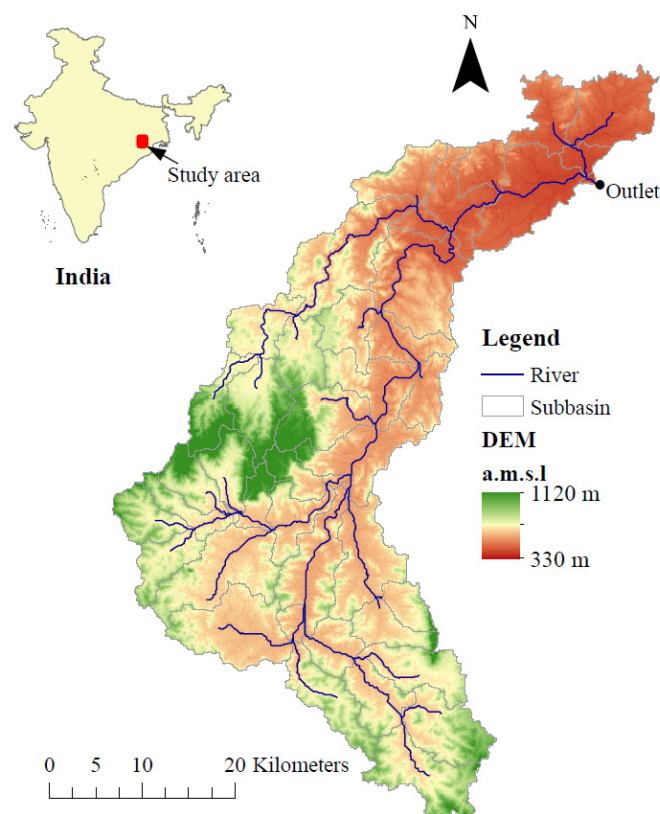
From the above discussion, it can be said that spatiotemporal and temporospatial model evaluation (1) should be used together to detect both spatial and temporal model issues, and (2) are the two different aspects of model evaluation. However, this statement has not been validated by the research conducted in the past. Therefore, the main objective of this study is to validate the two aforementioned statements. For these objectives, the Soil and Water Assessment Tool (SWAT) [6] was set up for the Upper Baitarani catchment located in a data-scare region in India. Satellite-derived actual evapotranspiration (ETa) data from multiple sources [29–31] were used as reference spatiotemporal data for model evaluation due to its global coverage and availability. Results from this study are also expected to promote the use of remotely sensed data for the spatial and temporal evaluation of distributed hydrological models.

## 2. Methodology

### 2.1. Study Area and Data

The Upper Baitarani catchment is in India with an area of about 1711 km<sup>2</sup> (Figure 1). The digital elevation model (DEM) from the Shuttle Radar Topography Mission shows that the elevation of the catchment varies in a wide range, from 330 to 1120 m above mean sea level (a.m.s.l). Despite the study area covering a relatively large area with high

variation in topography, there is only one weather station located in the area. Observed rainfall shows an average annual rainfall of 1165 mm with high temporal variation and out of which more than 80% of the rainfall occurs between June and October. The average monthly minimum and maximum temperature in the area are 11 °C and 34 °C, respectively. The dominant land uses/land covers in the area are agricultural, forest, and range grass, accounting for about 36%, 32%, and 25% of the study area, respectively. The dominant soil type in the area is sandy clay loam with low water holding capacity [32]. The observed weather data were procured from the India Meteorological Centre, Bhubaneswar, India. Furthermore, sunshine hours were used for calculating the solar radiation data for the considered catchment.

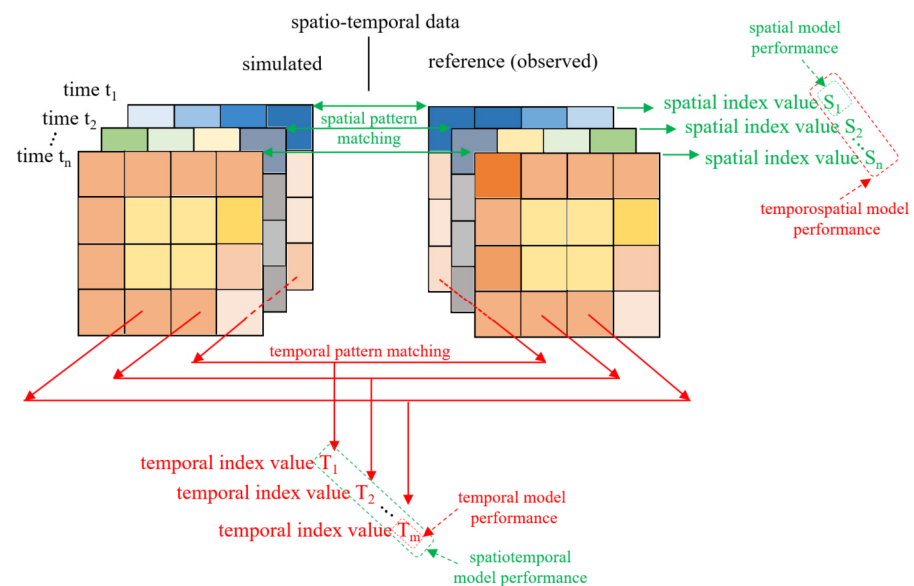


**Figure 1.** Location of the Upper Baitarani catchment and the Digital Elevation Model (DEM).

## 2.2. Spatiotemporal and Temporospatial Model Performance Evaluation

Evaluation of the spatial and temporal performance is the assessment of the simulated spatiotemporal data against the reference (or observed) spatiotemporal data. Here, the model output data that need to be evaluated are time-series data at every pixel (with an equal area) of the modeling domain (Figure 2).

In the spatiotemporal approach (Figure 2), the simulated time-series data at every pixel (spatial index *S*) are first evaluated against the reference time-series data at the corresponding pixel using a temporal performance index *T* (e.g., NSE, KGE, PBIAS, RMSE). The spatial variation of the temporal index *T* provides information on the spatiotemporal model performance. In the temporospatial approach (Figure 2), the simulated spatial patterns are evaluated against the reference spatial patterns at every time step using the spatial index *S* (e.g., SPAEF, RMSE). Information about the temporospatial model performance could be obtained from the temporal variation of the spatial index *S*.

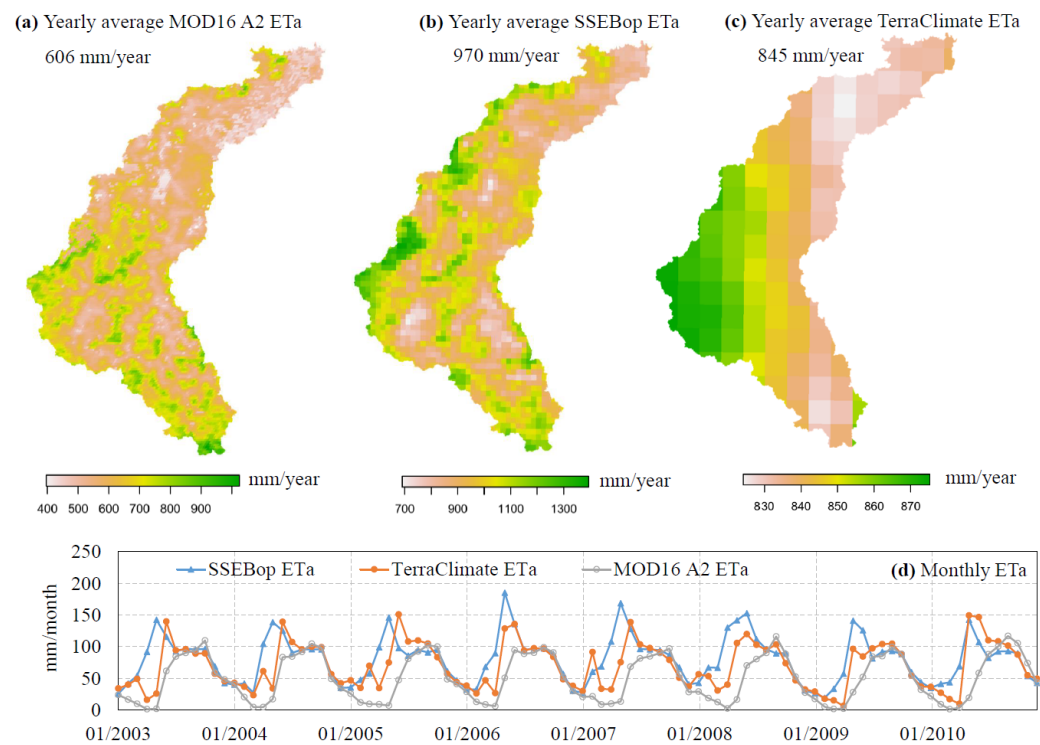


**Figure 2.** Schematics showing spatiotemporal and temporospatial approaches.

### 2.3. Reference Spatiotemporal ETa

In this study, global satellite-based ETa products from the Moderate Resolution Imaging Spectroradiometer (MOD16 A2) [29], the operational Simplified Surface Energy Balance model (SSEBop) [30], and the TerraClimate [31] were used. ETa from these products was estimated using different techniques and is available at different spatial and temporal resolutions. For example, MOD16 A2 ETa was derived using the Penman–Monteith approach with daily meteorological reanalysis data (air pressure, temperature, humidity, solar radiation) and 8-day remote-sensing vegetation indices (leaf area index, albedo, and fraction of photosynthetically active radiation). MOD16 A2 ETa is available at 0.5-km spatial resolution and 8-day interval. SSEBop ETa was calculated with the simplified energy balance model [30]. This model combines the estimated ET fraction based on MODIS thermal imagery and grass-reference potential ET. SSEBop ETa is available at 1-km resolution and monthly timestep. TerraClimate ETa was estimated based on the one-dimensional modified Thornthwaite–Mather climatic water-balance model with (1) monthly climate data from the WorldClim [33,34], the Climate Research Unit [35], and the Japanese 55-year Reanalysis [36], and (2) water storage capacity from the root zone storage capacity [37]. Global TerraClimate ETa is available at 4-km spatial resolution and monthly time step.

The differences in the absolute ETa between the three products are expected due to the different techniques and input data used to estimate ETa. The annual average ETa during the period 2003–2010 in the study area is 606 mm (MOD16 A2), 970 mm (SSEBop), and 845 mm (TerraClimate). The spatial and monthly ETa patterns also show their dissimilarities (Figure 3). However, there are some common spatial and temporal ETa patterns among these products, e.g., low ETa near the catchment outlet, high ETa in the western part of the study area, and high seasonal variation. Evaluating the accuracy of these products is not within the scope of this study. Here, we instead provide the novel techniques for spatiotemporal and temporospatial model evaluation. Future studies can evaluate the validity of different remote-sensing products before applying our approach for spatiotemporal and temporospatial model evaluations. In this study, the considered ETa products were used as reference data for model evaluation to (1) explore the SWAT model performance compared with different ETa products and (2) account for the uncertainty in the estimated ETa from different ETa products.



**Figure 3.** (a–c) Spatial and (d) temporal variations of ETa from MOD16 A2, SSEBop, and TerraClimate in the study area during the period from 2003 to 2010.

#### 2.4. The SWAT Model

SWAT is a distributed hydrological model used for evaluating the impacts of land use management practices on water, sediment, and nutrient yields [6]. SWAT has been widely used and tested in various catchments worldwide [38]. In SWAT, a catchment is divided into sub-catchments which are further sub-divided into hydrologic response units (HRUs). An HRU is a fraction of land with a unique combination of land use, soil type, and slope within a sub-catchment. SWAT simulates two phases of the hydrological cycle, the land phase and the routing phase. The land phase includes HRU-related processes (e.g., evapotranspiration, surface runoff, infiltration, lateral flow, groundwater recharge, and base flow). The routing phase simulates flow in water bodies (e.g., river, reservoir, and pond). SWAT provided outputs at different spatial levels, e.g., HRU, sub-catchment, or catchment. Outputs at the HRU level could be mapped to the HRU raster file, which contains the spatial information of HRUs, created during the model setup with ArcSWAT (e.g., Kim et al. [39]; Nguyen & Dietrich [40]).

#### 2.5. Model Setup, Parameter Variation, and Model Evaluation for ETa

##### 2.5.1. Model Setup

Based on the land use, soil type, and slope, the study area is divided into 45 catchments, which are further sub-divided into 875 hydrologic response units (HRUs). The model was set up to simulate for ten years at the daily time step with two years of warm-up (2001–2002) and eight years of model evaluation (2003–2010). In SWAT, ETa is calculated after potential evapotranspiration (ETp). There are several methods available for calculating ETp [6,41], for example, the Penman–Monteith [42–44], Priestly–Taylor [45], and Hargreaves [46]. Here, we selected the Penman–Monteith method as it was one of the widely used methods for evapotranspiration quantification [47]. Daily ETp is calculated as follows [6,42–44]:

$$ETp = \frac{\Delta \cdot (R_n - G) + \rho_a \cdot c_p \cdot (e_z^0 - e_z) / r_a}{\lambda \cdot (\Delta + \gamma \cdot (1 + r_c / r_a))}, \quad (1)$$



where  $\Delta$  is the slope of the saturated vapor pressure curve ( $\text{kPa } ^\circ\text{C}^{-1}$ ),  $R_n$  is the net radiation ( $\text{MJ m}^{-2}, \text{d}^{-1}$ ),  $G$  is the heat flux to the soil ( $\text{MJ m}^{-2} \text{d}^{-1}$ ),  $\rho_a$  is the air density ( $\text{kg m}^{-3}$ ),  $c_p$  is the specific heat at constant pressure ( $\text{MJ kg}^{-1} ^\circ\text{C}^{-1}$ ),  $e_z^0$  and  $e_z$  are the saturated and the actual vapor pressure at height  $z$  (kPa), respectively,  $r_a$  is the bulk surface aerodynamic resistance for water vapor ( $\text{s m}^{-1}$ ),  $\lambda$  is the latent heat of vaporization ( $\text{MJ kg}^{-1}$ ),  $\gamma$  is the psychrometric constant ( $\text{kPa } ^\circ\text{C}^{-1}$ ), and  $r_c$  is the canopy surface resistance ( $\text{s m}^{-1}$ ).

Daily simulated ETa from SWAT was aggregated to an 8-day interval for comparison with MOD16 A2 ETa and to a monthly interval for comparison with SSEBop and TerraClimate ETa. The aggregated ETa was then mapped to the HRU raster map of 90 m resolution that was created during the model setup using ArcSWAT (e.g., Kim et al. [39]; Nguyen & Dietrich [40]). ETa products from MOD16 A2, SSEBop and TerraClimate were spatially disaggregated to the same spatial resolution of SWAT ETa (90 m resolution) for comparison.

### 2.5.2. Parameter Variation and Model Performance Evaluation Scenarios

The objectives of parameter variation are to (1) search for a global optimal solution, and (2) find a relationship between spatiotemporal and temporospatial model performance in different model simulations. The model parameters (Table 1) were selected based on the literature review of the most common parameters for ETa calibration [19,20,23,48,49]. These parameters affect evaporation/evapotranspiration in different model compartments. For example, the soil evaporation compensation factor (ESCO) controls the maximum allowable soil evaporation with depth. Higher plant uptake compensation factors (EPCO) mean that plants can take water from deeper soil layers, resulting higher evapotranspiration. Higher groundwater “revap” coefficients mean that more water from the shallow aquifer could be moved up to the root zone for evaporation. The maximum canopy storage (CANMX) affects the amount of precipitation that can be held by plant canopy; therefore, it affects evapotranspiration. The effect of other parameters (Table 1) on evapotranspiration is not straightforward to interpret. In this study, some of the parameters were changed in a spatial approach according to land use and soil type (Table 1). Latin Hypercube Sampling (LHS) was used to generate 2000 parameter sets in a search for a global optimum solution. LHS has been demonstrated as an efficient sampling approach for searching for a global optimal solution for high-dimensional problems [23].

**Table 1.** List of model parameters and their ranges for parameter optimization.

Nr.	Parameters	Description	Parameter Range	
			Min	Max
1	r_CN2 <sup>FRSD</sup>	SCS runoff curve number of forest (FRSD),	−0.2	0.2
2	r_CN2 <sup>AGRL</sup>	agriculture (AGRL), and range grass	−0.2	0.2
3	r_CN2 <sup>RNGE</sup>	(RNGE) lands	−0.2	0.2
4	v_ESCO <sup>FRSD</sup>	Soil evaporation compensation factor of forest agriculture, and range grass lands	0	1
5	v_ESCO <sup>AGRL</sup>		0	1
6	v_ESCO <sup>RNGE</sup>		0	1
7	v_EPCO <sup>FRSD</sup>	Plant uptake compensation factor of forest agriculture, and range grass lands	0	1
8	v_EPCO <sup>AGRL</sup>		0	1
9	v_EPCO <sup>RNGE</sup>		0	1
10	v_GWQMN	Groundwater baseflow threshold (mm)	0	2000
11	v_GW_REVAP	Groundwater “revap” coefficient	0.02	0.2
12	v_REVAPMN	Groundwater “revap” threshold (mm)	0	500
13	r_SOL_AWC <sup>SOIL1</sup>	Soil available water content of soil classes 1 and 2	−0.2	0.2
14	r_SOL_AWC <sup>SOIL2</sup>		−0.2	0.2
15	r_SOL_K <sup>SOIL1</sup>	Soil hydraulic conductivity of soil classes 1 and 2 (mm/h)	−0.2	0.2
16	r_SOL_K <sup>SOIL2</sup>		−0.2	0.2
17	v_CANMX	Maximum canopy storage (mm)	0	5

The prefixes “v” and “r” indicate “replace” and “relative” change compared to the original values.

In this study, the optimal solution achieved from 2000 model runs was used as an example for demonstrating (1) how spatiotemporal and temporospatial model performance evaluation could be conducted, and (2) the benefits of conducting both spatiotemporal and temporospatial model evaluation (Section 3.1). The following multi-objective function (OF) was used for selecting the optimal solution:

$$OF = \text{minimize}(T_S + S_T), \quad (2)$$

$$S_T = \frac{\sum_{k=1}^{ntimesteps} T_{ETa}^k}{ntimesteps}, \quad (3)$$

$$T_S = \frac{\sum_{j=1}^{npixels} S_{ETa}^j}{npixels}, \quad (4)$$

where  $T_S$  and  $S_T$  are the spatiotemporal and temporospatial model performance indices, respectively,  $T_{ETa}^j$  and  $S_{ETa}^k$  are the statistical indices for temporal and spatial model performance, respectively, for actual evapotranspiration (ETa) at the pixel  $j$  and time step  $k$ ,  $npixels$  is the number of pixels, and  $k$  is the number of evaluation time steps. For this analysis, we selected a MOD16 A2 ETa product as reference data with  $S_{ETa}$  is SPAEF (Equation (6)) and  $T_{ETa}$  is NSE (Equation (7)) as an example in this case (Section 3.1).

To evaluate the relation between spatiotemporal and temporospatial model performance, we used different spatiotemporal ( $S_T$ ) and temporospatial ( $T_S$ ) performance indices (Table 2). For this evaluation, different ETa products (MOD16 A2, SEEBop, and TerraClimate) were used to draw reliable conclusions. The relation between spatiotemporal and temporospatial model performance was analyzed based on 2000 model runs (Section 3.2).

**Table 2.** List of spatial and temporal model performance statistics used in this study.

Spatiotemporal Statistic $S_T$ (Equation (3))		Temporospatial Statistic $T_S$ (Equation (4))	
Notation	Range and Ideal Value	Notation	Range and Ideal Value
$S_{-NSE}$	$[-1, \infty)$	$T_{-SPAEF}$	$[-1, \infty)$
$S_{-NSE}$	$[-1, \infty)$	$T_{-NSE}$	$[-1, \infty)$
$S_{RMSE}$	$[0, \infty)$	$T_{RMSE}$	$[0, \infty)$
$S_{RSR}$	$[0, \infty)$	$T_{RSR}$	$[0, \infty)$
$S_{aBIAS}$	$[0, \infty)$	$T_{aBIAS}$	$[0, \infty)$

Numbers in bold indicate the ideal value.

The spatiotemporal ( $S_T$ ) and temporospatial ( $T_S$ ) performance statistics (Table 2) are calculated using Equations (3) and (4), respectively, while the temporal T (e.g., SPAEF, NSE, RMSE, RSE, absolute bias aBIAS) and spatial S (e.g., KGE, NSE, RMSE, RSR, aBIAS) performance are calculated as follows:

$$KGE = 1 - \sqrt{(r-1)^2 + (\alpha-1)^2 + (\beta-1)^2}, \quad (5)$$

$$SPAEF = 1 - \sqrt{(r-1)^2 + (\alpha/\beta-1)^2 + (\gamma-1)^2}, \quad (6)$$

$$NSE = 1 - \frac{\sum_{i=1}^n (x_i^{sim} - x_i^{obs})^2}{\sum_{i=1}^n (x_i^{obs} - \bar{x}^{obs})^2}, \quad (7)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_i^{sim} - x_i^{obs})^2}{n}}, \quad (8)$$

$$RSR = \frac{\sqrt{\sum_{i=1}^n (x_i^{sim} - x_i^{obs})^2}}{\sqrt{\sum_{i=1}^n (x_i^{obs} - \bar{x}^{obs})^2}}, \quad (9)$$

$$aBIAS = |BIAS| = \left| 1 - \frac{\sum_{i=1}^n x_i^{sim}}{\sum_{i=1}^n x_i^{obs}} \right|, \quad (10)$$

with:

$$r = \frac{\sum_{i=1}^n (x_i^{obs} - \bar{x}^{obs}) \cdot (x_i^{sim} - \bar{x}^{sim})}{\sqrt{\sum_{i=1}^n (x_i^{obs} - \bar{x}^{obs})^2 \cdot \sum_{i=1}^n (x_i^{sim} - \bar{x}^{sim})^2}}, \quad (11)$$

$$\alpha = \frac{\sqrt{\sum_{i=1}^n (x_i^{sim} - \bar{x}^{sim})^2}}{\sqrt{\sum_{i=1}^n (x_i^{obs} - \bar{x}^{obs})^2}}, \quad (12)$$

$$\beta = \frac{\bar{x}^{sim}}{\bar{x}^{obs}}, \quad (13)$$

$$\gamma = \frac{\sum_{j=1}^m \min(K_j, L_j)}{\sum_{j=1}^m K_j}, \quad (14)$$

where  $r$  is the linear correlation between observed and simulated values (which in this study are the simulated and satellite-derived ETa),  $\alpha$  is the variability measure,  $\beta$  is the bias term,  $\gamma$  is the histogram intersection,  $x_j^{sim}$  and  $x_j^{obs}$  are the simulated and observed values,  $\bar{x}^{obs}$  is the mean of observed  $x$ ,  $K$  and  $L$  are the histograms of observed and simulated variables, respectively,  $m$  is the number of bins, and  $|\cdot|$  is the absolute sign.

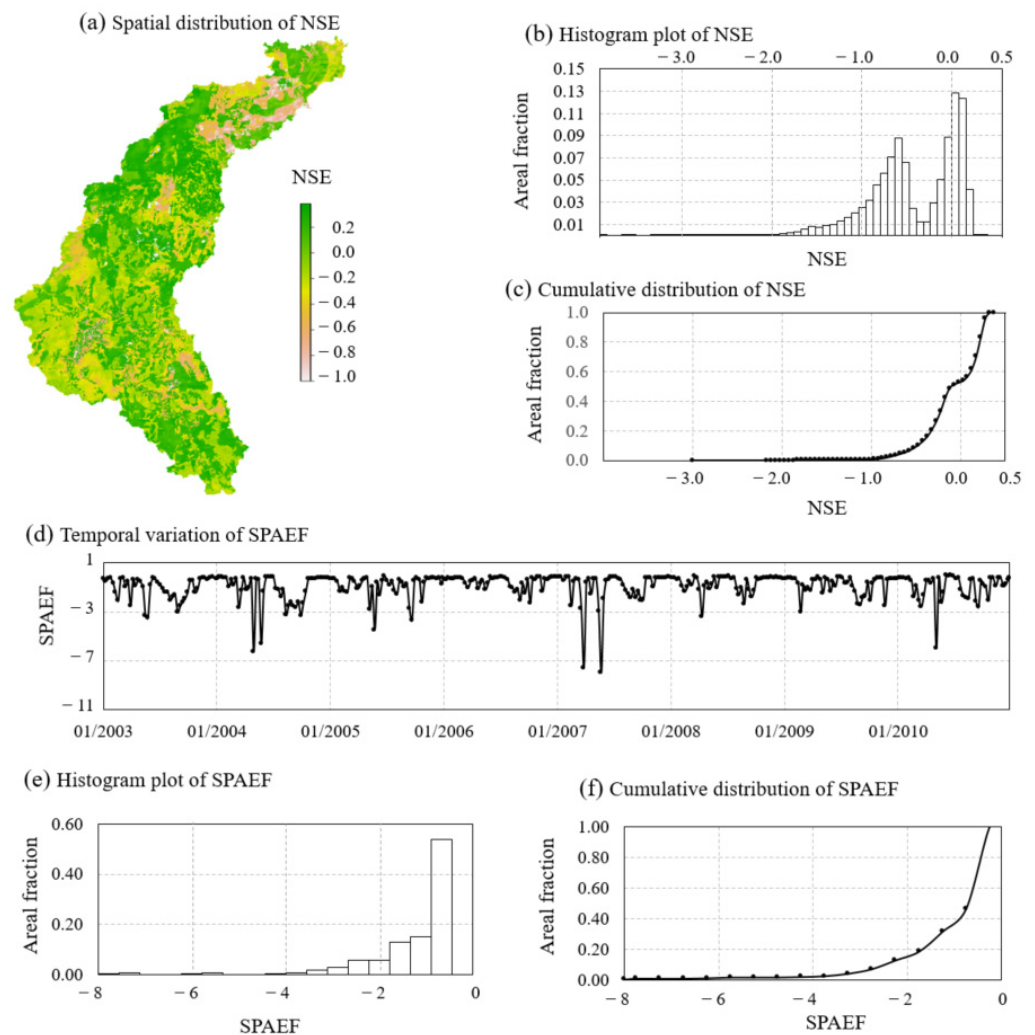
### 3. Results and Discussion

#### 3.1. Spatiotemporal and Temporospatial Model Performance

Figure 4 shows the best spatiotemporal and temporospatial model performance for ETa (with MOD16 A2 ETa as the reference dataset). It is seen that the spatial distribution of the NSE is far from homogeneous (Figure 4a), which could not be seen in a spatially lumped approach. From the spatial distribution of NSE, the areas with a relatively poor temporal model performance compared to others could be easily identified. This information could be used to (1) further improve the model performance and (2) detect the reason for poor temporal model performance in these areas. The histogram and cumulative plots provide useful statistical information about the temporal model performance (Figure 4b,c). The histogram (or cumulative) plot shows the fraction of the study area that has the NSE within a specific range (or below a specific value). The histogram plot is seen to be highly skewed toward the optimal NSE value (optimal NSE = 1). This is the desired model performance, in other words, a higher skewness of the NSE distribution toward the optimal point indicates a better model performance.

Despite the NSE being a temporal performance index, it does not provide information about the time when the model poorly performs. However, this can only be seen in the temporal variation of the SPAEF index (Figure 3d). For example, during April 2004 and April–May 2007, the simulated patterns are poorly matched with the reference data while in other periods the match is far better. The distribution of SPAEF indices is also highly skewed toward the optimal value as about 40% of the time step of the SPAEF indices are less than  $-1$  (Figure 4e,f).

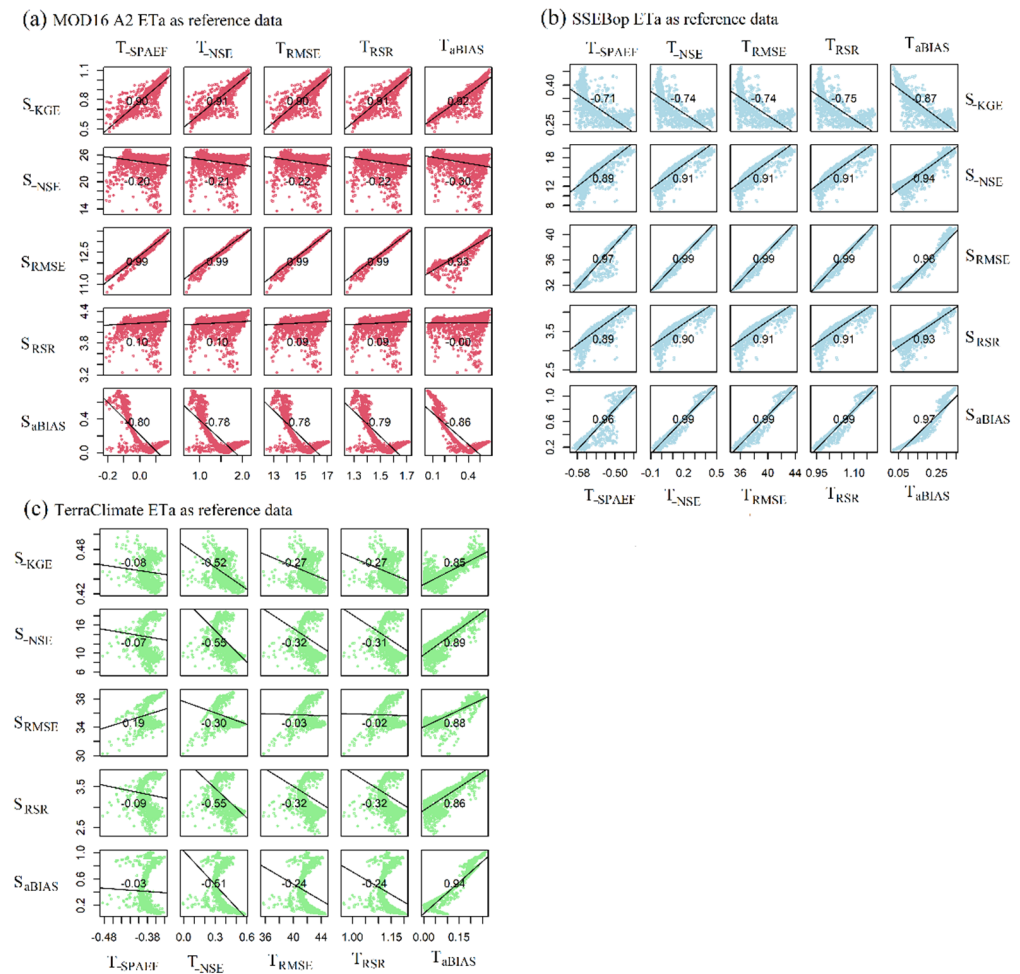




**Figure 4.** Spatiotemporal and temporospatial and model performance represented by (a) the spatial distribution of the NSE and (b,c) the histogram and cumulative plots of NSE, (d) the temporal variation of SPAEF, (e,f) the histogram and cumulative plots of SPAEF. The ideal value of NSE and SPAEF is 1.

### 3.2. The Relation between Spatiotemporal and Temporospatial Model Performance

Figure 5 shows that the relation between spatiotemporal ( $S_T$ ) and temporospatial ( $T_S$ ) model performance varies depending on the reference ETa data and the statistical indices used for evaluation. For example, when KGE and SPAEF were used to derive the spatiotemporal ( $S_{KGE}$ ) and temporospatial ( $T_{SPAEF}$ ) mode of performance statistics,  $S_{KGE}$  and  $T_{SPAEF}$  could be highly positively correlated (Figure 5a), highly negatively correlated (Figure 5b), or uncorrelated (Figure 5c). A high positive correlation between  $S_{KGE}$  and  $T_{SPAEF}$  indicates that an increase in the spatiotemporal model performance will likely lead to an increase in temporospatial model performance. However, a high negative correlation between  $S_{KGE}$  and  $T_{SPAEF}$  means that an increase in spatiotemporal model performance will likely result in a decrease in temporospatial model performance and vice versa. An uncorrelated  $S_{KGE}$  and  $T_{SPAEF}$  indicates that the spatiotemporal model performance cannot be inferred from the temporospatial model performance and vice versa. The results show that even if the same statistical index was used for spatiotemporal and temporospatial model evaluation (e.g.,  $S_{NSE}$  and  $T_{NSE}$ , Figure 5a) it does not always guarantee that an increase in the spatiotemporal model performance will lead to an increase in the temporospatial model performance (Figure 5a–c). Overall, the results show that both spatiotemporal and temporospatial model performance evaluation is needed.



**Figure 5.** Spatiotemporal and temporospatial model performance evaluation against different ETa products (a) MOD16 A2 ETa, (b) SSBop, and (c) TerraClimate. Results are from 2000 model runs, lines are the regression lines and numbers inside the plot are the correlation values. In all correlation plots, the points located closer to the lower-left corner indicate a better model performance.

#### 4. Conclusions

Distributed hydrological models have long been used for understanding the spatial and temporal hydrological responses of a catchment. To ensure that the model can provide meaningful spatial and temporal information, the simulated spatiotemporal data from these models should be evaluated against the reference data. However, studies have either focused on temporal or spatial model evaluation. In addition, temporal (or spatial) model evaluation is often conducted in a spatially (or temporally) lumped approach. The terms “spatiotemporal” and “temporospatial” model performance evaluations introduced in this study refer to the evaluation of (1) the temporal model performance in a spatially distributed approach and (2) the spatial model performance in a temporally distributed approach, respectively. Here, we demonstrated that both spatiotemporal and temporospatial model performance evaluations are needed. They are indeed two different aspects of model evaluation. The results showed that spatiotemporal (or temporospatial) model performance evaluation could help in detecting the locations (or the time) where (or when) the temporal (or spatial) patterns are poorly represented by the model. Additionally, an increase in the spatiotemporal model performance will not necessarily lead to an increase in the temporospatial model performance and vice versa. This further suggests that both spatiotemporal and temporospatial model performance evaluations are needed to achieve a reliable model performance. Overall, the proposed approach provides graphical and sta-

tistical indicators for a better understanding of spatiotemporal and temporospatial model behavior which will further help in improving the model performance.

**Author Contributions:** Conceptualization, T.V.N.; methodology, T.V.N., B.U. and D.A.T.; software, T.V.N.; formal analysis, T.V.N., D.A.T. and B.U.; writing—original draft preparation, T.V.N., D.A.T., T.B.T.P. and B.U.; writing—review and editing, T.V.N., D.A.T., B.U. and T.B.T.P.; visualization, T.V.N. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** The MOD16 A2, SSEbop, TerraClimate data used for model evaluation are freely available at <https://modis.gsfc.nasa.gov/data/dataproduct/mod16.php> (accessed 15 January 2022), and <https://earlywarning.usgs.gov/fews/product/460> (accessed 15 January 2022), and <http://www.climatologylab.org/terraclimate.html> (accessed 15 January 2022), respectively. The R code and SWAT model used in this study are available upon request via contacting the corresponding author.

**Acknowledgments:** We thank the three anonymous reviewers for their constructive comments that helped to considerably improve the quality of the manuscript.

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

## References

- Klingler, C.; Schulz, K.; Herrnegger, M. Large-Sample Data for Hydrology and Environmental Sciences for Central Europe. *Earth Sci. Data* **2021**, *13*, 4529–4565. [CrossRef]
- Brocca, L.; Tullo, T.; Melone, F.; Moramarco, T.; Morbidelli, R. Catchment Scale Soil Moisture Spatial-Temporal Variability. *J. Hydrol.* **2012**, *422–423*, 63–75. [CrossRef]
- Wilson, D.J.; Western, A.W.; Grayson, R.B. Identifying and Quantifying Sources of Variability in Temporal and Spatial Soil Moisture Observations. *Water Resour. Res.* **2004**, *40*, W02507. [CrossRef]
- Thomas, A. Spatial and Temporal Characteristics of Potential Evapotranspiration Trends over China. *Int. J. Climatol.* **2000**, *20*, 381–396. [CrossRef]
- Edmunds, W.M.; Fellman, E.; Goni, I.B.; Prudhomme, C. Spatial and Temporal Distribution of Groundwater Recharge in Northern Nigeria. *Hydrogeol. J.* **2002**, *10*, 205–215. [CrossRef]
- Neitsch, S.; Arnold, J.; Kiniry, J.; Williams, J. *Soil & Water Assessment Tool Theoretical Documentation Version 2009*; Texas Water Resources Institute: College Station, TX, USA, 2011. Available online: <https://swat.tamu.edu/media/99192/swat2009-theory.pdf> (accessed on 17 April 2022).
- 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]
- Beven, K.J.; Kirkby, M.J. A Physically Based, Variable Contributing Area Model of Basin Hydrology. *Hydrol. Sci. Bull.* **1979**, *24*, 43–69. [CrossRef]
- Liang, X.; Lettenmaier, D.P.; Wood, E.F.; Burges, S.J. A Simple Hydrologically Based Model of Land Surface Water and Energy Fluxes for General Circulation Models. *J. Geophys. Res.* **1994**, *99*, 14415–14428. [CrossRef]
- Kumar, R.; Samaniego, L.; Attinger, S. Implications of Distributed Hydrologic Model Parameterization on Water Fluxes at Multiple Scales and Locations. *Water Resour. Res.* **2013**, *49*, 360–379. [CrossRef]
- Samaniego, L.; Kumar, R.; Attinger, S. Multiscale Parameter Regionalization of a Grid-Based Hydrologic Model at the Mesoscale. *Water Resour. Res.* **2010**, *46*, W05523. [CrossRef]
- Campo, L.; Caparrini, F.; Castelli, F. Use of Multi-Platform, Multi-Temporal Remote-Sensing Data for Calibration of a Distributed Hydrological Model: An Application in the Arno Basin, Italy. *Hydrol. Process.* **2006**, *20*, 2693–2712. [CrossRef]
- Dembélé, M.; Ceperley, N.; Zwart, S.J.; Salvatore, E.; Mariethoz, G.; Schaeffli, B. Potential of Satellite and Reanalysis Evaporation Datasets for Hydrological Modelling under Various Model Calibration Strategies. *Adv. Water Resour.* **2020**, *143*, 103667. [CrossRef]
- Herman, M.R.; Nejadhashemi, A.P.; Abouali, M.; Hernandez-Suarez, J.S.; Daneshvar, F.; Zhang, Z.; Anderson, M.C.; Sadeghi, A.M.; Hain, C.R.; Sharifi, A. Evaluating the Role of Evapotranspiration Remote Sensing Data in Improving Hydrological Modeling Predictability. *J. Hydrol.* **2018**, *556*, 39–49. [CrossRef]
- Immerzeel, W.W.; Droogers, P. Calibration of a Distributed Hydrological Model Based on Satellite Evapotranspiration. *J. Hydrol.* **2008**, *349*, 411–424. [CrossRef]
- Jiang, L.; Wu, H.; Tao, J.; Kimball, J.S.; Alfieri, L.; Chen, X. Satellite-Based Evapotranspiration in Hydrological Model Calibration. *Remote Sens.* **2020**, *12*, 428. [CrossRef]
- Koch, J.; Demirel, M.C.; Stisen, S. The SPATIAL Efficiency Metric (SPAEF): Multiple-Component Evaluation of Spatial Patterns for Optimization of Hydrological Models. *Geosci. Model Dev.* **2018**, *11*, 1873–1886. [CrossRef]
- Mendiguren, G.; Koch, J.; Stisen, S. Spatial Pattern Evaluation of a Calibrated National Hydrological Model—A Remote-Sensing-Based Diagnostic Approach. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 5987–6005. [CrossRef]

19. Odusanya, A.E.; Mehdi, B.; Schürz, C.; Oke, A.O.; Awokola, O.S.; Awomeso, J.A.; Adejuwon, J.O.; Schulz, K. Multi-Site Calibration and Validation of SWAT with Satellite-Based Evapotranspiration in a Data-Sparse Catchment in Southwestern Nigeria. *Hydrol. Earth Syst. Sci.* **2019**, *23*, 1113–1144. [\[CrossRef\]](#)
20. Rajib, A.; Evenson, G.R.; Golden, H.E.; Lane, C.R. Hydrologic Model Predictability Improves with Spatially Explicit Calibration Using Remotely Sensed Evapotranspiration and Biophysical Parameters. *J. Hydrol.* **2018**, *567*, 668–683. [\[CrossRef\]](#)
21. 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. [\[CrossRef\]](#)
22. Stisen, S.; Koch, J.; Sonnenborg, T.O.; Refsgaard, J.C.; Bircher, S.; Ringgaard, R.; Jensen, K.H. Moving beyond Run-off Calibration—Multivariable Optimization of a Surface–Subsurface–Atmosphere Model. *Hydrol. Process.* **2018**, *32*, 2654–2668. [\[CrossRef\]](#)
23. Nguyen, V.T.; Dietrich, J.; Uniyal, B. Modeling Interbasin Groundwater Flow in Karst Areas: Model Development, Application, and Calibration Strategy. *Environ. Model. Softw.* **2020**, *124*, 104606. [\[CrossRef\]](#)
24. Nguyen, T.V.; Kumar, R.; Lutz, S.R.; Musolff, A.; Yang, J.; Fleckenstein, J.H. Modeling Nitrate Export From a Mesoscale Catchment Using StorAge Selection Functions. *Water Resour. Res.* **2021**, *57*, e2020WR028490. [\[CrossRef\]](#)
25. Nash, J.E.; Sutcliffe, J.V. River Flow Forecasting through Conceptual Models Part I—A Discussion of Principles. *J. Hydrol.* **1970**, *10*, 282–290. [\[CrossRef\]](#)
26. Gupta, H.V.; Kling, H.; Yilmaz, K.K.; Martinez, G.F. Decomposition of the Mean Squared Error and NSE Performance Criteria: Implications for Improving Hydrological Modelling. *J. Hydrol.* **2009**, *377*, 80–91. [\[CrossRef\]](#)
27. Roberts, N.M.; Lean, H.W. Scale-Selective Verification of Rainfall Accumulations from High-Resolution Forecasts of Convective Events. *Mon. Weather Rev.* **2008**, *136*, 78–97. [\[CrossRef\]](#)
28. Koch, J.; Jensen, K.H.; Stisen, S. Toward a True Spatial Model Evaluation in Distributed Hydrological Modeling: Kappa Statistics, Fuzzy Theory, and EOF-Analysis Benchmarked by the Human Perception and Evaluated against a Modeling Case Study. *Water Resour. Res.* **2015**, *51*, 1225–1246. [\[CrossRef\]](#)
29. Mu, Q.; Zhao, M.; Running, S.W. MODIS Global Terrestrial Evapotranspiration (ET) Product (MOD16A2/A3); Algorithm Theoretical Basis Document Collection 5. 2013; 66p. Available online: <https://modis-land.gsfc.nasa.gov/pdf/MOD16ATBD.pdf> (accessed on 17 April 2022).
30. Senay, G.B.; Bohms, S.; Singh, R.K.; Gowda, P.H.; Velpuri, N.M.; Alemu, H.; Verdin, J.P. Operational Evapotranspiration Mapping Using Remote Sensing and Weather Datasets: A New Parameterization for the SSEB Approach. *J. Am. Water Resour. Assoc.* **2013**, *49*, 577–591. [\[CrossRef\]](#)
31. Abatzoglou, J.T.; Dobrowski, S.Z.; Parks, S.A.; Hegewisch, K.C. TerraClimate, a High-Resolution Global Dataset of Monthly Climate and Climatic Water Balance from 1958–2015. *Sci. Data* **2018**, *5*, 170191. [\[CrossRef\]](#)
32. Uniyal, B.; Dietrich, J.; Vu, N.Q.; Jha, M.K.; Arumí, J.L. Simulation of Regional Irrigation Requirement with SWAT in Different Agro-Climatic Zones Driven by Observed Climate and Two Reanalysis Datasets. *Sci. Total Environ.* **2019**, *649*, 846–865. [\[CrossRef\]](#)
33. Fick, S.E.; Hijmans, R.J. WorldClim 2: New 1-Km Spatial Resolution Climate Surfaces for Global Land Areas. *Int. J. Climatol.* **2017**, *37*, 4302–4315. [\[CrossRef\]](#)
34. Hijmans, R.J.; Cameron, S.E.; Parra, J.L.; Jones, P.G.; Jarvis, A. Very High Resolution Interpolated Climate Surfaces for Global Land Areas. *Int. J. Climatol.* **2005**, *25*, 1965–1978. [\[CrossRef\]](#)
35. Harris, I.; Jones, P.D.; Osborn, T.J.; Lister, D.H. Updated High-Resolution Grids of Monthly Climatic Observations—The CRU TS3.10 Dataset. *Int. J. Climatol.* **2014**, *34*, 623–642. [\[CrossRef\]](#)
36. Kobayashi, S.; Ota, Y.; Harada, Y.; Ebata, A.; Moriya, M.; Onoda, H.; Onogi, K.; Kamahori, H.; Kobayashi, C.; Endo, H.; et al. The JRA-55 Reanalysis: General Specifications and Basic Characteristics. *J. Meteorol. Soc. Jpn.* **2015**, *93*, 5–48. [\[CrossRef\]](#)
37. Wang-Erlandsson, L.; Bastiaanssen, W.G.M.; Gao, H.; Jägermeyr, J.; Senay, G.B.; Van Dijk, A.I.J.M.; Guerschman, J.P.; Keys, P.W.; Gordon, L.J.; Savenije, H.H.G. Global Root Zone Storage Capacity from Satellite-Based Evaporation. *Hydrol. Earth Syst. Sci.* **2016**, *20*, 1459–1481. [\[CrossRef\]](#)
38. Arnold, J.G.; Fohrer, N. SWAT2000: Current Capabilities and Research Opportunities in Applied Watershed Modelling. *Hydrol. Process.* **2005**, *19*, 563–572. [\[CrossRef\]](#)
39. Kim, N.W.; Chung, I.M.; Won, Y.S.; Arnold, J.G. Development and Application of the Integrated SWAT-MODFLOW Model. *J. Hydrol.* **2008**, *356*, 1–16. [\[CrossRef\]](#)
40. Nguyen, V.T.; Dietrich, J. Modification of the SWAT Model to Simulate Regional Groundwater Flow Using a Multicell Aquifer. *Hydrol. Process.* **2018**, *32*, 939–953. [\[CrossRef\]](#)
41. Rafiei, V.; Ghahramani, A.; An-Vo, D.A.; Mushtaq, S. Modelling Hydrological Processes and Identifying Soil Erosion Sources in a Tropical Catchment of the Great Barrier Reef Using SWAT. *Water* **2020**, *12*, 2179. [\[CrossRef\]](#)
42. Monteith, J.L. Evaporation and Environment. *Symp. Soc. Exp. Biol.* **1965**, *19*, 205–234.
43. Allen, R.G. A Penman for All Seasons. *J. Irrig. Drain. Eng.* **1986**, *112*, 348–368. [\[CrossRef\]](#)
44. Allen, R.G.; Jensen, M.E.; Wright, J.L.; Burman, R.D. Operational Estimates of Reference Evapotranspiration. *Agron. J.* **1989**, *81*, 650–662. [\[CrossRef\]](#)
45. Priestley, C.H.B.; Taylor, R.J. On the Assessment of Surface Heat Flux and Evaporation Using Large-Scale Parameters. *Mon. Weather Rev.* **1972**, *100*, 81–92. [\[CrossRef\]](#)
46. Hargreaves, G.H.; Samani, Z.A. Reference Crop Evapotranspiration from Temperature. *Appl. Eng. Agric.* **1985**, *1*, 96–99. [\[CrossRef\]](#)
47. Howell, T.; Evett, S.R. *The Penman-Monteith Method*; USDA Agricultural Research Service: Bushland, TX, USA, 2001.

- 
48. Abiodun, O.O.; Guan, H.; Post, V.E.A.; Batelaan, O. Comparison of MODIS and SWAT Evapotranspiration over a Complex Terrain at Different Spatial Scales. *Hydrol. Earth Syst. Sci.* **2018**, *22*, 2775–2794. [[CrossRef](#)]
  49. Kannan, N.; White, S.M.; Worrall, F.; Whelan, M.J. Sensitivity Analysis and Identification of the Best Evapotranspiration and Runoff Options for Hydrological Modelling in SWAT-2000. *J. Hydrol.* **2007**, *332*, 456–466. [[CrossRef](#)]