remote sensing

MDPI

Article

Estimates of the Land Surface Hydrology from the Community Land Model Version 5 (CLM5) with Three Meteorological Forcing Datasets over China

Dayang Wang [1], Dagang Wang [2,*], Yiwen Mei [2], Qing Yang [3], Mingfei Ji [1], Yuying Li [1], Shaobo Liu [1], Bailian Li [4], Ya Huang [5] and Chongxun Mo [6]

[1] Overseas Expertise Introduction Center for Discipline Innovation of Watershed Ecological Security in the Water Source Area of the Middle Route of South-to-North Water Diversion, School of Water Resource and Environmental Engineering, Nanyang Normal University, Nanyang 473061, China; wangdy58@mail2.sysu.edu.cn (D.W.); jimfdy@gmail.com (M.J.)
[2] School of Geography and Planning, Sun Yat-sen University, Guangzhou 510006, China; meiyw3@mail.sysu.edu.cn
[3] School of Freshwater Sciences, University of Wisconsin-Milwaukee, Milwaukee, WI 53201, USA; yang2473@uwm.edu
[4] International Center for Ecology and Sustainability, University of California, Riverside, CA 93106, USA
[5] State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing 211098, China; hygccw@hhu.edu.cn
[6] College of Civil Engineering and Architecture, Guangxi University, Nanning 530004, China
* Correspondence: wangdag@mail.sysu.edu.cn

Abstract: The land surface model (LSM) is extensively utilized to simulate terrestrial processes between land surface and atmosphere in the Earth system. Hydrology simulation is the key component of the model, which can directly reflect the capability of LSM. In this study, three offline LSM simulations were conducted over China using the Community Land Model version 5.0 (CLM5) driven by different meteorological forcing datasets, namely China Meteorological Forcing Dataset (CMFD), Global Soil Wetness Project Phase 3 (GSWP3), and bias-adjusted ERA5 reanalysis (WFDE5), respectively. Both gridded and in situ reference data, including evapotranspiration (ET), soil moisture (SM), and runoff, were employed to evaluate the performance levels of three CLM5-based simulations across China and its ten basins. In general, all simulations realistically replicate the magnitudes, spatial patterns, and seasonal cycles of ET over China when compared with remote-sensing-based ET observations. Among ten basins, Yellow River Basin (YRB) is the basin where simulations are the best, supported by the higher KGE value of 0.79. However, substantial biases occur in Northwest Rivers Basin (NWRB) with significant overestimation for CMFD and WFDE5 and underestimation for GSWP3. In addition, both grid-based or site-based evaluations of SM indicate that systematic wet biases exist in all three CLM5 simulations for shallower soil layer over nine basins of China. Comparatively, the performance levels in simulating SM for deeper soil layer are slightly better. Moreover, all three types of CLM5 simulate reasonable runoff spatial patterns, among which CMFD can capture more detailed information, but GSWP3 presents more comparable change trends of runoff when compared to the reference data. In summary, this study explored the capacity of CLM5 driven by different meteorological forcing data, and the assessment results may provide important insights for the future developments and applications of LSM.

Keywords: community land model; meteorological forcing; evapotranspiration; soil moisture; runoff; China

Citation: Wang, D.; Wang, D.; Mei, Y.; Yang, Q.; Ji, M.; Li, Y.; Liu, S.; Li, B.; Huang, Y.; Mo, C. Estimates of the Land Surface Hydrology from the Community Land Model Version 5 (CLM5) with Three Meteorological Forcing Datasets over China. *Remote Sens.* **2024**, *16*, 550. https://doi.org/10.3390/rs16030550

Academic Editor: Hatim Sharif

Received: 14 December 2023
Revised: 22 January 2024
Accepted: 28 January 2024
Published: 31 January 2024

1. Introduction

Land surface process plays a vital role in connecting the water cycle and energy process at the interface between land and atmosphere [1–3]. For instance, evapotranspiration (ET)

ensures a continuous water vapor supply to the atmosphere and affects the local climate through precipitation [4–6]. Soil moisture (SM) controls the partitioning of net radiation into sensible heat flux and latent heat flux at the soil surface [7,8]. Runoff collects residual water from the precipitation and ET, maintaining the water balance on Earth and providing fresh water for agricultural irrigation, industrial processes, and domestic consumption [4,9–11]. Consequently, acquiring more precise terrestrial information is essential for gaining a deeper understanding of land surface process and evolution of terrestrial ecosystems.

In recent decades, numerous measurements and their associated approaches have been proposed and developed to estimate the terrestrial variables. Therefore, various land surface data products have been constructed and utilized in land surface and climate studies [12]. For example, in situ measurements can provide the most accurate terrestrial elements at point and local scales. However, directly measuring hydrological variables on a larger scale is unrealistic due to the limited instrument coverage and highly heterogeneous land surface condition [13,14]. Satellite-based remote sensing retrievals can be used to derive hydrology variables on a larger spatial scale, but they only offer static estimations instead of depicting the dynamic process. Additionally, the accuracy of data heavily relies on retrieval algorithms and the quality of other relevant observations [15,16]. The emergence of land surface model (LSM) provides powerful help towards the purpose. LSM not only generates continuous spatiotemporal terrestrial data but also describes the dynamic variations of the physical process. As a crucial component of the Earth system model (ESM), LSM is a promising tool for simulating and understanding the interactions between the land surface and atmosphere in the Earth system [17,18]. LMS is a comprehensive numerical model including various biophysical and biochemical processes, among which hydrological process is the most fundamental and key component [19–21]. The performance of hydrology simulation can directly determine the ability of LSM to a certain extent. LSM has been widely used to provide estimates of terrestrial hydrological elements on regional and global scales [22–25].

The Community Land Surface Model (CLM) is one of the most popular LSMs developed and maintained by National Center for Atmospheric Research (NCAR). Serving as the land component of the Community Earth System Model (CESM), CLM has been well-test and broadly utilized in hydrological and ecological processes studies [17,26]. CLM5, the latest version of CLM, has undergone significant upgrades in various aspects, including terrestrial water, energy, carbon, and nitrogen cycle, based on its predecessors. CLM5 demonstrates improved performance compared to earlier versions such as CLM4.5 and CLM4.0, particularly in key hydrological processes, as evidenced by validation studies. For example, Lawrence et al. [17] reported general improvements and modification from CLM4.5 to CLM5, revealing the superiority of CLM5 in multiple aspects. Song et al. [27] modeled land surface processes in a mountainous rainforest in Costa Rica and found that CLM5 alleviates some errors existing in CLM4.5 through more reliable parameterization, when compared to observed canopy flux and micrometeorological data. Cheng et al. [28] documented that CLM5 exhibits improved ability in capturing terrestrial biogeochemical dynamic over the Contiguous United States (CONUS) due to the parametric and structural updates.

Despite great efforts to improve the performance of CLM over the years, substantial model biases still remain [14,29–31]. These biases are primarily influenced by uncertainties in atmospheric forcing and land surface parameterization. Meteorological forcing serves as the upper boundary condition for CLM and directly impacts the simulations of the energy and water cycles in CLM. Recent advancements in meteorological measurements and associated data fusion methods facilitate generating various atmospheric forcing products. For instance, the Global Soil Wetness Project (GSWP) produced the first global continuous gridded meteorological forcing dataset, with GSWP version 3 encompassing a century-long comprehensive and extensive set of data [23]. Recently, Cucchi et al. [32] developed the WFDE5 forcing data using the WATCH Forcing Data (WFD) methodology applied to surface meteorological variables from the ERA5 reanalysis. Validation results

demonstrated that WFDE5 exhibits lower mean absolute error and higher correlation compared to previous reanalysis data for all elements. Wang et al. [33] found that global reanalysis-based atmospheric forcings produced excessively low SM values in northwestern China and excessively high SM values in northeastern China. By contrast, these biases can be reduced by employing meteorological forcing obtained through the merging of regional station-based and remote-sensed observations [33]. Therefore, significant efforts have been devoted to developing regional meteorological forcing datasets using diverse merging methodologies [34]. In CONUS, the North American Land Data Assimilation System (NLDAS) provides a regional high-resolution spatiotemporal atmospheric dataset that has been widely utilized to drive a hydrological model and land surface model [30,35]. In China, Shi et al. [36] developed the China Meteorological Administration Land Data Assimilation System (CLDAS) atmospheric forcing dataset, which was used to drive Noah-MP LSM over China [34]. The results indicated that CLDAS helps in improving Noah-MP-based simulations in most areas over China and its eight river basins [34]. Additionally, He et al. [37] established the China Meteorological Forcing Dataset (CMFD), which is a high-spatial–temporal resolution gridded near-surface meteorological dataset for studies of land surface processes in China. The dataset was constructed through the fusion of satellite-based remote sensing products, reanalysis datasets and in situ observation data, and has proven to be reliable for LSM simulation in China [38–40].

The above-mentioned meteorological forcing datasets have their own merits, and several previous studies reported the simulation results of LSM driven by different atmospheric forcing datasets. Wang et al. [15] conducted CLM4.5 simulations using four atmospheric forcing datasets globally, three of which include precipitation adjusted by Global Precipitation Climatology Project (GPCP) monthly product. The estimations were validated and evaluated using observation-based datasets. Liu et al. [41] explored the impact of the CMFD and PRIN (the global forcing dataset developed by Princeton University) on two versions of CLM (CLM3.5 and CLM4.5) over China, and found that the pair of CMFD + CLM4.5 outperformed other paired simulation experiments. Ma and Wang [42] revealed the capability of CLM5 forced by CMFD in estimating ET, SM, and snow cover fraction across mainland China using site-based observations. These studies provided valuable references for the meteorological forcing development and the associated model application. However, these modeling assessments either focused solely on previous versions of CLM (e.g., CLM3.5 and CLM4.5) or CLM5 driven by single meteorological forcing dataset. Systematical estimations of CLM5 forcing using different atmospheric forcing datasets have not been conducted. Furthermore, some CLM-modeled hydrological variables, such as runoff, have not been thoroughly investigated due to a lack of gridded reference data.

In the current study, we conducted three types of model simulation experiments over China, in which CLM5 was driven by CMFD, GSWP3, and WFDE5, respectively. We then employed various grid-based estimations and station-based observations as the reference data to evaluate the performance in ET, SM, and runoff simulations from different CLM5 configurations. The paper is structured as follows: Section 2 provides a brief description of the land surface model and forcing data. Section 3 introduces the reference data and methods used in this study. The model evaluation results are presented in Section 4. Section 5 discusses some issues, and Section 6 provides the conclusions of this study.

2. Land Surface Model and Forcing Data

2.1. CLM5

The development of CLM has been a time-consuming and ongoing effort over the past few decades. CLM5 is the latest version, which incorporates significant updates from CLM4.5, particularly in land surface hydrology, plant hydraulics, and snow parameterization schemes [17]. CLM5 has been adopted as the land component of a subset of ESMs participating in phase 6 of the Coupled Model Intercomparison Project (CMIP6), due to its outstanding ability to simulate global energy, water, and carbon budgets. In CLM5, in

order to accurately depict high heterogeneity of land surface, each grid cell is classified into multiple sub-grid land units, and each land unit consists of multiple snow/soil columns occupied with diverse plant functional types (PFTs). Land surface processes for each land unit, column, and PFT are simulated independently, and each sub-grid unit retains its own prognostic variables [17]. CLM5 runs based on specific surface data derived from multiple sources. The leaf area index (LAI) for different PFTs was obtained from Moderate-Resolution Imaging Spectroradiometer (MODIS) satellite data. Canopy top and bottom heights were retrieved from Bonan [43]. Soil texture was calculated using the International Geosphere-Biosphere Programme (IGBP) soil dataset, which includes approximately 5000 soil mapping units. More detailed information can be found in Lawrence et al. [17]. In this study, CLM5 in its offline model is used.

2.2. Meteorological Forcing Data

Three different meteorological forcings are used in this study: CMFD, GSWP3, and WFDE5. The CMFD was developed through fusion of numerous remote sensing products, reanalysis datasets and in situ station observations [37]. The CMFD covers seven near-surface meteorological elements at a temporal resolution of 3 h and a spatial resolution of 0.1° × 0.1° for China from 1979 to 2018. Due to its high quality of continuous spatial–temporal coverage, CMFD is widely used in hydrological models and LSMs [31,41,44].

The GSWP3 is a global gridded meteorological forcing dataset produced by the Institute of Industrial Science, University of Tokyo. It is based on the 20th Century Reanalysis version 2, which was performed with National Centers for Environmental Prediction (NCEP) model. The GSWP3 forcing dataset is a 3-hourly 0.5° forcing product covering the period of 1901–2014. It serves as the default forcing dataset for CLM5 in offline mode [17,45]. Temperature, precipitation, wind, shortwave radiation and longwave radiation within GSWP3 were bias-corrected using several other datasets, including Climate Research Unit (CRU), Global Precipitation Climatology Centre (GPCC), and Surface Radiation Budget datasets. GSWP3 has been broadly utilized in the hydro-energy-eco processes simulation and investigation for a longer time span [46,47].

The WFDE5 provides bias-corrected reconstruction of near-surface variables. It is designed to serve as a meteorological forcing dataset for land surface and hydrological models developed by the European Centre for Medium-Range Weather Forecasts (ECMWF). It was generated using the WFD methodology applied to ERA5 reanalysis dataset. The data were adjusted based on an elevation correction method and corrected using a series of data including CRU, CPCC and precipitation gauge observations. The WFDE5 dataset exhibits lower mean absolute error and higher correlation than its predecessors through a series of evaluation against globally distributed site-based observations for all elements [32]. The WFDE5 was provided at the hourly time step and the 0.5° spatial resolution spanning from 1979 to 2019.

2.3. Experimental Setup

In this study, we carried out three types of CLM5 simulation experiments, in which CMFD, GSWP3 and WFDE5 atmospheric forcing datasets were used to drive CLM5 in its offline mode, and they were referred as CMFD, GSWP3 and WFDE5, respectively. Considering that the inconsistences in the spatial resolution can introduce uncertainties, we upscaled CMFD data to 0.5° × 0.5° grid cells using a simple arithmetic average to match GSWP3 and WFDE5. Consequently, all three CLM5 simulations were ran at a spatial resolution of 0.5° × 0.5°. The model integration time step was set to 1800 s, and the monthly outputs were recorded and used for evaluation. To achieve an equilibrium initial condition and steady state, CLM5 simulations were performed in the overlap period of three forcing datasets (1980–2014) with the first 5-year (1980–1984) run disregarded as spinup. Therefore, only the simulation results from 1985 to 2014 were used in the following comparisons, analyses and evaluations.

In this study, CLM5 simulations were conducted using the prescribed satellite phenology (SP) mode, in which the vegetation phenology parameters are from globally available satellite data in the default CLM5 surface datasets. In SP mode, the vegetation states (e.g., canopy bottom heights, top heights, LAI) are derived through daily linear interpolation between adjacent monthly values. More detailed information can be found in Lawrence et al. [48].

3. Data and Method

3.1. Validation Data

In this study, a suite of comprehensive datasets, including remote-sensing-based observations, in situ measurements and reanalysis datasets over China were utilized to validate the CLM5 simulations. Specifically, we evaluated the impact of meteorological forcing datasets on ET, SM and runoff. For the evaluation at different regional scales, we divided mainland China into ten basins. They are Songhua River Basin (SHRB), Liao River Basin (LRB), Hai River Basin (HaiRB), Huai River Basin (HuaiRB), Yellow River Basin (YRB), Yangtze River Basin (YZRB), Pearl River Basin (PRB), Southeast Rivers Basin (SERB), Southwest Rivers Basin (SWRB), and Northwest Rivers Basin (NWRB). The spatial distribution of these basins is provided in Figure 1.

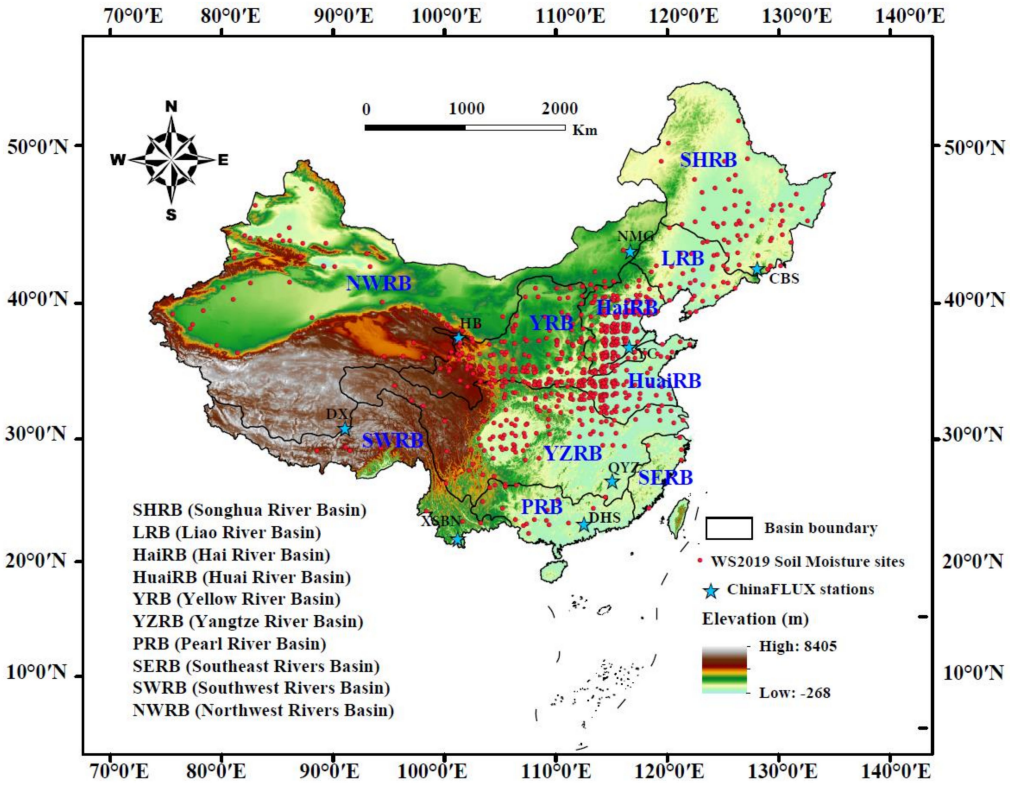

Figure 1. The ten basins map and the locations of ET and SM observation sites. The river basins considered in this study are Songhua River Basin (SHRB), Liao River Basin (LRB), Hai River Basin (HaiRB), Huai River Basin (HuaiRB), Yellow River Basin (YRB), Yangtze River Basin (YZRB), Pearl River Basin (PRB), Southeast Rivers Basin (SERB), Southwest Rivers Basin (SWRB) and Northwest Rivers Basin (NWRB).

3.1.1. ET

a. GLEAM ET

Global Land Evaporation Amsterdam Model (GLEAM) evapotranspiration data are used to evaluate the simulated ET in three types of CLM5 simulations. The GLEAM was derived from reanalysis radiation and air temperature, a combination of gauge-based, reanalysis and satellite-based precipitation, as well as satellite-based vegetation optical

depth. It covers a 43-year period from 1980 to 2022. GLEAM maximizes the recovery of information related to ET from the available stack of climatic and environmental observations [49,50]. In GLEAM, the Priestley and Taylor equation was adopted to calculate potential evaporation (PET) based on observations of surface net radiation and near-surface air temperature, which are achieved from remote sensing-based products. PET is converted into actual evaporation by applying a multiplicative evaporative stress factor derived from microwave observations of the vegetation optical depth. Moreover, the extensive usage of microwave remote sensing products in GLEAM ensures the accuracy of ET estimation under diverse weather conditions. Recent evaluations demonstrated that the GLEAM ET product outperforms other remote-sensed ET products in estimating terrestrial evapotranspiration [51–54]. From the perspective of globe, Zhu et al. [52] comprehensively evaluated the accuracy and uncertainty of five remote sensing-based global ET products (e.g., MOD16, SSEBop, GLEAM, AVHRR, and BESS) at multiple scales against in situ observations from 94 worldwide flux tower, and the results show that GLEAM outperforms other products. Yao et al. [53] also reported that the performance of GLEAM is the best among the three actual ET products (GLEAM, NOAH, and CR), with high accuracy and low uncertainty in China. The outstanding performance of GLEAM is probably attributed to its detailed and comprehensive parameterization. As a consequence, GLEAM is commonly used as a benchmark for validating and assessing land surface models and hydrological models [14,28–30]. Currently, GLEAM incorporates the latest data on precipitation, soil moisture, vegetation optical depth, radiation, and temperature. Due to the reliability and applicability, GLEAM ET was selected as the gridded-based ET reference data in this study. The GLEAM datasets are provided at a $0.25^\circ \times 0.25^\circ$ grid cell with daily, monthly, and yearly temporal scales. In this study, GLEAM version 3.6 ET data are adopted to evaluate the performance levels of CLM5 driven by using three forcing datasets.

b. Eddy covariance measurements of ET

Considering the relatively accurate estimation of latent heat flux (ET in energy units) provided by Eddy covariance (EC) measurements at specific sites, the EC-based ET observations of 8 flux towers from the ChinaFLUX were also adopted for validation. ChinaFLUX is a vital observation and research network that utilized ET and chamber methods to quantify the exchanges of water vapor and energy between terrestrial ecosystem and atmosphere. It plays an important role in the network AsiaFlux [55–57]. Monthly observed data from the 8 stations were used, and the heat flux data with unit of W/m^2 were converted to equivalent ET with unit of mm/mon using Equation (1).

$$ET = LE/\lambda \quad (1)$$

where λ is the latent heat of vaporization assumed to be a constant of 2.45 MJ/kg [57]. The spatial distribution of the 8 stations is shown in Figure 1 and the detailed information is listed in Table 1.

Table 1. The information of the eight flux measurement sites from ChinaFLUX used in this study. Precipitation (P), Temperature (T), and Aridity index (AI) denote the mean annual precipitation, temperature, and aridity index, respectively.

Sites	Elevation (m)	Ecosystem Type	P (mm/yr)	T (°C)	AI
HB	3250	Alpine meadow	535	−1.7	1.29
NMG	1189	Temperate steppe	1493	−0.4	2.26
DX	4333	Alpine steppe-meadow	450	6.5	1.73
CBS	738	Temperate mixed forest	713	3.6	0.79
QYZ	102	Subtropical planted coniferous forest	1542	17.9	0.72
DHS	300	South subtropical evergreen broadleaved forest	1956	21.0	0.67
XSBN	750	Tropic seasonal rainforest	1493	21.8	0.80
YC	28	Temperate farmland	582	13.1	1.78

3.1.2. SM

a. ITPLDAS SM

ITPLDAS is a soil moisture dataset for China obtained through assimilating the observed brightness temperature from the Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E) into an LSM developed by Yang et al. [58,59]. The dataset used the automatically calibrating parameters land data assimilation system (ITPLDAS, hereafter ITP). Its precision in estimating volumetric water content is within ±5% compared to ground-measured value, as validated and assessed in the Tibetan Plateau [54,55]. This soil moisture product provides three layers (0–5 cm, 5–20 cm, and 20–100 cm) of soil water content with a unit of m^3/m^3. All data are accessible at 0.25° spatial resolution and daily time scale, spanning from June 2002 to September 2011 over China. In this study, the daily SM data were converted into monthly and annual data to evaluate the performance levels of CLM5-based SM simulations.

b. WS2019 SM

The in situ SM reference data used in this study are obtained from the WS2019 SM product developed by Wang et al. [60]. These in situ SM observations were collected by the Chinese Meteorological Administration (CMA) at over 1000 sites across China from 1991 to 2013. The measurements were taken at five depths (10, 20, 50, 70, and 100 cm) using the gravimetric technique. Wang et al. [60] conducted strict quality control measures to the raw SM measurements by using flied capacity and bulk density values derived from observation-based soil texture derived soil texture datasets. Finally, more than 35,000 volumetric monthly SM records from 732 stations covering five soil layers were obtained. The data span from 1992 to 2013. Due to reliability and convenience, the SM product has been broadly utilized in nature eco-environment and agriculture investigations and assessments [61–63]. Figure 1 displays the geographic locations of these sites.

3.1.3. Runoff

a. CNRD Runoff

In this study, we employed the China Natural Runoff Dataset (CNRD) version 1.0 as the reference data to evaluate the three runoff simulations of CLM5. CNRD is a recently developed, quality-controlled gridded runoff reconstruction dataset developed by Gou et al. [47]. It offers daily, monthly, and annual 0.25° × 0.25° runoff estimates for the period of 1961–2018 over China. The dataset was generated using the Variable Infiltration Capacity (VIC) model trained by incorporating parameter sensitivity analysis, optimization, and regionalization with numerous catchments gauge-based observations. CNRD holds high potential for exploring long-term trends and spatial–temporal variations in water resources across China. Additionally, it contributes to the development and improvement of LSM [47,64]. To date, only the monthly runoff data are available.

b. CWRB Runoff

We also used the China's Water Resources Bulletin (CWRB) runoff data for the period of 2005–2014, collected and archived by the Ministry of Water Resources of China (MWRC), as a runoff reference [11]. The bulletin provides a series of water resources statistics, including precipitation, runoff, water supply, water use and consumption, and eco-environmental water use. These statistics are derived from monitoring facilities located across ten river basins in China. The runoff data in the bulletin are presented in a unit of m^3, which was converted to mm based on the basin area for comparison in this study.

3.2. Evaluation Method

We evaluated CLM5 by comparing simulated ET, runoff, and SM with the reference data. Because of the inconsistency in the spatial resolution between the reference data and CLM simulation outputs, the scale conversion is needed before comparison. For GLEAM, ITP and CNRD products, they all have a finer resolution (0.25°) than CLM simulation results. We transformed them to the same spatial resolution of CLM results (0.5°) using grid box average method. As a result, a grid-to-grid comparison is performed at 0.5° spatial

resolution. Moreover, for in situ observations, we performed a grid-to-site comparison, wherein the grid-based simulation results of CLM5 are directly compared to the site-based observations located in the corresponding grid at the same time scale (e.g., monthly, seasonally, and/or yearly). If there is more than one station located in a grid, the average value of the observations from these stations is adopted for comparison. Since the soil layers of CLM5-simulated SM are not exactly consistent with the measurement layers of the ITP SM and CMA sites, we applied weighted average method according to the soil depths to interpolate the SM into different soil layers (10, 20, 50, 70, and 100 cm) [25]. The interpolation process involved Equation (2), where SM_i denotes the SM after linearly interpolating the simulation results to the ith observation layer; M_j represents the SM simulation in the jth model layer; $T_{j,i}$ represents the soil thickness of the intersection of model soil layer j; and the observation layers i, b, and a denote the starting and ending numbers of the model soil layers that intersect observation layer i.

$$\mathrm{SM}_i = \frac{\sum_{j=b}^{a}\left(M_j \cdot T_{j,i}\right)}{\sum_{j=b}^{a} T_{j,i}} \tag{2}$$

Five metrics, including bias, relative bias (RB), root mean square error (RMSE), correlation coefficient (Corr), and Kling–Gupta efficiency (KGE), were used to evaluate the model performance. The calculations for these metrics are as follows:

$$\mathrm{Bias} = \frac{1}{N}\sum_{i=1}^{N}(S_i - R_i) \tag{3}$$

$$\mathrm{RB} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{S_i - R_i}{R_i}\right) \tag{4}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(S_i - R_i)^2} \tag{5}$$

$$\mathrm{Corr} = \frac{\sum_{i=1}^{N}\left(S_i - \overline{S}\right)\left(R_i - \overline{R}\right)}{\sqrt{\sum_{i=1}^{N}\left(S_i - \overline{S}\right)^2 \sum_{i=1}^{i=N}\left(R_i - \overline{R}\right)^2}} \tag{6}$$

where N represents the total number of grid cells or sites; S_i and R_i represent the simulated and reference values, respectively; the subscript i denotes the sample number at the ith position; and $\overline{S}$ and $\overline{R}$ denote the spatial or temporal mean of the modeled and reference values, respectively.

$$\mathrm{KGE} = 1 - \sqrt{(Corr - 1)(\beta - 1)(\gamma - 1)} \tag{7}$$

where β and γ are defined as:

$$\beta = \frac{\overline{S}}{\overline{R}} \tag{8}$$

$$\gamma = \frac{\sigma_S / \overline{S}}{\sigma_R / \overline{R}} \tag{9}$$

where σ_S and σ_R denote the standard deviations of simulated and reference values, respectively. The KGE is a comprehensive and objective statistical metric that provides diagnostic insights into model performance by merging correlations, biases, and variability. A KGE score larger than 0.5 indicates relatively good performance, the optimal score is 1. Moreover, common trend analysis methods such as Mann–Kendall trend detection and Sen's slope method were also used in this study. Detailed information on these methods can be seen in previous publications [65–67].

4. Results

In the following sections, the intercomparisons of three atmospheric forcing datasets were firstly conducted, and then the corresponding terrestrial simulations, namely CMFD, GSWP3 and WFDE5, were evaluated across different spatial and temporal scales. The evaluation focuses are ET, SM, and runoff.

4.1. Intercomparison of Meteorological Forcing Datasets

Table 2 presents the annual averages of meteorological elements from the three atmospheric forcing datasets during the period of 1985–2014. On a national scale, all three datasets have comparable atmospheric elements averaged values, with the relative difference between the maximum and minimum values for each element being less than 10%. The largest discrepancy lies in air temperature, ranging from 6.44 to 7.07 °C, while the smallest difference is in humidity, with all values converging at 0.006 kg/kg. Moreover, serving as the sources of water and energy supply for terrestrial processes, the spatial distributions of precipitation and solar radiation are further investigated. Figure 2 illustrates the multi-year averaged precipitation obtained from CMFD, GSWP3, and WFDE5. Overall, the average values of precipitation from three forcing datasets are very similar, which are 601, 602, and 596 mm/yr, respectively. Spatially, all of them show consistent spatial patterns, which decrease from the southeastern coastal areas to the northwestern inland of China. The highest annual precipitation in southernmost China could be above 2000 mm/yr, while the lowest precipitation in the northwest of China is below 100 mm/yr. Nevertheless, some discrepancies do exist. For instance, both CMFD and GSWP3 have relatively higher precipitation in southwest China (e.g., Sichuan Province and Yunnan Province), whereas these features are not presented in WFDE5. In addition, WFDE5 shows more smoother spatial transitions from southeastern China (high-value areas) to northwestern China (low-value areas), compared with CMFD and GSWP3. On the basin scale, the highest inconsistency among these three atmospheric forcing datasets can be found in the SWRB, where domain-averaged precipitation amounts are 754, 936, and 863 mm/yr for CMFD, GSWP3, and WFDE5, respectively. By contrast, the highest consistency is located in LRB, with domain-averaged precipitation amounts of 565, 569, and 567 mm/yr for CMFD, GSWP3, and WFDE5, respectively. In terms of solar radiation, CMFD and GSWP3 have nearly identical annual averages, and both are less than that of WFDE5, as displayed in Figure 3. Notably, all three forcing datasets show similar spatial patterns of solar radiation, with the high-value center in western Tibetan Plateau and low-value center in central YZRB (e.g., Sichuan Basin). However, WFDE5 shows larger amplitude of solar radiation variation, compared with CMFD and GSWP3. For instance, the highest solar radiation is located in SWRB, where the domain-averaged value of WFDE5 is 222 W/m^2, which is obviously larger than that of GSWP3 (205 W/m^2) and CMFD (209 W/m^2). These differences among different atmospheric forcing datasets could be propagated to land surface processes by models, directly or indirectly affecting the hydrological process simulations to some extent.

Table 2. Annual averages of meteorological elements from different meteorological forcing datasets for the period of 1985–2014 over China.

Elements	CMFD	GSWP3	WFDE5
Precipitation (mm/yr)	601	602	596
Solar radiation (W/m^2)	179.64	179.71	190.43
Downward longwave radiation (W/m^2)	284.08	287.03	278.98
Air temperature (°C)	6.44	7.07	7.01
Humidity (kg/kg)	0.006	0.006	0.006
Wind speed (m/s)	2.45	2.48	2.65
Surface pressure (kPa)	83.76	83.29	83.26

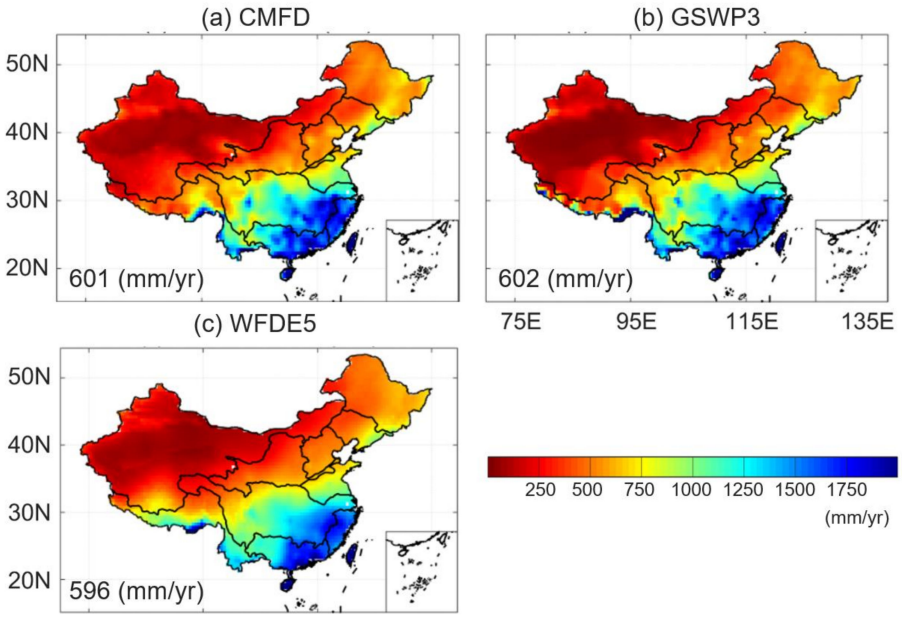

Figure 2. The spatial distribution of multi-year averaged precipitation derived from (**a**) CMFD, (**b**) GSWP3, and (**c**) WFDE5 during the period of 1985–2014. The numbers in the lower left corner indicate the domain average.

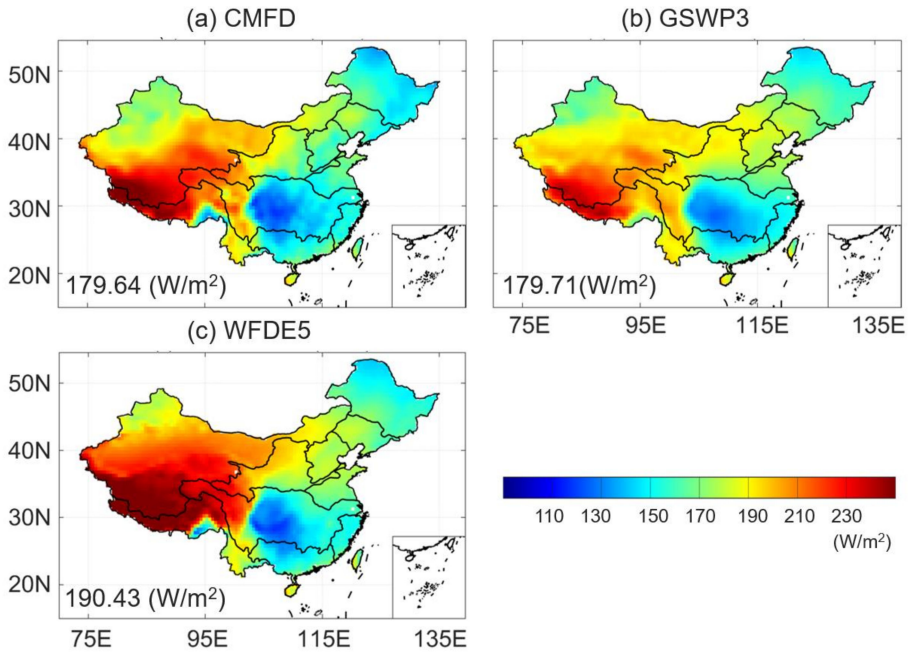

Figure 3. The spatial distribution of multi-year averaged solar radiation derived from (**a**) CMFD, (**b**) GSWP3, and (**c**) WFDE5 during the period of 1985–2014. The numbers in the lower left corner indicate the domain average.

4.2. ET

Figure 4 displays the density scatter plot of the multiyear mean values (1985–2014) of CLM5-based simulated ET (i.e., CMFD, GSWP3, and WFDE5) and GLEAM-based ET. Generally, the scatter points are consistently distributed around the 1:1 line, indicating a strong correlation between the modeled ET and reference ET. The Corr values for CMFD, GSWP3, and WFDE5 are 0.94, 0.92, and 0.92, respectively, and all KGE values exceed 0.85. To investigate the spatial patterns of ET, gridded multiyear average annual ET simulated by CMFD, GSWP3, and WFDE5, and their RBs against GLEAM in China are presented

in Figure 5. Overall, all three ET simulations capture the spatial patterns of ET reasonably well, exhibiting a gradual decrease from southeast to northwest. In southeast China, the highest annual ET values can reach above 800 mm/yr and even exceed 1200 mm/yr in parts of the Hainan province, while in northwest China, the lowest ET values are below 100 mm/yr. In addition, the domain-averaged values of simulated ET are 331, 336, and 343 mm/yr for CMFD, GSWP3, and WFDE5, respectively, which are all close to GLEAM ET values of 344 mm/yr. All three experiments show outstanding performance levels in south, central north, and northeast China, while presenting diverse performance levels in northwest China, as shown in Figure 5e–g. In most areas of northwest China, both CMFD and WFDE5 significantly overestimate ET with the RB values larger than 50%, while GSWP3 underestimates ET with an RB value less than −50%. To investigate the difference among these experiments across different basins of China, the statistical metrics for all the experiments against GLEAM ET over China and its ten basins were calculated and are listed in Table 3. The statistical metrics vary with regions. All three types of simulations perform well in resembling GLEAM ET in most basins. Specifically, the most outstanding performance occurs in the YRB, with RB values of −0.4%, 2.7%, and 0.7%, and KGE values of 0.82, 0.80, and 0.77 for CMFD, GSWP3, and WFDE5, respectively. In contrast, the simulations fail to accurately capture ET spatial pattern in NWRB. The RB values in this region are 55.8%, 5.5%, and 60.1%, with most areas deviating by more than ±50% from GELAM ET. These significant differences can be attributed to two reasons. Firstly, the NWRB comprises diverse climate zones, including arid, semi-arid, semi-humid, and humid climate zones. As a result, the factors affecting ET are considerably more complex than other basins. Secondly, these biases may result from the forcing data. The gridded meteorological elements in the three forcing datasets are either derived from gauge-based observations or corrected by observations. However, the scarcity of station in the NWRB hampers the guarantee of high-quality climate data. Despite overestimation and underestimation existing at the regional scale in all three experiments, both CMFD and GSWP3 show slightly better performance levels than WFDE5 in terms of KGE values over ten basins.

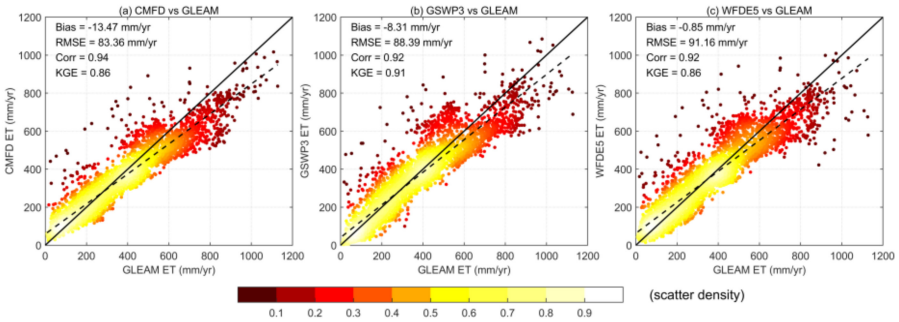

Figure 4. The density scatter plots between GLEAM-based ET and simulated ET forced by (**a**) CMFD, (**b**) GSWP3, and (**c**) WFDE5 across China for the 1985–2014 period. The black solid line is the 1:1 line, and the black dashed line is the linear fit line. The statistical metrics including Bias, RMSE, Corr, and KGE are domain-averaged values shown in the top-left corner.

The seasonal variations in ET over China and its ten basins were also analyzed. As shown in Figure 6a, the ET in summer (June–July–August) is notably higher than in other months due to sufficient precipitation and high temperature, which is captured by all three CLM5-simulated ET and GLEAM ET. Nevertheless, discrepancies still exist between CLM5 simulations and GLEAM estimations in several basins. For example, all three types of simulations overestimate ET during the warm season (May to August) in the LRB, HaiRB and HuaiRB (Figure 6c–e), especially in HuaiRB, the simulated ET surpasses GLEAM ET with an RB of approximately 40%. Conversely, CLM5 underestimates ET in YZRB and PRB (Figure 6g,h) during the cold season (January–February–March) and in SERB throughout the year (Figure 6i) compared to GLEAM data. Overall, the three CLM5-based

simulations show comparable performance levels in the seasonal cycles of ET over most basins, with WFDE5 presenting slightly better results than other experiments, displaying a closer cycle pattern in relation to GLEAM ET.

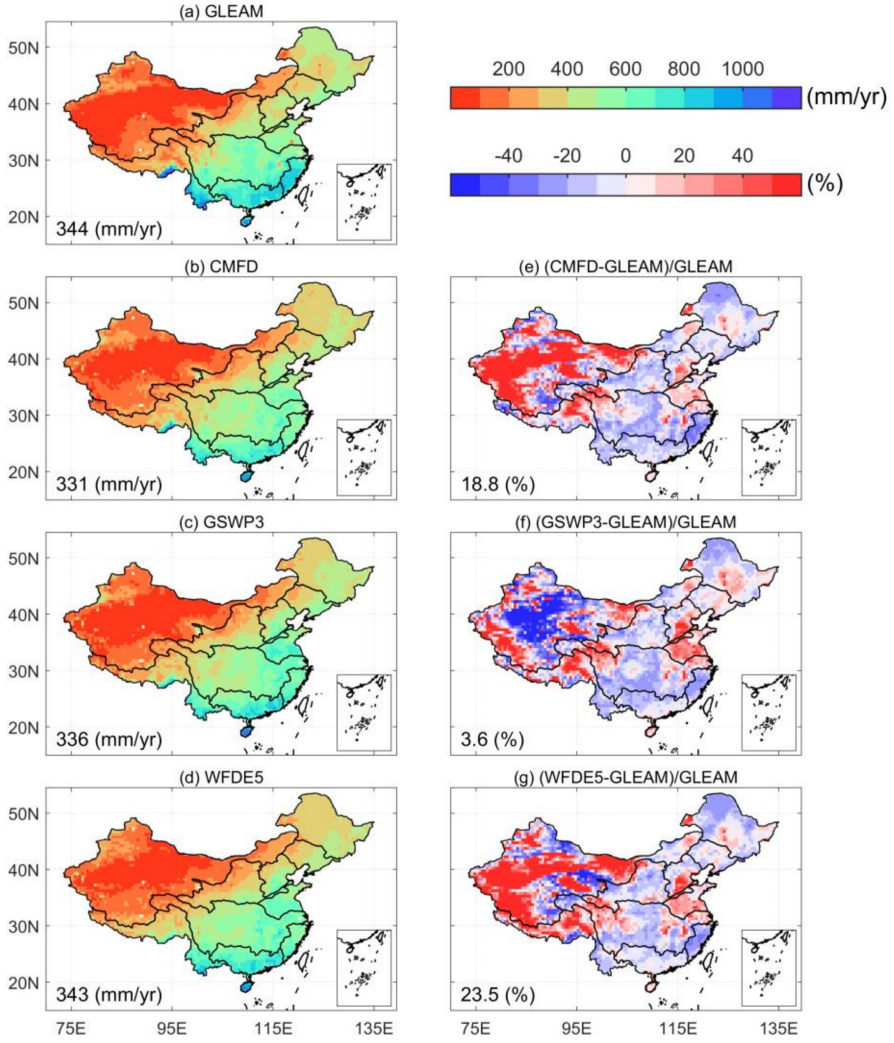

Figure 5. Multi-year averaged ET derived from (**a**) GLEAM, (**b**) CMFD, (**c**) GSWP3, and (**d**) WFDE5, and the RB between (**e**) CMFD and GLEAM, (**f**) GSWP3 and GLEAM, and (**g**) WFDE5 and GLEAM for the period of 1985–2014. The numbers in the lower left corner indicate the domain average.

Table 3. Statistical metrics (Bias, RB, RMSE, Corr, and KGE) of averaged annual ET between simulations (CMFD, GSWP3, and WFDE5) and GLEAM observations over China and its ten basins during the period of 1985–2014.

Regions	Simulations	Bias (mm/yr)	RB (%)	RMSE (mm/yr)	Corr	KGE
China	CMFD	−13.5	18.8	83.4	0.94	0.86
	GSWP3	−8.3	3.6	88.4	0.92	0.91
	WFDE5	−0.8	23.5	91.2	0.92	0.86
SHRB	CMFD	−33.7	−4.2	72.7	0.35	0.29
	GSWP3	−11.9	1.1	70.8	0.36	0.35
	WFDE5	−36.3	−4.8	77.1	0.29	0.24
LRB	CMFD	−24.7	−5.2	50.9	0.82	0.75
	GSWP3	9.2	2.7	44.8	0.85	0.84
	WFDE5	−13.3	−2.2	49.5	0.79	0.75

Table 3. *Cont.*

Regions	Simulations	Bias (mm/yr)	RB (%)	RMSE (mm/yr)	Corr	KGE
HaiRB	CMFD	15.6	4.5	76.3	0.44	0.17
	GSWP3	51.0	13.3	106.9	0.40	−0.05
	WFDE5	34.6	9.4	104.3	0.33	−0.16
HuaiRB	CMFD	51.1	12.7	91.6	0.38	0.36
	GSWP3	133.3	29.5	159.9	0.39	0.34
	WFDE5	93.5	21.3	124.9	0.38	0.35
YRB	CMFD	−10.8	−0.4	60.5	0.82	0.82
	GSWP3	−0.9	2.7	63.7	0.80	0.80
	WFDE5	−6.5	0.7	66.7	0.80	0.77
YZRB	CMFD	−43.3	−5.0	102.4	0.78	0.71
	GSWP3	−36.2	−3.6	103	0.76	0.72
	WFDE5	−17.4	−0.6	100.4	0.75	0.74
PRB	CMFD	−65.2	−7.4	117.1	0.55	0.54
	GSWP3	−49.1	−5.6	120.2	0.58	0.46
	WFDE5	−31.1	−2.6	106.6	0.50	0.50
SERB	CMFD	−192.3	−23.0	220.5	−0.04	−0.12
	GSWP3	−142.7	−16.8	172	0.24	0.14
	WFDE5	−168.6	−20.0	197.3	0.09	0.00
SWRB	CMFD	−24.7	4.2	98.5	0.96	0.79
	GSWP3	−12.1	10.7	117.5	0.94	0.71
	WFDE5	10.0	21.2	137.5	0.91	0.62
NWRB	CMFD	18.6	55.8	56.2	0.77	0.65
	GSWP3	−1.9	5.5	57.8	0.78	0.75
	WFDE5	19.5	60.1	65.2	0.70	0.60

To further quantitatively explore the performance levels in ET simulations of CLM5 driven by three forcing datasets, we also evaluated the ET results using eight stations' observation data from ChinaFLUX. The monthly mean ET time series from simulations and observations are illustrated in Figure 7, and statistical metrics are provided in Table S1. All three CLM5-based simulations capture the fluctuation patterns of monthly averaged ET at all sites, especially at CBS, QYZ, and DHS (Figure 7d–f). The averaged Corr value at these sites can reach up to 0.93, 0.91, and 0.85, and the averaged KGE values can exceed 0.86, 0.81, and 0.75 for CMFD, GSPW3, and WFDE5, respectively (Table S1). However, remarkable underestimation of ET occurs at HB, NMG, and DX (Figure 7a–c), with averaged RB values ranging from −67.8% to −24.8% for all types of simulations. This indicates that CLM5 fails to accurately model the physical process associated with ET in arid areas. In contrast, a systematic overestimation can be found at XSBN, with RB values of 35.8%, 39.0%, and 42.5% for CMFD, GSWP3, and WFDE5, respectively. A possible reason for this overestimation can be uncertainties in the LAI data within the model used in the computation of transpiration (the predominant components of ET). The parameterization scheme of ET within CLM5 does not accurately represent the ET-related processes in such an environment characterized by comprehensive multilayer canopies.

4.3. SM

In this section, we evaluated the SM of three CLM5-based simulation experiments using gridded reference data (ITP SM) and gauge-based measurements (WS2019 SM). Notably, the first nine soil layers (0–1.063 m) in CLM5 were adjusted to match the referenced depths (0–5, 5–20, and 20–100 cm for ITP SM, and 0–10, 10–20, 40–50, 60–70, and 90–100 cm for WS2019 SM). This alignment was achieved using the weight averaged of the soil layer thicknesses [25], as described in Section 3.2. Because of the strong spatial heterogeneity of SM, assessing the representativeness of station measurements was a challenging task.

In this study, we employed the average of the simulated SM values over China and its ten basins when compared with site-observed SM. In addition, we only selected grid cells containing at least one observation station to ensure meaningful comparisons.

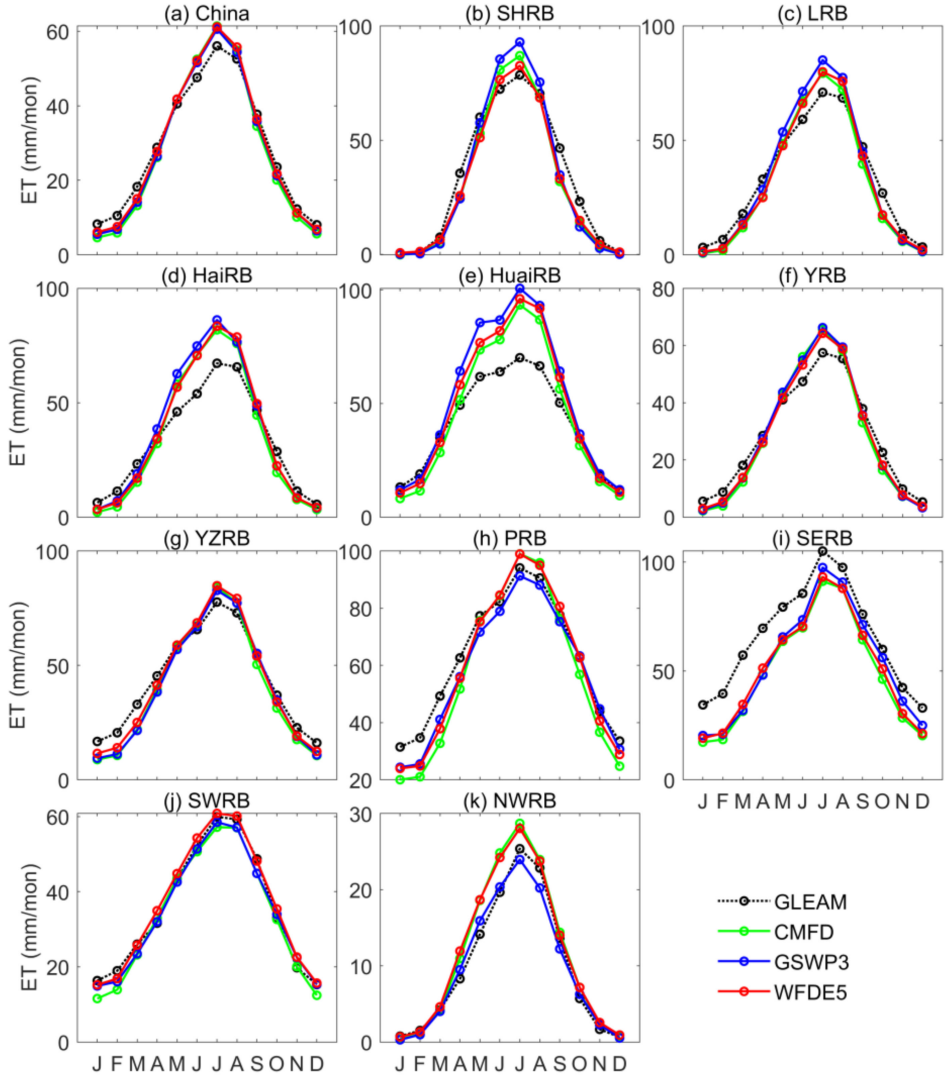

Figure 6. Seasonal cycles of ET from the GLEAM, CMFD, GSWP3, and WFDE5 over (**a**) China and its ten basins (**b**) SHRB, (**c**) LRB, (**d**) HaiRB, (**e**) HuaiRB, (**f**) YRB, (**g**) YZRB, (**h**) PRB, (**i**) SERB, (**j**) SWRB, and (**k**) NWRB during the period of 1985–2014.

Figure 8 illustrates the spatial distribution of multi-year (2003–2010) averaged SM in the 0–5 cm soil layer for ITP SM, three CLM5-simulated SMs, and their differences. Compared with the ITP SM, all three simulated SMs generally capture the spatial pattern (increasing gradually from northwest to southeast) and dry/wet centers of the SM in most areas over China (Figure 8a–c). However, some differences in spatial variations are also evident. The domain-averaged SM values in the 0–5 cm layer over China for CMFD, GSWP3, and WFDE5 are 0.233, 0.226, and 0.230 m^3/m^3, respectively, all remarkably higher than ITP SM (0.164 m^3/m^3). This indicates that CLM5 substantially overestimates SM in the shallow soil layer. The wet biases in the 0–5 cm SM mainly result from overestimations occurring in specific regions, such as HuaiRB, HaiRB, YZRB, YRB, SWRB, and NWRB. In particular, in HuaiRB, the RB values are 87.6%, 84.8%, and 86.8% for CMFD, GSWP3, and WFDE5, and the KGE values are negative against ITP SM. It suggests the poor skill of CLM5 in SM modeling for the 0–5 cm over HuaiRB (Table 4). Comparably, the overestimations of SM in CLM5 can also be documented in previous studies [28,42,68]. Nevertheless, CLM5

underestimates SM in the 0–5 cm soil layer over SERB, southeast of YZRB and north of SHRB. In addition, we also compared spatial patterns of SM simulations in the deeper soil layers (e.g., 5–20 cm and 20–100 cm) with ITP-based SM (Figures S1 and S2). Different from the results of 0–5 cm SM, both 5–20 cm and 20–100 cm SM present closer domain-averaged SM to ITP SM values, although slight underestimations exist in parts of China. For instance, the biases of 5–20 cm SM are only −0.014, −0.020, and −0.016 m^3/m^3 for CMFD, GSWP3, and WFDE5 against ITP SM, respectively. Overall, the CLM5 performs better in deeper soil layers with higher mean KGE values (0.54 for 5–20 cm SM, and 0.52 for 20–100 cm SM, respectively) than that in shallower soil layer (0.43 for 0–5 cm SM) based on CMFD, GSWP3, and WFDE5, relative to ITP-based SM over China (Tables 4, S1 and S2).

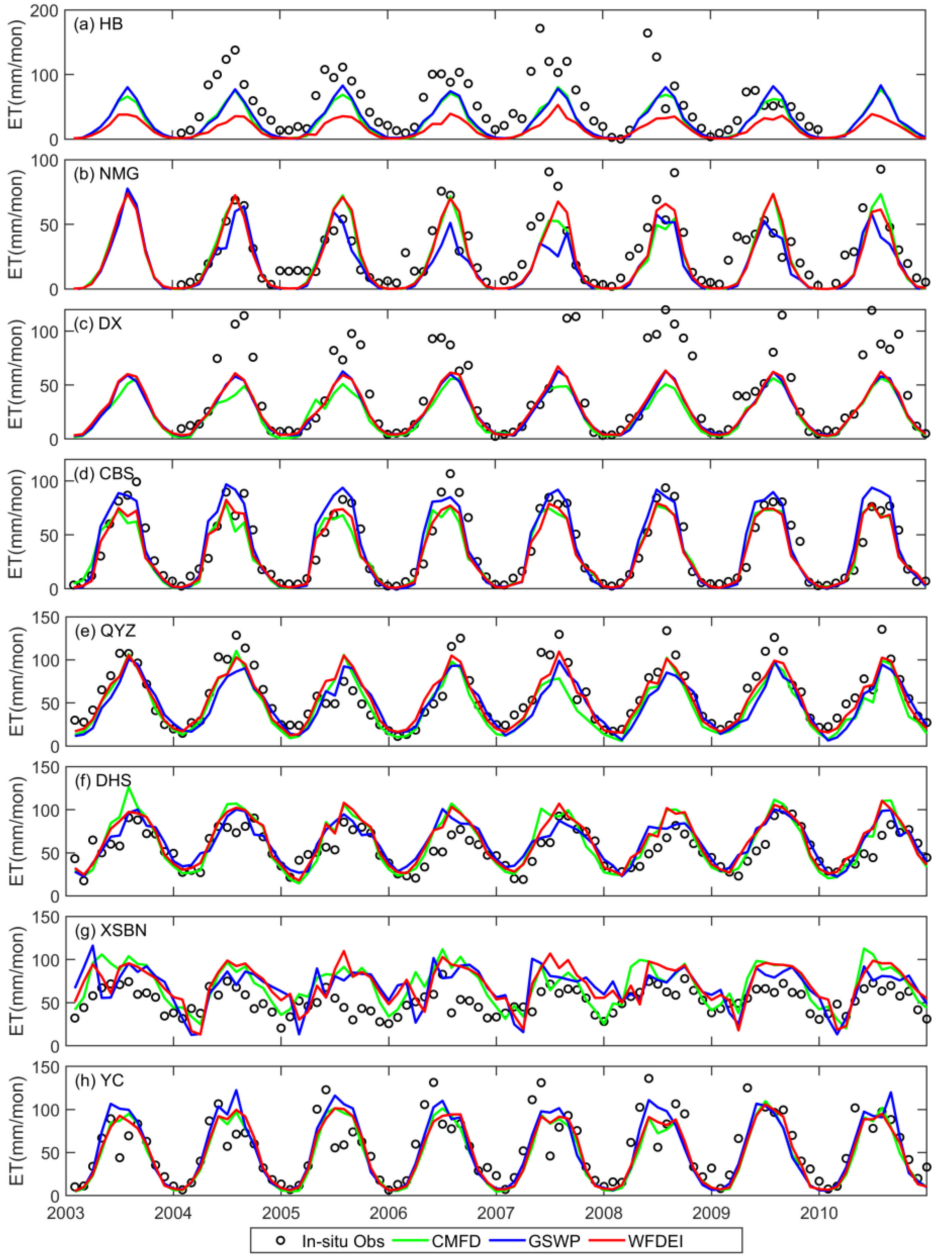

Figure 7. Comparison of monthly mean ET simulations from CMFD, GSWP3, and WFDE5 with ET measurements from ChinaFLUX at the 8 stations (**a**) HB, (**b**) NMG, (**c**) DX, (**d**) CBS, (**e**) QYZ, (**f**) DHS, (**g**) XSBN, and (**h**) YC distributed in China during the period of 2003–2010.

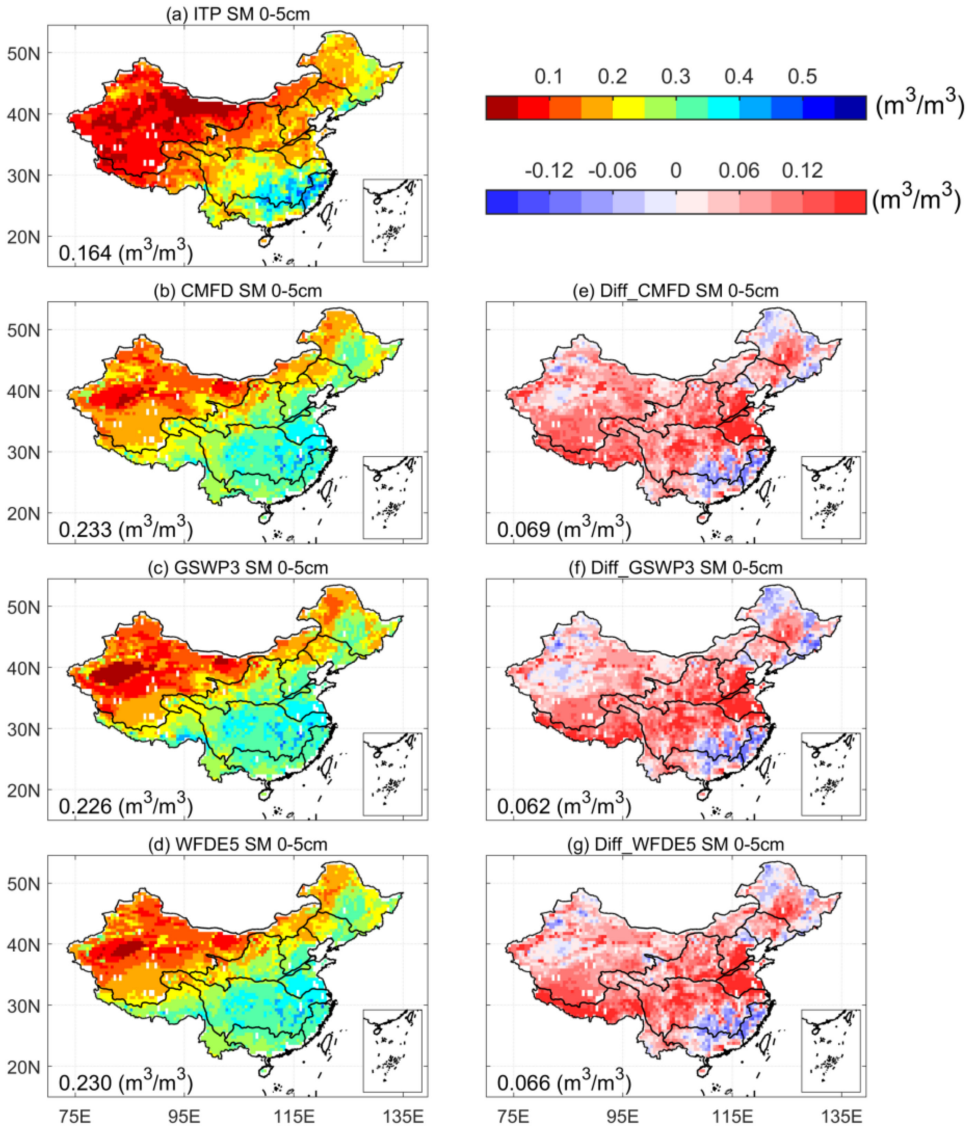

Figure 8. The spatial distribution of multi-year averaged SM at the 0–5 cm depth for the (**a**) ITP, (**b**) CMFD, (**c**) GSWP3, and (**d**) WFDE5, and (**e**) CMFD minus ITP, (**f**) GSWP3 minus ITP, and (**g**) WFDE5 minus ITP during the period of 2003–2010. The numbers in the lower left corner indicate the domain average.

Table 4. Statistical metrics (Bias, RB, RMSE, Corr, and KGE) of averaged annual SM for the 0–5 cm soil layer between simulations (CMFD, GSWP3, and WFDE5) and ITP-based SM over China and its ten basins for the period of 2003–2010.

Regions	Simulations	Bias (m^3/m^3)	RB (%)	RMSE (m^3/m^3)	Corr	KGE
China	CMFD	0.069	72.6	0.094	0.75	0.40
	GSWP3	0.062	61.7	0.093	0.73	0.47
	WFDE5	0.066	69.3	0.095	0.73	0.42
SHRB	CMFD	0.030	17.6	0.069	0.23	0.18
	GSWP3	0.017	11.8	0.071	0.10	0.03
	WFDE5	0.032	19.3	0.070	0.19	0.16
LRB	CMFD	0.047	30.5	0.070	0.61	0.42
	GSWP3	0.040	27.3	0.068	0.55	0.40
	WFDE5	0.051	33.1	0.075	0.54	0.37

Table 4. *Cont.*

Regions	Simulations	Bias (m^3/m^3)	RB (%)	RMSE (m^3/m^3)	Corr	KGE
HaiRB	CMFD	0.097	62.9	0.123	−0.39	−0.49
	GSWP3	0.095	62.3	0.124	−0.40	−0.51
	WFDE5	0.092	60.5	0.123	−0.45	−0.55
HuaiRB	CMFD	0.161	87.6	0.166	0.07	−0.32
	GSWP3	0.155	84.8	0.161	0.09	−0.28
	WFDE5	0.159	86.8	0.165	0.05	−0.32
YRB	CMFD	0.091	61.8	0.103	0.54	0.25
	GSWP3	0.094	63.5	0.107	0.53	0.22
	WFDE5	0.081	56.4	0.098	0.37	0.17
YZRB	CMFD	0.066	35.5	0.097	0.53	0.29
	GSWP3	0.072	38.8	0.103	0.47	0.22
	WFDE5	0.068	36.7	0.100	0.48	0.26
PRB	CMFD	0.018	12.1	0.069	0.39	0.07
	GSWP3	0.021	13.2	0.073	0.25	−0.03
	WFDE5	0.013	10.6	0.069	0.35	0.02
SERB	CMFD	−0.038	−5.7	0.087	−0.07	−0.21
	GSWP3	−0.045	−7.5	0.091	−0.09	−0.24
	WFDE5	−0.043	−6.8	0.089	−0.08	−0.22
SWRB	CMFD	0.102	87.7	0.112	0.68	0.07
	GSWP3	0.115	98.1	0.129	0.50	−0.01
	WFDE5	0.117	104.1	0.131	0.36	−0.17
NWRB	CMFD	0.072	124.0	0.086	0.48	−0.12
	GSWP3	0.051	91.1	0.070	0.46	0.13
	WFDE5	0.062	111.0	0.080	0.39	−0.05

We also assessed the performance levels of CMFD, GSWP3, and WFDE5 in estimating SM using in situ SM observations (WS2019 SM). Figure 9 depicts the comparison of monthly anomalies of in situ SM and CLM5-simulated SM in the 0–10 cm soil layer over China and its ten basins during the period of 1992–2013. In general, all three CLM5 experiments exhibit similar wet or dry fluctuations as site-based SM observations over China during the period of 1992–2013 (Figure 9a), with Corr values of 0.84, 0.83, and 0.83 for CMFD, GSWP3, and WFDE5, respectively. In terms of basins, the best performance is obtained in HuaiRB, with the averaged Corr value greater than 0.82 and mean RMSE value less than 0.200 m^3/m^3 derived from three CLM5 simulations. It indicates that the wet biases in HuaiRB are systematic, supported by the higher correlation of SM anomalies but higher RB values of SM. By contrast, the simulated SM in SERB shows the worst performance, with a mean RMSE value of 0.076 m^3/m^3, and all Corr values being positive.

Moreover, the seasonal cycles of site-measured SM and CLM5-simulated SM in the 0–10 cm soil layer over China are displayed in Figure 10. Apparently, the systematic wet biases occur in most basins of China. The domain-averaged in situ SM in China is 0.259 m^3/m^3, whereas all the simulated SM values exceed 0.310 m^3/m^3. For basins, SERB shows the most considerable differences in seasonality and magnitudes (Figure 10i). The peak value of the observed SM over SERB occurred in October, while the modeled SM peaks in May. In addition, the domain-averaged SM measurements are 0.287 m^3/m^3, but the SM simulations are 0.381, 0.379, and 0.380 m^3/m^3 for CMFD, GSWP3, and WFDE5 over SERB, respectively. Notably, the performance levels of CLM5 are also undesirable over NWRB even though the values of SM in the 0–10 cm soil layer are lower. This may result from the systematic bias in the model parameterization schemes. Additionally, CLM5 mostly considers soil textures for hydraulic parameters and neglects biological activities [28,69], leading to uncertainties and deficiencies in SM modeling. In comparison, CLM5 is able to accurately simulate 0–10 cm SM seasonality as well as amplitude over

PRB (Figure 10h). Moreover, we also compared the deeper soil layers (40–50 cm and 90–100 cm) SM simulation derived from the three CLM5 simulations with in situ SM observations (Figures S3–S6). The results indicate that the performance levels of CLM5 in simulating SM for deeper soil layers are better than for shallower soil layers, which are consistent with the result obtained from ITP SM.

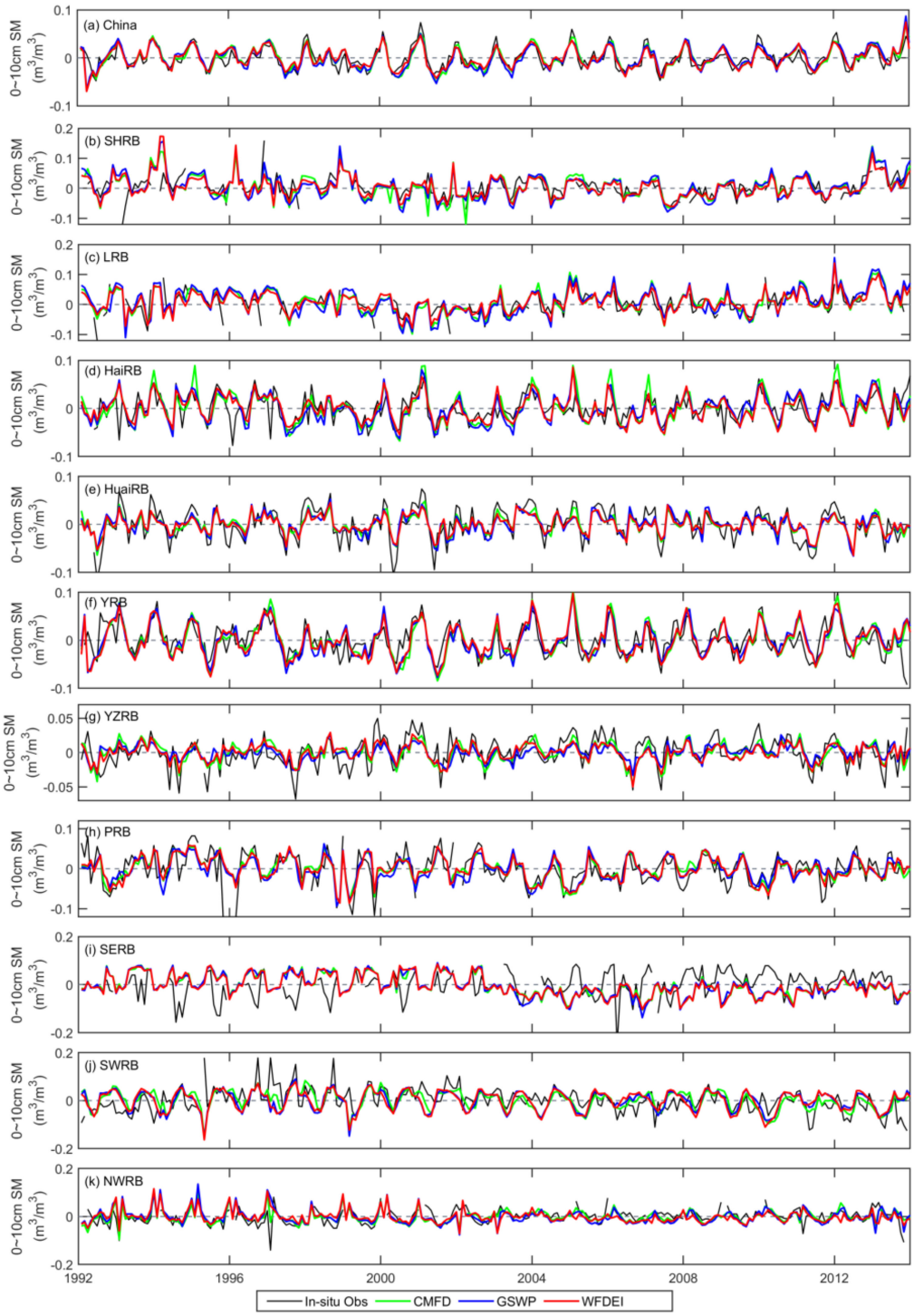

Figure 9. Monthly anomalies of WS2019 (in situ Obs) and CLM5-simulated (CMFD, GSWP3, and WFDE5) SM at the 0–10 cm depth over (**a**) China and its ten basins (**b**) SHRB, (**c**) LRB, (**d**) HaiRB, (**e**) HuaiRB, (**f**) YRB, (**g**) YZRB, (**h**) PRB, (**i**) SERB, (**j**) SWRB, and (**k**) NWRB during the period from 1992 to 2013.

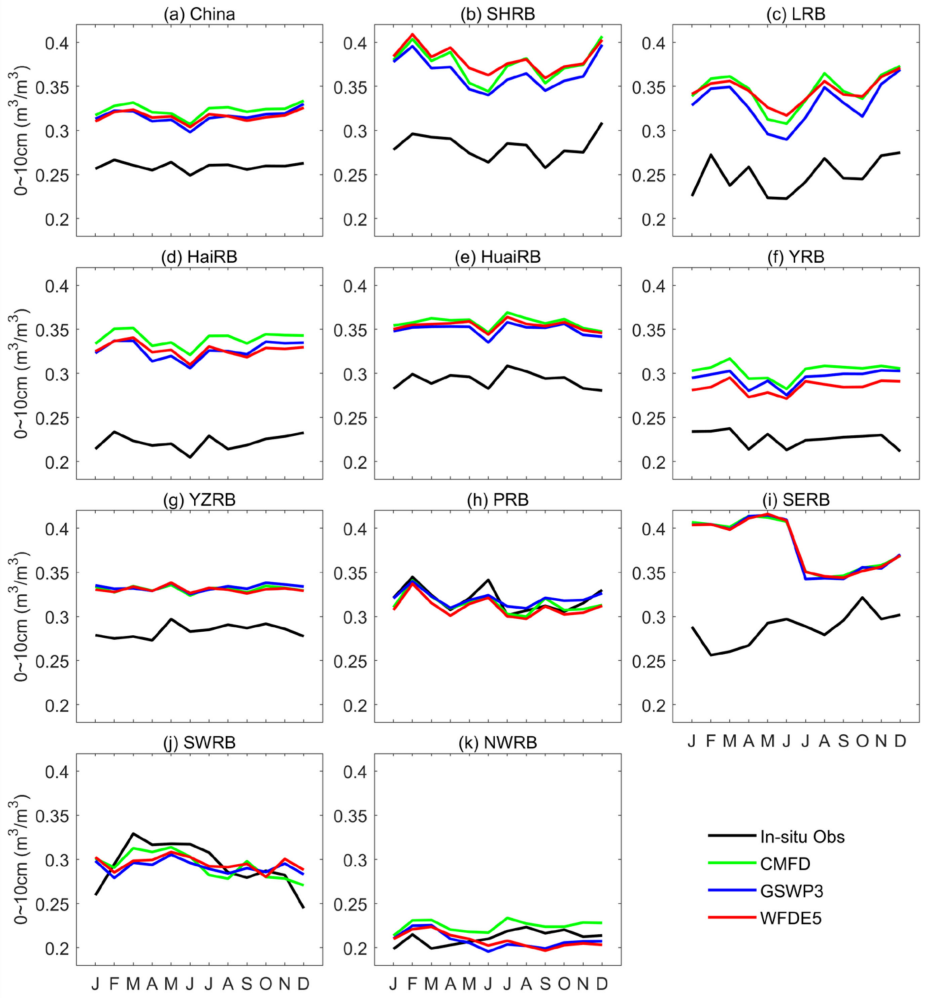

Figure 10. Seasonal cycles of site-based measured SM and CLM5-based simulated SM in the 0–10 cm soil layer over (**a**) China and its ten river basins (**b**) SHRB, (**c**) LRB, (**d**) HaiRB, (**e**) HuaiRB, (**f**) YRB, (**g**) YZRB, (**h**) PRB, (**i**) SERB, (**j**) SWRB, and (**k**) NWRB from 1992 to 2013.

4.4. Runoff

In this study, reanalysis-based CNRD and site-based CWRB runoff data were used as reference data to assess the CLM5-simulated runoff forced by three meteorological datasets. Figure 11 illustrates the spatial patterns of multi-year (1985–2014) averaged runoff for the CNRD, CMDF, GSWP3, and WFDE5 over China, along with the trends of them using Sen's slope and Mann–Kendall test, respectively. All the simulated runoffs present similar spatial distributions to CNRD-derived runoff, with KGE values of 0.80, 0.78, and 0.75 for CMFD, GSWP3, and WFDE5, respectively (Table 5). The high-value runoff centers are produced by the monsoon climate across SERB, eastern PRB, and southeastern PRB, whereas lower runoff values are observed across northwestern NWRB. CLM5 models lower domain-averaged runoff values (257, 255, and 241 mm/yr for CMFD, GSWP3, and WFDE5, respectively) than CNRD-based runoff (302 mm/yr). However, it should be noted that several detailed features were not well captured by all three of the CLM5 simulations. For instance, the snowmelt runoff produced by the special geographic environment of northwestern China, where two basins (Junggar basin and Tarim basin) are sandwiched between three mountains (Kunlun Mountains, Tianshan Mountain, and Altai Mountains), should be relative higher runoff values. These features are partially captured by CMFD, but overlooked in GSWP3 and WFDE5, possibly due to the use of the same meteorological forcing data (CMFD) used in CNRD.

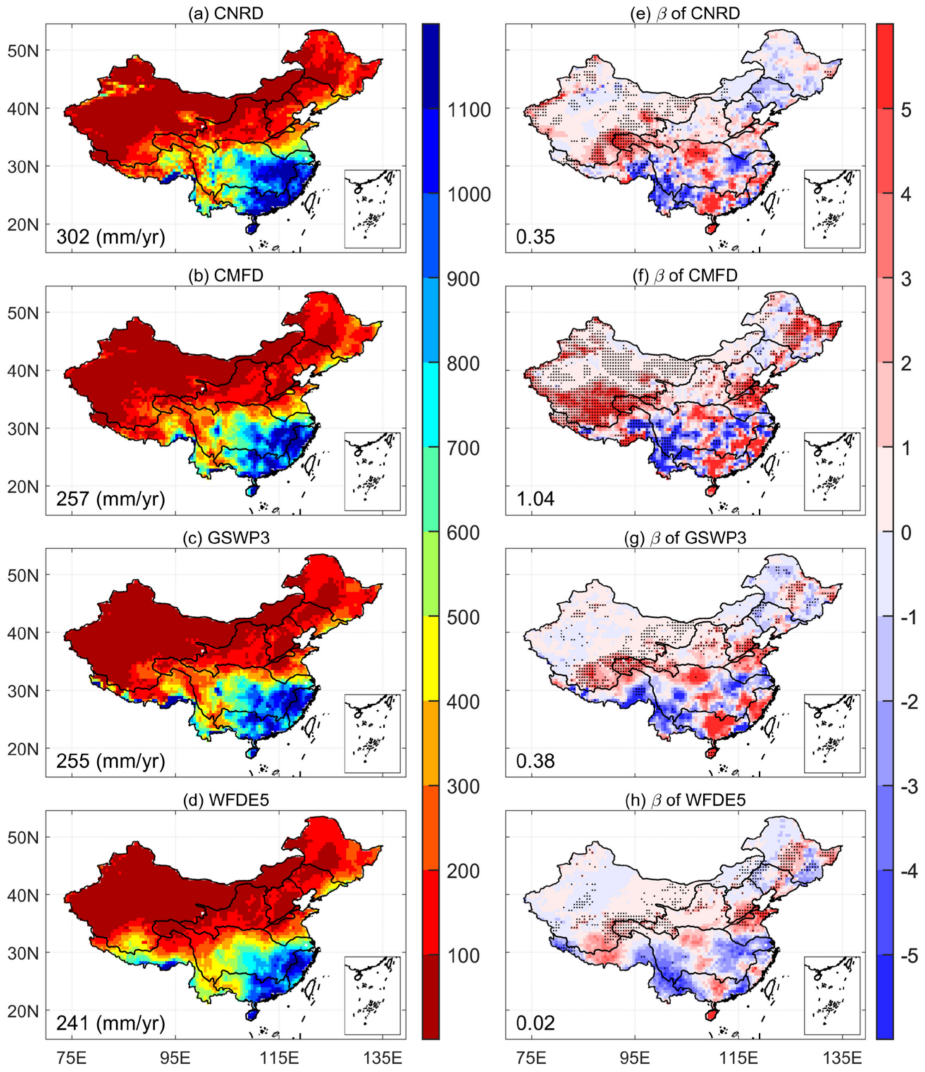

Figure 11. The spatial patterns of multi-year averaged runoff for (**a**) CRND, (**b**) CMFD, (**c**) GSWP3, and (**d**) WFDE5, and the β values derived from (**e**) CRND, (**f**) CMFD, (**g**) GSWP3, and (**h**) WFDE5 based on Sen's slope method during the period of 1985–2014. The positive values (blue grids) represent increasing trend of runoff, and negative values (red grids) represent decreasing trend runoff. The grid cells with statistically significant ($p < 0.05$) trends are marked with black dots using Mann–Kendall test. The numbers in the lower left corner indicate the domain average.

Table 5. Statistical metrics (Bias, RB, RMSE, Corr, and KGE) of averaged annual runoff between simulations (CMFD, GSWP3, and WFDE5) and CNRD-derived runoff over China and its ten basins during the period of 1985–2014.

Regions	Simulations	Bias (mm/yr)	RB (%)	RMSE (mm/yr)	Corr	KGE
China	CMFD	−45.7	−85.6	192.2	0.88	0.80
	GSWP3	−47.6	−99.4	155.4	0.92	0.78
	WFDE5	−60.8	−90.1	211.7	0.86	0.75
SHRB	CMFD	15.6	17.2	65.3	0.75	0.72
	GSWP3	−10.7	0.9	53.3	0.78	0.76
	WFDE5	8.5	15.4	56.5	0.77	0.76
LRB	CMFD	−14.7	−13.1	66.4	0.91	0.60
	GSWP3	−39.1	−23.7	68.4	0.90	0.57
	WFDE5	−19.0	−11.7	63.4	0.89	0.70

Table 5. *Cont.*

Regions	Simulations	Bias (mm/yr)	RB (%)	RMSE (mm/yr)	Corr	KGE
HaiRB	CMFD	−33.4	34.3	111.2	0.03	−0.06
	GSWP3	−64.9	5.9	138.0	−0.11	−0.78
	WFDE5	−73.8	5.6	142.8	−0.31	−0.93
HuaiRB	CMFD	−85.7	−23.6	105.1	0.92	0.65
	GSWP3	−161.3	−45.1	177.7	0.88	0.21
	WFDE5	−138.7	−37.7	153.8	0.90	0.47
YRB	CMFD	−24.8	−18.3	65.3	0.77	0.70
	GSWP3	−33.6	−27.1	75.0	0.71	0.60
	WFDE5	−60.4	−29.3	103.1	0.47	0.31
YZRB	CMFD	−80.5	−12.2	159.4	0.91	0.79
	GSWP3	−85.3	−12.6	159.1	0.91	0.81
	WFDE5	−151.5	−23.9	212.3	0.88	0.70
PRB	CMFD	−104.3	−7.3	203.0	0.87	0.77
	GSWP3	−111.7	−6.6	237.8	0.80	0.68
	WFDE5	−163.7	−13.8	235.3	0.88	0.76
SERB	CMFD	13.4	2.2	179.9	0.37	0.33
	GSWP3	−31.9	−1.8	202.8	0.19	0.15
	WFDE5	−61.8	−4.6	177.0	0.25	0.25
SWRB	CMFD	−134.0	1.2	548.3	0.57	0.34
	GSWP3	35.8	163.8	343.9	0.85	0.81
	WFDE5	−58.6	142.4	550.2	0.53	0.27
NWRB	CMFD	−27.6	−229.0	72.7	0.62	0.38
	GSWP3	−45.0	−292.3	91.7	0.45	−0.55
	WFDE5	−21.1	−261.0	109.6	0.37	−0.06

In order to investigate the change in modeled runoff during the past years, the trends of runoff simulated by the three types of CLM5 and derived from CNRD are also displayed in Figure 11e–h. We applied β values, calculated using Sen's slope method, to estimate the trend and magnitude of runoff and applied the Mann–Kendall test to determine the significance level. Positive β values indicate increasing trends (red grids), while negative values indicate decreasing trends (blue grids). Grid cells with statistically significant ($p < 0.05$) trends are marked with black dots. Generally, increasing trends in runoff are mainly observed in northern YZRB, central SERB, and central and southern PRB, while decreasing trends are mainly distributed in southern SWRB, southwestern YZRB, and western PRB in CNRD. Although all three CLM5 simulations capture a comparable spatial pattern to CNRD, some differences exist among them. For example, CMFD presents more fragmented spatial patterns and sudden transitions between increasing and decreasing trends, while WFDE5 exhibits more smoother transitions in trend distribution compared to CMFD and GSWP3, with lower increasing and decreasing amplitudes. In terms of significance level, CNRD-derived runoff shows more significant changes in NWRB than in other regions, and these spatial patterns are successfully simulated by all three CLM5 experiments. In generally, GSWP3 simulates a more similar spatial distribution of black dots and a closer domain-averaged value of β to CNRD than CMFD and WFDE5.

We also assessed three CLM5-simulated runoffs employing the independent observations, CWRB-based runoff, for the period of 2005–2014. Figure 12 displays the annual time series of basin domain-averaged runoff from CWRB, CMFD, GSWP3, and WFDE5. Overall, the timing of dry and wet years in three CLM5 simulations closely matches that of CWRB, with biases of −1.6, −7.2, and −28.4 mm/yr and RMSE values 37.5, 38.9, and 46.9 mm/yr for CMFD, GSWP3, and WFDE5, respectively, in comparison to CWRB (Figure 12). However, the performance levels of CLM5 differ among basins, and the statistical metrics are provided in Table S4. For instance, almost all three CLM5 simulations dramatically overes-

timate the runoff over HaiRB, HuaiRB, and YRB, especially for CMFD, whose RB values are 157.2%, 64.0%, and 82.8% against CWRB, respectively. Nevertheless, runoff values are underestimated by all three CLM5 simulations in SWRB, where all the RB values are less than −30%. The best performance of CLM5 occurs in YZRB, with corresponding KGE values of 0.90, 0.90, and 0.78 for CMFD, GSWP3, and WFDE5, respectively (Table S4). In contrast, the worst performance occurs in SWRB, where the KGE values of CMFD and WFDE5 are negative. In general, GSWP3 presents a more reliable and robust performance in runoff simulation over most basins of China when compared to CWRB referenced dataset.

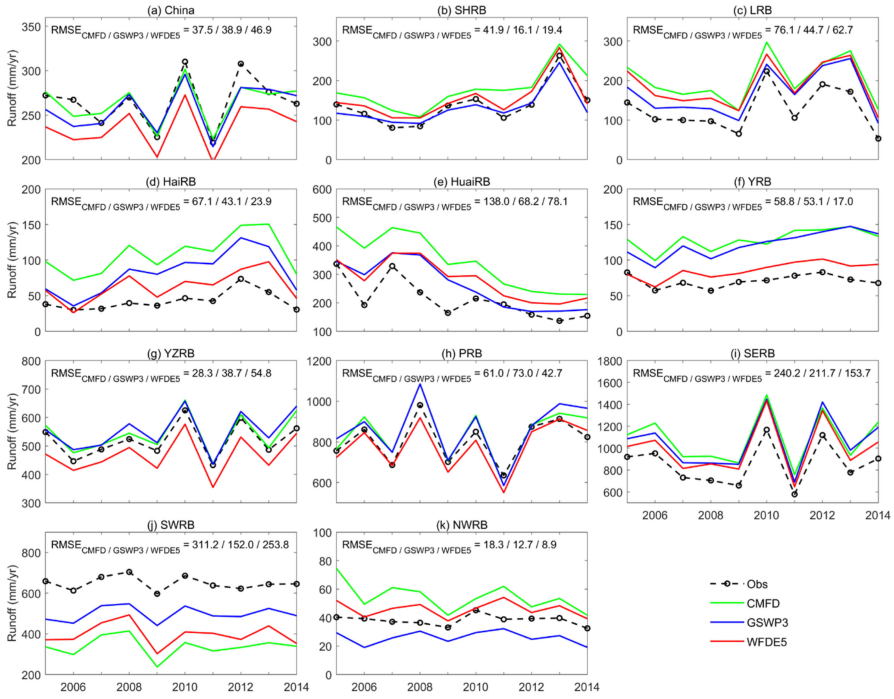

Figure 12. The annual time series of domain-averaged runoff obtained from CWRB, CMFD, GSWP3, and WFDE5 over (**a**) China and its ten basins (**b**) SHRB, (**c**) LRB, (**d**) HaiRB, (**e**) HuaiRB, (**f**) YRB, (**g**) YZRB, (**h**) PRB, (**i**) SERB, (**j**) SWRB, and (**k**) NWRB during the period from 2005 to 2014. The numbers represent the RMSE of CMFD, GSWP3, and WFDE5 against CWRB, respectively.

5. Discussion

This study employed three meteorological forcing datasets, namely CMFD, GSWP3, and WFDE5, to drive CLM5 to simulate land surface process. Diverse grid-based and gauge-based observations were taken as reference data to assess the performance levels of different types of CLM5 in ET, SM, and runoff simulations. Although a number of useful results have been obtained, there are still several issues worth discussing.

For ET simulations, all three types of CLM5 simulations can capture the general ET spatial pattern over China, but they performed poorly in NWRB. Both CMFD and WFDE5 show substantial overestimation, while GSWP3 presents obvious underestimation when compared to GLEAM ET. Furthermore, the comparison with ChinaFLUX sites also reveals that all three CLM5 simulations produce lower ET values than the site-based ET observations at HB and DX (sites located in NWRB). Thus, the validation results using independent reference data are consistent with each other. Actually, the poor performance levels in ET simulation over northwestern China exist not only in CLM5, but also in other LSMs. For instance, Sun et al. [70] evaluated the spatial patterns of ET estimates from three LSMs, namely Noah, Catchment LSM, and VIC, based on ChinaFLUX site-based observations. Their results also presented systematic negative biases at the HB and DX sites, particularly at HB, where the relative error of the monthly ET can reach up to about −50%, which is in agreement with the RB values derived in this study. Despite distinctions in the model of ET, model structures and model parametrizations, generally, all LSMs apply the

Penman–Monteith approach for potential ET (PET) computations, and then evaporation and transpiration are transformed by scaling PET [71]. The specific scaling methods and model parameterizations may differ in each LSM. The undesirable performance levels in modeling ET over northwestern China within almost all LSMs indicate that the ET processes are highly complex and susceptible, easily influenced by frequent regional climate condition, vegetation degradation, and fragile ecological environment. Consequently, there is significant potential for LSMs to improve the accuracy of ET simulation in northwestern China. In addition, in this study, the simulated ET generally showed overestimation in XSBN, which was also reported in previous studies [51,70]. Sun et al. [70] investigated the spatial pattern of ET estimates from three land surface models over China, and they found that Noah- and CLSM-modeled ET showed systematic positive biases at the XSBN site. Similar overestimations were also demonstrated in Bai et al. [51]. Moreover, the general overestimations are probably related to the satellited phenology (SP) mode of CLM5 in this study. In SP mode, the vegetation features, including canopy height of bottom and top, LAI, were obtained from daily linear interpolation between adjacent monthly values. The XSBN is located in tropical forest with dense vegetation, and thus, reasonable LAI retrievals may be challenging. In addition, the climatological LAI data were applied in SP mode only with a seasonal cycle but without interannual variation, which may also exert a profound impact on ET simulation in CLM5. For SM simulations, there are always systematic wet biases for shallower soil layers when compared to both grid-based and site-based SM measurements. This phenomenon has also been reported in previous studies. For example, Liu et al. [41] conducted a series of simulation experiments with two versions of CLM (CLM3.5 and CLM4.5) and two meteorological forcing datasets (PRIN and CMFD). They found that almost all experiments simulated higher SM for the 0–10 cm soil layer than in situ observations. They concluded that meteorological forcing is the most influential factor governing SM. In CLM5, a dry surface layer was introduced in the evaporation resistance parameterization, which seemed to alleviate the ET biases over arid and semi-arid regions. However, this introduction increased total water storage, resulting in overestimation in SM simulation by Deng et al. [68]. In this study, although three meteorological forcing datasets were employed to drive CLM5, discrepancies among different types of simulations do not appear distinct, which implies that meteorological forcing may not be the sole determinant affecting SM. In reality, SM is highly sensitive to soil condition (e.g., soil texture, soil hydraulic properties, soil structures, LAI, and land use categories). Among these factors, it is difficult to identify which one plays a dominant role, as they all contribute in some way to the uncertainty in modeling SM and may even compensate for each other. Additionally, the soil hydraulic information used in LSM usually only considers soil texture but overlooks the effect of soil structures, such as biopores and soil aggregates, created by biological activities [69,72,73]. These differences could partly account for the mismatch between simulated and observed SM. Therefore, continuous efforts are necessary to develop a more comprehensive and accurate soil moisture parameterization.

For runoff simulations, we took runoff instead of streamflow as the evaluated hydrological variable. There are two reasons for this selection. First, the streamflow data are almost unavailable in China. Second, the runoff routing parameterization schemes in different land surface models are influenced by various topographic factors (e.g., flow direction, channel length, topographic and river slopes). Comparatively, runoff, as the portion of precipitation deducting ET and change in soil moisture, can be well estimated and constructed by using a hydrological model. Therefore, the CNRD runoff was used as the reference data to evaluate different types of CLM5 experiments, due to its high-quality and continuous spatiotemporal characteristics. CNRD was generated using VIC model driven by CMFD. Nevertheless, the differences between CNRD runoff and CLM5-simulated runoff driven by CMFD also exists (Figure 11a,b). Specifically, VIC was able to simulate snowmelt runoff occurring in northwestern NWRB (Junggar basin and Tarim basin) and southeastern NWRB (Qaidam Basin), while CLM5 appears to struggle with capturing these processes. Runoff exerts an important influence on other water budget fluxes (e.g., ET and SM), as

runoff partitioning and surface energy partitioning inherently interact closely with each other. It is essential to note that the hydrological parameters in the model significantly affect runoff modeling. In the past decade, various approaches have been proposed to improve the parameters calibration, which provide valuable insights for future research to calibrate relevant hydrological parameters (e.g., maximum fractional saturated area, decay factor representing the distribution of surface runoff and sub-surface runoff) and enhance runoff simulations of LSM [74,75]. Therefore, extra attention should be given to develop a reliable runoff calibration scheme to improve the performance of CLM5 in runoff simulation [76–78].

Regarding the results derived from three types of CLM5 experiments, the discrepancies among different simulations do exist, which probably result from the distinctions in meteorological components (e.g., precipitation, solar shortwave, air temperature). Moreover, there is no consensus on which dataset is the best for simulating hydrological processes over China, because each forcing dataset has its own advantages and disadvantages. Thus, merging different climate products with various error correction and uncertainty analysis to obtain more accurate meteorological datasets may be a feasible alternative. In the case of precipitation, diverse precipitation products can be further integrated via data fusion, which can fully synthesize multi-source information and take advantage of their respective strengths to improve the quality of precipitation [79,80]. In addition, artificial intelligence algorithms may also be an effective way to achieve the data fusion well, with the booming development of computer science technology.

In this study, several limitations should be noted. Firstly, in order to evaluate the performance levels of the three types of CLM5 at same spatial resolution, we upscaled the CMFD meteorological forcing dataset from 0.1° grid cells to 0.5° using a simple arithmetic average to match GSWP3 and WFDE5. This treatment may introduce some uncertainties, and the advantages of CMFD may not be fully realized. Nevertheless, CMFD still demonstrates better performance in many regions. For example, CMFD exhibits the highest Corr value of 0.94 for simulating ET over China against GLEAM ET, and the highest KEG value is also observed for CMFD over YRB. Liu et al. [41] ran CLM3.5 with two atmospheric forcing dataset, and reported that mean Corr values of ET improves from 0.86 (driven by PRIN) to 0.89 (driven by CMFD) over China against reference ET data. Secondly, to partially mitigate the uncertainties from single reference data, we employed both site-based observations (ChinaFLUX for ET, WS2019 for SM) and grid-based estimations (GLEAM for ET, ITPLDAS for SM) to evaluate the performance levels of CLM5. However, due to the sparse gauge distribution, there may be fewer sites located in a certain grid cell and the representation of site data may be limited, which may also introduce potential biases and uncertainties to the results. Therefore, it is essential to develop higher-quality gridded ET and SM benchmark products by blending multiple datasets. These endeavors would be crucial for the assessment and improvement of LSMs in future studies. Thirdly, SM is observed at fixed depths, but it is nonlinear in space. We adjusted the multiple soil layer thicknesses of CLM5-simulated SM to agree with the observed soil layers using a weight-average approach based on soil layer thicknesses to calculate SM. However, this adjustment for SM is inherently inaccurate, and thus, biases and uncertainties may also exist.

6. Conclusions

This study conducted three offline LSM simulations over China using CLM5 driven by different meteorological forcing datasets, namely CMFD, GSWP3, and WFDE5, respectively. We then used various grid-based estimations and gauge-based observations as the reference data to assess the performance levels of these three CLM5 experiments in hydrological variables, such as ET, SM, and runoff, across China and its ten basins. The main conclusions are summarized as follows:

Firstly, all three types of CLM5 reasonably simulate ET values and spatial patterns over China, exhibiting higher Corr values (all are greater than 0.92) and KGE values (all are greater than 0.85) against GLEAM ET. As for basins, the best performance is observed in

Yellow River Basin (YRB), while the worst performance occurs in northwest rivers basin (NWRB) of China. Substantial biases exist in most areas of NWRB, with either overestimations or underestimations, which are mainly stemming from complex climate conditions and low-quality atmospheric forcing data. These undesirable performance levels are also evident when compared with ChinaFLUX site-based ET data, particularly at stations located in NWRB. In general, CMFD and GSWP show slightly better performance levels than WFDE5 in terms of the KGE values in ten basins. Secondly, although both CLM5-simulated SM and ITP-derived SM show a gradually increasing trend from northwest to southeast China, there are still visible discrepancies between the two. Specifically, systematic wet biases are prevalent in CLM5 simulations for shallower soil layers (the top 0–10 cm) across most areas of China when compared to grid-based SM estimations as well as site-based SM measurements. Comparatively, all CLM5 performance levels in SM simulations for deeper soil layers are better, with reduced RMSE values and improved KGE values varying with regions. Thirdly, all three types of CLM5 demonstrate reasonable runoff spatial patterns. CMFD effectively captures more detailed information, such as snowmelt runoff over northwestern China, while GSWP3 simulates more comparable spatial distributions of runoff change trends to CNRD runoff. Yangtze River Basin (YZRB) shows the best performance with KGE values exceeding 0.78, while Northwest Rivers Basin (NWRB) exhibits the worst performance, with negative KGE values when taking in situ runoff as reference. Overall, GSWP3 shows a more reliable skill in runoff simulation over most areas of China.

The choice of atmospheric forcing data is crucial, but not the sole determinant to LSM performance, as it heavily depends on model parameterizations and structures. Ongoing efforts should be made to the physically based model development, particularly in regions with complex conditions, such as northwestern China. The biases and limitations mentioned in this study provide motivations and directions for model improvements in the future. Meanwhile, the higher-quality and grid-based reference data also should be developed to achieve more comprehensive and robust validations and evaluations of LSM.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/rs16030550/s1, Figure S1: The spatial distribution of multi-year averaged SM at 5–20 cm depth for (a) ITP, (b) CMFD, (c) GSWP3, and (d) WFDE5, and (e) CMFD minus ITP, (f) GSWP3 minus ITP, and (g) WFDE5 minus ITP during the period of 2003–2010. The numbers in the lower left corner indicate the domain average; Figure S2: Same as Figure S1, but for SM at the 20–100 cm soil layer depth; Figure S3: Monthly anomalies of WS2019 (in situ Obs) and CLM5-simulated (CMFD, GSWP3 and WFDE5) SM at the 40–50 cm depth over China and its ten sub-basins during the period from 1992 to 2013; Figure S4: Seasonal cycles of site-based measured SM and CLM5-based simulated SM in the 40–50 cm soil layer over (a) China and its ten river basins (b) SHRB, (c) LRB, (d) HaiRB, (e) HuaiRB, (f) YRB, (g) YZRB, (h) PRB, (i) SERB, (j) SWRB, and (k) NWRB from 1992 to 2013; Figure S5: Same as Figure S3, but for SM in the 90–100 cm soil layer; Figure S6: Same as Figure S4, but for SM in the 90–100 cm soil layer; Table S1: Statistical metrics (Bias, RB, RMSE, Corr, and KGE) of averaged annual ET between simulations (CMFD, GSWP3, and WFDE5) and ChinaFLUX observed ET at eight sites during the period of 2003–2010; Table S2: Statistical metrics (Bias, RB, RMSE, Corr, and KGE) of averaged annual SM for the 5–20 cm soil layer between simulations (CMFD, GSWP3, and WFDE5) and ITP-based SM over China and its ten sub-basins for the period of 2003–2010; Table S3: Same as Table S2, but for soil moisture in the 20–100 cm layer; Table S4: Statistical metrics (Bias, RB, RMSE, Corr, and KGE) of averaged annual runoff between simulations (CMFD, GSWP3, and WFDE5) and CWRB-based measurements over China and its ten sub-basins for the period of 1985–2014.

Author Contributions: Conceptualization, D.W. (Dagang Wang) and D.W. (Dayang Wang); methodology, D.W. (Dayang Wang); software, D.W. (Dagang Wang) and D.W. (Dayang Wang); validation, D.W. (Dayang Wang) and D.W. (Dagang Wang); formal analysis, D.W. (Dayang Wang), D.W. (Dagang Wang) and Y.M.; resources, D.W. (Dagang Wang), D.W. (Dayang Wang), Q.Y., M.J. and Y.L.; data curation, D.W. (Dayang Wang), S.L. and B.L.; Writing—original draft preparation, D.W. (Dayang Wang); review and editing, D.W. (Dagang Wang), D.W. (Dayang Wang) and Q.Y.; visualization, D.W. (Dayang Wang); supervision, D.W. (Dagang Wang), Y.M., Y.H. and C.M.; project admin-

istration, D.W. (Dagang Wang), D.W. (Dayang Wang) and Y.H.; funding acquisition, D.W. (Dagang Wang), D.W. (Dayang Wang), S.L. and Y.H. All authors have read and agreed to the published version of the manuscript.

Funding: This study was financially supported by the High-Level Talent Introduction Research Project of the Nanyang Normal University (No. 2023ZX017), the National Natural Science Foundation of China (No. 52079151), the National Natural Science Foundation Projects of International Cooperation and Exchanges (No. 52111540261), the Open Project of Laboratory of Nanyang Normal University (No. SYKF2023072), the Natural Science Foundation of Jiangsu Province (No. BK20230957), and the Jiangsu Funding Program for Excellent Postdoctoral Talent (No. 2022ZB147).

Data Availability Statement: The CMFD meteorological forcing data used in this study are available from https://data.tpdc.ac.cn, accessed on 12 December 2023. The GSWP3 data are downloadable at https://data.isimip.org, accessed on 12 December 2023. The WFDE5 data are provided at https://cds.climate.copernicus.eu, accessed on 12 December 2023. The GLEAM ET data are available from https://www.gleam.eu, accessed on 12 December 2023. The ChinaFLUX data are from http://www.cnern.org.cn, accessed on 12 December 2023. The ITPLDAS SM data are available from https://data.tpdc.ac.cn/, accessed on 12 December 2023. The WS2019 SM data are accessible by emailing the corresponding author (wangaihui@mail.iap.ac.cn) of Ma and Wang (2022). The CNRD runoff data are downloadable at https://data.tpdc.ac.cn/, accessed on 8 December 2023. The CWRB runoff data are available from https://www.mwr.gov.cn, accessed on 8 December 2023.

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

References

1. Dickinson, R.E. Land-atmosphere interaction. *Rev. Geophys.* **1995**, *33*, 917–922. [CrossRef]
2. Koster, R.D.; Dirmeyer, P.A.; Guo, Z.; Bonan, G.; Chan, E.; Cox, P.; Gordon, C.T.; Kanae, S.; Kowalczyk, E.; Lawrence, D.; et al. Regions of Strong Coupling Between Soil Moisture and Precipitation. *Science* **2004**, *305*, 1138–1140. [CrossRef] [PubMed]
3. Oleson, K.W.; Niu, G.Y.; Yang, Z.L.; Lawrence, D.M.; Thornton, P.E.; Lawrence, P.J.; Stöckli, R.; Dickinson, R.E.; Bonan, G.B.; Levis, S.; et al. Improvements to the Community Land Model and their impact on the hydrological cycle. *J. Geophys. Res. Biogeosci.* **2008**, *113*, 1–26. [CrossRef]
4. Oki, T.; Kanae, S. Global Hydrological Cycles and World Water Resources. *Science* **2006**, *313*, 1068–1072. [CrossRef]
5. Trenberth, K.E.; Fasullo, J.T.; Kiehl, J. Earth's Global Energy Budget. *Bull. Am. Meteorol. Soc.* **2009**, *90*, 311–324. [CrossRef]
6. Zhao, M.; Liu, Y.; Konings, A.G. Evapotranspiration frequently increases during droughts. *Nat. Clim. Change* **2022**, *12*, 1024–1030. [CrossRef]
7. Colliander, A.; Reichle, R.H.; Crow, W.T.; Cosh, M.H.; Chen, F.; Chan, S.; Das, N.N.; Bindlish, R.; Chaubell, J.; Kim, S.; et al. Validation of Soil Moisture Data Products from the NASA SMAP Mission. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2022**, *15*, 364–392. [CrossRef]
8. Gao, L.; Gao, Q.; Zhang, H.; Li, X.; Chaubell, M.J.; Ebtehaj, A.; Shen, L.; Wigneron, J.-P. A deep neural network based SMAP soil moisture product. *Remote Sens. Environ.* **2022**, *277*, 113059. [CrossRef]
9. Veldkamp, T.I.E.; Wada, Y.; Aerts, J.C.J.H.; Döll, P.; Gosling, S.N.; Liu, J.; Masaki, Y.; Oki, T.; Ostberg, S.; Pokhrel, Y.; et al. Water scarcity hotspots travel downstream due to human interventions in the 20th and 21st century. *Nat. Commun.* **2017**, *8*, 15697. [CrossRef]
10. Munia, H.A.; Guillaume, J.H.A.; Mirumachi, N.; Wada, Y.; Kummu, M. How downstream sub-basins depend on upstream inflows to avoid scarcity: Typology and global analysis of transboundary rivers. *Hydrol. Earth Syst. Sci.* **2018**, *22*, 2795–2809. [CrossRef]
11. Miao, C.; Gou, J.; Fu, B.; Tang, Q.; Duan, Q.; Chen, Z.; Lei, H.; Chen, J.; Guo, J.; Borthwick, A.G.L.; et al. High-quality reconstruction of China's natural streamflow. *Sci. Bull.* **2021**, *67*, 547–556. [CrossRef]
12. Bonan, G.B.; Doney, S.C. Climate, ecosystems, and planetary futures: The challenge to predict life in Earth system models. *Science* **2018**, *359*, eaam8328. [CrossRef]
13. Seneviratne, S.I.; Corti, T.; Davin, E.L.; Hirschi, M.; Jaeger, E.B.; Lehner, I.; Orlowsky, B.; Teuling, A.J. Investigating soil moisture–climate interactions in a changing climate: A review. *Earth Sci. Rev.* **2010**, *99*, 125–161. [CrossRef]
14. Wang, D.; Wang, D.; Mo, C. The Use of Remote Sensing-Based ET Estimates to Improve Global Hydrological Simulations in the Community Land Model Version 5.0. *Remote Sens.* **2021**, *13*, 4460. [CrossRef]
15. Wang, A.; Zeng, X.; Guo, D. Estimates of Global Surface Hydrology and Heat Fluxes from the Community Land Model (CLM4.5) with Four Atmospheric Forcing Datasets. *J. Hydrometeorol.* **2016**, *17*, 2493–2510. [CrossRef]
16. Chen, Y.; Yuan, H. Evaluation of nine sub-daily soil moisture model products over China using high-resolution in situ observations. *J. Hydrol.* **2020**, *588*, 125054. [CrossRef]
17. Lawrence, D.M.; Fisher, R.A.; Koven, C.D.; Oleson, K.W.; Swenson, S.C.; Bonan, G.; Collier, N.; Ghimire, B.; Kampenhout, L.; Kennedy, D.; et al. The Community Land Model Version 5: Description of New Features, Benchmarking, and Impact of Forcing Uncertainty. *J. Adv. Model. Earth Syst.* **2019**, *11*, 4245–4287. [CrossRef]

18. Fisher, R.A.; Koven, C.D. Perspectives on the Future of Land Surface Models and the Challenges of Representing Complex Terrestrial Systems. *J. Adv. Model. Earth Syst.* **2020**, *12*, e2018MS001453. [CrossRef]
19. Cherkauer, K.A.; Lettenmaier, D.P. Hydrologic effects of frozen soils in the upper Mississippi River basin. *J. Geophys. Res. Atmos.* **1999**, *104*, 19599–19610. [CrossRef]
20. Andreadis, K.M.; Storck, P.; Lettenmaier, D.P. Modeling snow accumulation and ablation processes in forested environments. *Water Resour. Res.* **2009**, *45*, 1–13. [CrossRef]
21. Hamman, J.J.; Nijssen, B.; Bohn, T.J.; Gergel, D.R.; Mao, Y.X. The Variable Infiltration Capacity model version 5 (VIC-5): Infrastructure improvements for new applications and reproducibility. *Geosci. Model Dev.* **2018**, *11*, 3481–3496. [CrossRef]
22. Ghimire, B.; Riley, W.J.; Koven, C.D.; Mu, M.; Randerson, J.T. Representing leaf and root physiological traits in CLM improves global carbon and nitrogen cycling predictions. *J. Adv. Model. Earth Syst.* **2016**, *8*, 598–613. [CrossRef]
23. van Kampenhout, L.; Lenaerts, J.T.M.; Lipscomb, W.H.; Sacks, W.J.; Lawrence, D.M.; Slater, A.G.; van den Broeke, M.R. Improving the Representation of Polar Snow and Firn in the Community Earth System Model. *J. Adv. Model. Earth Syst.* **2017**, *9*, 2583–2600. [CrossRef]
24. Blyth, E.M.; Arora, V.K.; Clark, D.B.; Dadson, S.J.; De Kauwe, M.G.; Lawrence, D.M.; Melton, J.R.; Pongratz, J.; Turton, R.H.; Yoshimura, K.; et al. Advances in Land Surface Modelling. *Curr. Clim. Change Rep.* **2021**, *7*, 45–71. [CrossRef]
25. Ou, M.; Zhang, S. Evaluation and Comparison of the Common Land Model and the Community Land Model by Using In Situ Soil Moisture Observations from the Soil Climate Analysis Network. *Land* **2022**, *11*, 126. [CrossRef]
26. Yokohata, T.; Kinoshita, T.; Sakurai, G.; Pokhrel, Y.; Ito, A.; Okada, M.; Satoh, Y.; Kato, E.; Nitta, T.; Fujimori, S.; et al. MIROC-INTEG-LAND version 1: A global biogeochemical land surface model with human water management, crop growth, and land-use change. *Geosci. Model Dev.* **2020**, *13*, 4713–4747. [CrossRef]
27. Song, J.; Miller, G.R.; Cahill, A.T.; Aparecido, L.M.T.; Moore, G.W. Modeling land surface processes over a mountainous rainforest in Costa Rica using CLM4.5 and CLM5. *Geosci. Model Dev.* **2020**, *13*, 5147–5173. [CrossRef]
28. Cheng, Y.; Huang, M.; Zhu, B.; Bisht, G.; Zhou, T.; Liu, Y.; Song, F.; He, X. Validation of the Community Land Model Version 5 Over the Contiguous United States (CONUS) Using In Situ and Remote Sensing Data Sets. *J. Geophys. Res. Atmos.* **2021**, *126*, e2020JD033539. [CrossRef]
29. Parr, D.; Wang, G.; 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]
30. Wang, D.; Wang, G.; Parr, D.T.; Liao, W.; Xia, Y.; Fu, C. Incorporating remote sensing-based ET estimates into the Community Land Model version 4.5. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 3557–3577. [CrossRef]
31. Yang, S.; Li, R.; Wu, T.; Wu, X.; Zhao, L.; Hu, G.; Zhu, X.; Du, Y.; Xiao, Y.; Zhang, Y.; et al. Evaluation of soil thermal conductivity schemes incorporated into CLM5.0 in permafrost regions on the Tibetan Plateau. *Geoderma* **2021**, *401*, 115330. [CrossRef]
32. Cucchi, M.; Weedon, G.P.; Amici, A.; Bellouin, N.; Lange, S.; Müller Schmied, H.; Hersbach, H.; Buontempo, C. WFDE5: Bias-adjusted ERA5 reanalysis data for impact studies. *Earth Syst. Sci. Data* **2020**, *12*, 2097–2120. [CrossRef]
33. Wang, A.; Zeng, X. Sensitivities of terrestrial water cycle simulations to the variations of precipitation and air temperature in China. *J. Geophys. Res.* **2011**, *116*, 1–11. [CrossRef]
34. Liu, J.; Shi, C.; Sun, S.; Liang, J.; Yang, Z.-L. Improving Land Surface Hydrological Simulations in China Using CLDAS Meteorological Forcing Data. *J. Meteorol. Res.* **2020**, *33*, 1194–1206. [CrossRef]
35. Tesfa, T.K.; Li, H.Y.; Leung, L.R.; Huang, M.; Ke, Y.; Sun, Y.; Liu, Y. A subbasin-based framework to represent land surface processes in an Earth system model. *Geosci. Model Dev.* **2014**, *7*, 947–963. [CrossRef]
36. Shi, C.; Jiang, L.; Zhang, T.; Xu, B.; Han, S. Status and Plans of CMA Land Data Assimilation System (CLDAS) Project. *EGU Gen. Assem.* **2014**, *16*, EGU2014-5671.
37. He, J.; Yang, K.; Tang, W.; Lu, H.; Qin, J.; Chen, Y.; Li, X. The first high-resolution meteorological forcing dataset for land process studies over China. *Sci. Data* **2020**, *7*, 25. [CrossRef] [PubMed]
38. Liu, J.G.; Xie, Z.H. Improving simulation of soil moisture in China using a multiple meteorological forcing ensemble approach. *Hydrol. Earth Syst. Sci.* **2013**, *17*, 3355–3369. [CrossRef]
39. Lin, X.; Zhang, H.; Zhang, X.-j.; Liu, J.; Chen, J.; Shao, Q.; Chen, B.; Sun, S. Modeling Evapotranspiration over China's Landmass from 1979 to 2012 Using Multiple Land Surface Models: Evaluations and Analyses. *J. Hydrometeorol.* **2017**, *18*, 1185–1203.
40. Lu, H.; Zheng, D.; Yang, K.; Yang, F. Last-decade progress in understanding and modeling the land surface processes on the Tibetan Plateau. *Hydrol. Earth Syst. Sci.* **2020**, *24*, 5745–5758. [CrossRef]
41. Liu, J.; Yang, Z.-L.; Jia, B.; Wang, L.; Wang, P.; Xie, Z.; Shi, C. Elucidating Dominant Factors Affecting Land Surface Hydrological Simulations of the Community Land Model over China. *Adv. Atmos. Sci.* **2022**, *40*, 235–250. [CrossRef]
42. Ma, X.; Wang, A. Systematic Evaluation of a High-Resolution CLM5 Simulation over Continental China for 1979–2018. *J. Hydrometeorol.* **2022**, *23*, 1879–1897. [CrossRef]
43. Bonan, G.B. The land surface climatology of the NCAR land surface model coupled to the NCAR community climate model. *J. Clim.* **1998**, *11*, 1307–1326. [CrossRef]
44. Zhao, F.; Wang, X.; Wu, Y.; Singh, S.K. Prefectures vulnerable to water scarcity are not evenly distributed across China. *Commun. Earth Environ.* **2023**, *4*, 145. [CrossRef]
45. Hanasaki, N.; Oki, T.; Guo, Z.; Zhao, M.; Gao, X.; Dirmeyer, P.A. GSWP-2: Multimodel Analysis and Implications for Our Perception of the Land Surface. *Bull. Am. Meteorol. Soc.* **2006**, *87*, 1381–1398.

46. Hou, Y.; Guo, H.; Yang, Y.; Liu, W. Global Evaluation of Runoff Simulation from Climate, Hydrological and Land Surface Models. *Water Resour. Res.* **2022**, *59*, e2021WR031817. [CrossRef]
47. Gou, J.; Miao, C.; Samaniego, L.; Xiao, M.; Wu, J.; Guo, X. CNRD v1.0: A High-Quality Natural Runoff Dataset for Hydrological and Climate Studies in China. *Bull. Am. Meteorol. Soc.* **2021**, *102*, E929–E947. [CrossRef]
48. Lawrence, D.M.; Thornton, P.E.; Oleson, K.W.; Bonan, G.B. The Partitioning of Evapotranspiration into Transpiration, Soil Evaporation, and Canopy Evaporation in a GCM: Impacts on Land–Atmosphere Interaction. *J. Hydrometeorol.* **2007**, *8*, 862–880. [CrossRef]
49. Miralles, D.G.; Holmes, T.R.H.; De Jeu, R.A.M.; Gash, J.H.; Meesters, A.G.C.A.; Dolman, A.J. Global land-surface evaporation estimated from satellite-based observations. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 453–469. [CrossRef]
50. Martens, B.; Miralles, D.G.; Lievens, H.; van der Schalie, R.; de Jeu, R.A.M.; Fernández-Prieto, D.; Beck, H.E.; Dorigo, W.A.; Verhoest, N.E.C. GLEAM v3: Satellite-based land evaporation and root-zone soil moisture. *Geosci. Model Dev.* **2017**, *10*, 1903–1925. [CrossRef]
51. Bai, P.; Liu, X. Intercomparison and evaluation of three global high-resolution evapotranspiration products across China. *J. Hydrol.* **2018**, *566*, 743–755. [CrossRef]
52. Zhu, W.; Tian, S.; Wei, J.; Jia, S.; Song, Z. Multi-scale evaluation of global evapotranspiration products derived from remote sensing images: Accuracy and uncertainty. *J. Hydrol.* **2022**, *611*, 127982. [CrossRef]
53. Yao, T.; Lu, H.; Yu, Q.; Feng, S.; Xue, Y.; Feng, W. Uncertainties of three high-resolution actual evapotranspiration products across China: Comparisons and applications. *Atmos. Res.* **2023**, *286*, 106682. [CrossRef]
54. Jia, Y.; Li, C.; Yang, H.; Yang, W.; Liu, Z. Assessments of three evapotranspiration products over China using extended triple collocation and water balance methods. *J. Hydrol.* **2022**, *614 Pt B*, 128594. [CrossRef]
55. Yin, L.; Tao, F.; Chen, Y.; Liu, F.; Hu, J. Improving terrestrial evapotranspiration estimation across China during 2000–2018 with machine learning methods. *J. Hydrol.* **2021**, *600*, 126538. [CrossRef]
56. Li, X.; Zou, L.; Xia, J.; Dou, M.; Li, H.; Song, Z. Untangling the effects of climate change and land use/cover change on spatiotemporal variation of evapotranspiration over China. *J. Hydrol.* **2022**, *612*, 128189. [CrossRef]
57. Bai, P. Comparison of remote sensing evapotranspiration models: Consistency, merits, and pitfalls. *J. Hydrol.* **2023**, *617*, 128856. [CrossRef]
58. Yang, K.; Chen, Y.; He, J.; Zhao, L.; Lu, H.; Qin, J.; Zheng, D.; Li, X. Development of a daily soil moisture product for the period of 2002–2011 in Chinese mainland. *Sci. China Earth Sci.* **2020**, *63*, 1113–1125. [CrossRef]
59. Yang, K.; Zhu, L.; Chen, Y.; Zhao, L.; Qin, J.; Lu, H.; Tang, W.; Han, M.; Ding, B.; Fang, N. Land surface model calibration through microwave data assimilation for improving soil moisture simulations. *J. Hydrol.* **2016**, *533*, 266–276. [CrossRef]
60. Wang, A.; Shi, X. A Multilayer Soil Moisture Dataset Based on the Gravimetric Method in China and Its Characteristics. *J. Hydrometeorol.* **2019**, *20*, 1721–1736. [CrossRef]
61. Sang, Y.; Ren, H.-L.; Shi, X.; Xu, X.; Chen, H. Improvement of Soil Moisture Simulation in Eurasia by the Beijing Climate Center Climate System Model from CMIP5 to CMIP6. *Adv. Atmos. Sci.* **2021**, *38*, 237–252. [CrossRef]
62. Peng, D.; Zhou, Q.; Tang, X.; Yan, W.; Chen, M. Changes in soil moisture caused solely by vegetation restoration in the karst region of southwest China. *J. Hydrol.* **2022**, *613*, 128460. [CrossRef]
63. Zhou, Y.; Chen, H.; Sun, S. Assessing and comparing the subseasonal variations of summer soil moisture of satellite products over eastern China. *Int. J. Climatol.* **2023**, *43*, 3925–3946. [CrossRef]
64. Gou, J.; Miao, C.; Duan, Q.; Tang, Q.; Di, Z.; Liao, W.; Wu, J.; Zhou, R. Sensitivity Analysis-Based Automatic Parameter Calibration of the VIC Model for Streamflow Simulations Over China. *Water Resour. Res.* **2020**, *56*, e2019WR025968. [CrossRef]
65. Mann, H.B. Nonparametric tests against trend. *Econometrica* **1945**, *13*, 245–259. [CrossRef]
66. Sen, P.K. Estimates of the Regression Coefficient Based on Kendall's Tau. *J. Am. Stat. Assoc.* **1968**, *63*, 1379–1389. [CrossRef]
67. Zhao, R.; Wang, H.; Chen, J.; Fu, G.; Zhan, C.; Yang, H. Quantitative analysis of nonlinear climate change impact on drought based on the standardized precipitation and evapotranspiration index. *Ecol. Indic.* **2021**, *121*, 107107. [CrossRef]
68. Deng, M.; Meng, X.; Lyv, Y.; Zhao, L.; Li, Z.; Hu, Z.; Jing, H. Comparison of Soil Water and Heat Transfer Modeling Over the Tibetan Plateau Using Two Community Land Surface Model (CLM) Versions. *J. Adv. Model. Earth Syst.* **2020**, *12*, e2020MS002189. [CrossRef]
69. Fatichi, S.; Or, D.; Walko, R.; Vereecken, H.; Young, M.H.; Ghezzehei, T.A.; Hengl, T.; Kollet, S.; Agam, N.; Avissar, R. Soil structure is an important omission in Earth System Models. *Nat. Commun.* **2020**, *11*, 522. [CrossRef]
70. Sun, R.; Duan, Q.; Wang, J. Understanding the spatial patterns of evapotranspiration estimates from land surface models over China. *J. Hydrol.* **2021**, *595*, 126021. [CrossRef]
71. Kumar, S.; Holmes, T.; Mocko, D.M.; Wang, S.; Peters-Lidard, C. Attribution of flux partitioning variations between land surface models over the continental U.S. *Remote Sens.* **2018**, *10*, 751. [CrossRef]
72. Cheng, Y.; Ogden, F.L.; Zhu, J. Characterization of sudden and sustained base flow jump hydrologic behaviour in the humid seasonal tropics of the Panama Canal Watershed. *Hydrol. Process.* **2019**, *34*, 569–582. [CrossRef]
73. Cheng, Y.; Ogden, F.L.; Zhu, J.; Bretfeld, M. Land Use-Dependent Preferential Flow Paths Affect Hydrological Response of Steep Tropical Lowland Catchments with Saprolitic Soils. *Water Resour. Res.* **2018**, *54*, 5551–5566. [CrossRef]

74. Hou, Z.; Huang, M.; Leung, L.R.; Lin, G.; Ricciuto, D.M. Sensitivity of surface flux simulations to hydrologic parameters based on an uncertainty quantification framework applied to the Community Land Model. *J. Geophys. Res. Atmos.* **2012**, *117*, 1–18. [CrossRef]
75. Ren, H.; Hou, Z.; Huang, M.; Bao, J.; Sun, Y.; Tesfa, T.; Ruby Leung, L. Classification of hydrological parameter sensitivity and evaluation of parameter transferability across 431 US MOPEX basins. *J. Hydrol.* **2016**, *536*, 92–108. [CrossRef]
76. Gong, W.; Duan, Q.; Li, J.; Wang, C.; Di, Z.; Ye, A.; Miao, C.; Dai, Y. Multiobjective adaptive surrogate modeling-based optimization for parameter estimation of large, complex geophysical models. *Water Resour. Res.* **2016**, *52*, 1984–2008. [CrossRef]
77. Xia, Y.; Mocko, D.M.; Wang, S.; Pan, M.; Kumar, S.V.; Peters-Lidard, C.D.; Wei, H.; Wang, D.; Ek, M.B. Comprehensive Evaluation of the Variable Infiltration Capacity (VIC) Model in the North American Land Data Assimilation System. *J. Hydrometeorol.* **2018**, *19*, 1853–1879. [CrossRef]
78. Xu, T.; Guo, Z.; Xia, Y.; Ferreira, V.G.; Liu, S.; Wang, K.; Yao, Y.; Zhang, X.; Zhao, C. Evaluation of twelve evapotranspiration products from machine learning, remote sensing and land surface models over conterminous United States. *J. Hydrol.* **2019**, *578*, 124105. [CrossRef]
79. Baez-Villanueva, O.M.; Zambrano-Bigiarini, M.; Beck, H.E.; McNamara, I.; Ribbe, L.; Nauditt, A.; Birkel, C.; Verbist, K.; Giraldo-Osorio, J.D.; Thinh, N.X. RF-MEP: A novel random forest method for merging gridded precipitation products and ground-based measurements. *Remote Sens. Environ.* **2020**, *239*, 111606. [CrossRef]
80. Wei, L.; Jiang, S.; Ren, L.; Yuan, S.; Liu, Y.; Yang, X.; Wang, M.; Zhang, L.; Yu, H.; Duan, Z. An extended triple collocation method with maximized correlation for near global land precipitation fusion. *Geophys. Res. Lett.* **2023**, *50*, 2023GL105120. [CrossRef]