remote sensing

MDPI

Article

Comparison and Evaluation of Multi-Source Evapotranspiration Datasets in the Yarlung Zangbo River Basin

Yao Jiang [1,2], Zihao Xia [1], Lvyang Xiong [1,2,*] and Zongxue Xu [3]

[1] School of Civil Engineering and Architecture, Nanchang University, Nanchang 330031, China; jiangyao313@ncu.edu.cn (Y.J.); xia_zh123@email.ncu.edu.cn (Z.X.)
[2] Key Laboratory of Poyang Lake Environment and Resources Utilization, Ministry of Education, Nanchang University, Nanchang 330031, China
[3] College of Water Sciences, Beijing Normal University, Beijing 100875, China; zxxu@bnu.edu.cn
* Correspondence: xionglvyang@ncu.edu.cn

Highlights

What are the main findings?

- By comparing ten different ET datasets with ET estimates derived from the terrestrial water balance method in terms of spatiotemporal variations across the Yarlung Zangbo River basin, it was found that GLASS-ET and GLEAM-ET perform relatively well, whereas Han-ET and Chen-ET exhibits greater discrepancies in these aspects.

What is the implication of the main finding?

- Potential weaknesses are present in all ET datasets within high-altitude regions with complex terrain.
- The performance of ET datasets is highly dependent on the regional characteristics, algorithms and forcing data accuracy.

Abstract

Evapotranspiration (ET) data products has greatly facilitated the hydrological research in complex basins, and various ET datasets have been produced and applied. The applicability and reliability of ET dataset is significant for regional studies. Therefore, this study compared ET datasets from multisource remote sensing (GLEAM, MOD16, GLASS, PML-V2, Han, Chen and Ma), machine learning (Jung) and reanalysis products (ERA5-Land, MERRA2) for the Yarlung Zangbo River basin (YZB). ET was estimated using the terrestrial water balance (TWB) and was taken as baseline for comparisons of different ET datasets in terms of spatial distribution and temporal variation. Results indicate that (1) the TWB-based ET estimates are rational with acceptable uncertainties; (2) the multi-source ET datasets exhibit good correlations with TWB-ET across the entire basin (r = 0.78–0.90) in term of annual variation, with GLEAM-ET performing the best (r = 0.88, *RMSE* = 14.24 mm, *Rbias* = 18.55%); (3) Spatially, PML-ET and Ma-ET show higher consistency with TWB-ET, and temporally, MOD16-ET and GLASS-ET better capture the changing trend; (4) A comprehensive evaluation using the linear weighted method reveals that GLASS-ET and GLEAM-ET perform relatively well in all aspects and are reliable datasets for ET research in the YZB. These findings provide a scientific basis for ET estimation and data selection in the YZB, offering important references for ET analysis and hydrological research.

Keywords: evapotranspiration; terrestrial water balance; evapotranspiration products; GRACE; performance evaluation

Check for updates

Academic Editor: Liang Sun

Received: 22 October 2025
Revised: 24 December 2025
Accepted: 29 December 2025
Published: 4 January 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Evapotranspiration (ET) is a critical process of water and energy exchange among the hydrosphere, atmosphere, and biosphere, and constitutes an essential component of the regional water cycle [1,2]. The variations in ET not only directly influence runoff generation but also have significant impacts on regional climate regulation and vegetation productivity [3,4]. Therefore, accurate estimation of ET is essential for advancing understanding of hydrological processes [5]. However, conventional ground-based observation methods face limitations due to sparse spatial coverage and inadequate site representativeness, especially in high-altitude regions characterized by complex topography and climatic conditions, thereby making the accurate estimation of ET still a major challenge [6].

In recent years, remote sensing-based and reanalysis-driven ET data products have experienced rapid development, becoming important alternative methods for studying the spatial distribution and dynamic changes of regional ET, particularly in data-scarce regions [7]. Presently, the widely used ET datasets include: MOD16 ET products derived from Moderate Resolution Imaging Spectroradiometer (MODIS) satellites [8], GLEAM (Global Land Evaporation Amsterdam Model) products [9], PML-V2 (Penman Monteith Leuning-V2) [10], GLDAS-Noah (Global Land Data Assimilation System-Noah model) [11], and ERA5 (5th Generation of European Reanalysis-Land part) [12], among others. However, due to substantial differences in model structure, input data, and computational methods, these datasets exhibit considerable uncertainties in their ET estimation results across different regions and ecosystems [13,14]. For instance, Zuo et al. [15] compared and analyzed the performance of six ET data products in China, revealing significant differences among them. Therefore, a systematic evaluation of the applicability of different ET products in specific regions not only helps to clarify their accuracy characteristics and application scopes but also provides important references for ET data fusion and hydrological process simulation [16,17]. Particularly under the backdrop of intensifying global climate change and the increasingly pronounced water supply-demand imbalance, obtaining reliable and suitable ET data has become an urgent requirement for water resource management, ecosystem maintenance, and water security in complex basins [18].

The Yarlung Zangbo River is one of the major river systems on the Tibetan Plateau (TP) and an important transboundary river in South Asia. The Yarlung Zangbo River basin (YZB) is characterized by pronounced topographic undulations, a large altitude gradient, unique and highly variable climatic conditions, diverse vegetation types, and complex underlying surfaces. This complex environmental heterogeneity results in substantial spatial variations in ET within the basin, bringing considerable challenges for accurate ET estimation. Considering the distinct advantages and limitations of each product, along with their varying performances under specific environmental conditions, no consensus on exists regarding a universally superior product across all study areas [19,20].

Many studies have conducted evaluations of ET data in the TP. For example, the evaluation on eight major ET products by Meng et al. [21] found that CLM-BGCDV (Community Land Model–biogeochemical dynamic vegetation) better represented the response of ET to climate change by incorporating dynamic vegetation processes, whereas the ITP (Institution of Tibetan Plateau Research) products (i.e., Han-ET in this study) significantly overestimated and MOD16 underestimated the seasonal amplitude of ET. Li et al. [22] integrated terrestrial and atmospheric water balance approaches to evaluate four ET products (AVHRR: Advanced Very High Resolution Radiometer, GLEAM, MOD16, GLDAS-Noah) across the five major river basins of the TP. Their results revealed that in the YZB, all products exhibited substantial deviations from the water balance baseline, which was attributed to the general neglect of sublimation processes in current ET estimation methods. The study of Cheng et al. [23] concluded that remote sensing products exhibited significantly lower

uncertainty compared to land surface model and reanalysis products. They attributed performance differences to variations in algorithm structure and input data, noting that remote sensing products are better able to capture actual surface conditions, while model-based products are more sensitive to uncertainties in parameterization schemes. Liu [24] systematically evaluated three remote sensing ET products (GLEAM, ZHANG, CSIRO) across 16 river basins in the TP. Their findings indicated that in the YZB, all products generally overestimated summer ET and demonstrated weaker performance in capturing interannual variations than in simulating multi-year mean patterns. Product performance was influenced not only by climatic conditions but also by local surface characteristics and algorithmic structure. Li et al. [25] further demonstrated in their evaluation of ET products (JRA: Japanese 25 year Reanalysis, ZHANG, GLDAS, MODIS) in the TP that while the seasonal dynamics of each ET product were consistent, significant discrepancies in magnitude were presented. These deviations mainly originated from uncertainties in input data--for example, the overestimation of ET by MODIS in the upper reaches of the Yellow River resulted from excessive short-wave radiation inputs. From these studies, it is evident that the performance of ET products varies significantly across different regions of the TP due to the combined influence of regional climate conditions, underlying surface characteristics, algorithms, and driving data, with different dominant factors across regions. Substantial uncertainty exists among the various ET products, especially in areas with large glacier cover. However, systematic comparative analyses specifically focused on the YZB, a distinct geographical unit, remain limited. In particular, existing evaluation studies mostly focus on the overall scale of the TP, making it difficult to precisely capture the spatial variability of ET within the basin due to differences in topography, climate, and underlying surface characteristics. This limitation thus hinders more in-depth and detailed hydrological researches in the YZB.

Consequently, this study aims to conduct a systematic evaluation of the performance of multi-source ET datasets in the YZB. It will thoroughly investigate the accuracy and applicability of different ET datasets in the YZB, including terrestrial water balance-based estimates, remote sensing-based inversion products, and reanalysis products. The specific research contents are as follows: (1) estimating ET in the YZB using GRACE data and the terrestrial water balance method (TWB); (2) comparing the spatio-temporal variation characteristics of different ET datasets within the YZB; (3) quantitively evaluating the performance of these ET datasets in the YZB based on TWB-ET. The results of this study can provide reliable data support for the ET and hydrology research in the YZB, and offer valuable references for the selection and fusion of ET datasets in high-altitude and complex regions.

2. Material and Methods

2.1. Study Area

The YZB is located in southern TP, encompassing an area of 0.24 million km^2 within China (Figure 1). It is one of the world's highest river basins, with an average elevation of over 4000 m. The Yarlung Zangbo River flows from west to east, receiving many tributaries along its course, including major ones such as the Nianchu River, Lhasa River, Nyang River, and Parlung Zangbo River (Figure 1). The climate is influenced by warm and moist air currents from the Bay of Bengal, westerly circulation patterns, and the unique plateau geography. The climate thus exhibits significant regional variability, with average annual temperature decreasing from 5–9 °C in the downstream (east) to 0–3 °C in the upstream (west). Annual precipitation also varies markedly, ranging from over 1000 mm in the downstream to 200–400 mm in the upstream [26]. Additionally, the YZB is characterized by

intense solar radiation, low atmospheric pressure, low humidity, and a clear alternation of dry and wet seasons.

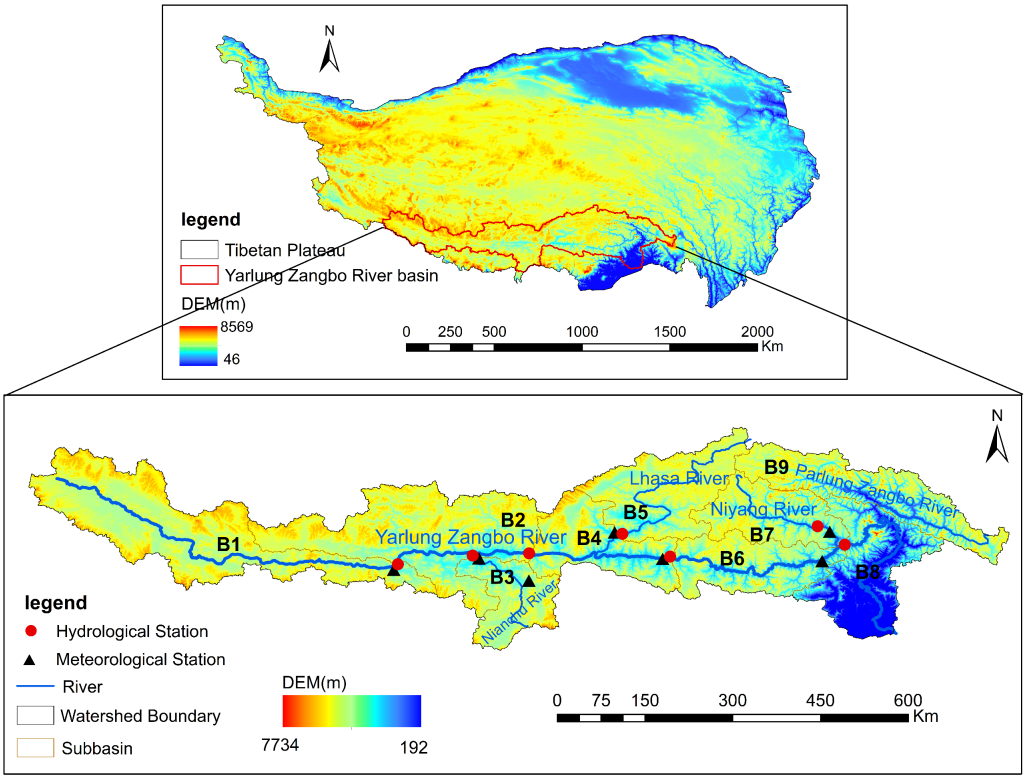


Figure 1. Geographic location of Yarlung Zangbo River basin (the number in the figure represents different sub-basins in Table 1).

Table 1. Basic information of 9 sub-basins in the Yarlung Zangbo River basin.

	Sub-Basin	Name	River	Hydrological Station
Upper reaches	B1	Lhatse	Yarlung Zangbo River	Lhatse
Middle reaches	B2	Nugesha	Yarlung Zangbo River	Nugesha
	B3	Shigatse	Nianchu River	Shigatse
	B4	Yangcun	Yarlung Zangbo River	Yangcun
	B5	Lhasa	Lhasa River	Lhasa
	B6	Nuxia	Yarlung Zangbo River	Nuxia
	B7	Gengzhang	Nyang River	Gengzhang
Lower reaches	B8	Motuo	Yarlung Zangbo River	
	B9	Parlung	Parlung Zangbo River	

Based on the distribution of the rivers and hydrological stations within the basin, this study divided the YZB into 9 sub-basins (Figure 1 and Table 1). Sub-basin B1 represents the upper reaches of the YZB, controlled by the Lhatse hydrological station. Sub-basins B2 to B5 are located in the middle reaches of the YZB, with runoff monitored by the hydrological stations of Nugesha, Yangcun, and Nuxia, respectively. The B3, B6 and B7 correspond to the Nianchu River basin, the Lhasa River basin, and the Nyang River basin, respectively. Sub-basins B8 and B9 are situated in the lower reaches of the YZB (Table 1). Due to the lack of runoff data for the lower reaches (B8 and B9), only the middle and upper reaches of the YZB (i.e., B1–B7) were included in the ET estimation using the TWB method and in the quantitative evaluation of different datasets.

2.2. Data Description

2.2.1. Precipitation and Runoff

Precipitation data used the China Meteorological Forcing Dataset (CMFD) [27] obtained from the National Tibetan Plateau Data Center (http://data.tpdc.ac.cn), with a spatial resolution of 0.1° and a temporal resolution of 3 h (Table 2). The dataset used in this study spans from 2003 to 2015 and was aggregated into monthly estimates. Additionally, observed precipitation data (Table 2) were collected from 11 meteorological stations located in the YZB (Figure 1). These observation data were utilized to calibrate the CMFD data, thereby improving the accuracy of precipitation data.

Runoff data were obtained from the daily measured runoff volumes at 7 hydrological stations (Figure 1), provided by the Water Resources Survey Bureau of the Tibet Autonomous Region, China. These sequences encompass the period of 2003–2015 without any missing data segments. The runoff depth in each sub-basin was calculated by dividing the total monthly outflow volume by the corresponding catchment area.

2.2.2. Terrestrial Water Storage

The terrestrial water storage (TWS) change data were obtained from the China Regional Precipitation-based Reconstructed Terrestrial Water Storage Change Dataset (2002–2019) [28], which has a spatial resolution of 0.25° and a temporal resolution of one month (Table 2). The data was generated by constructing a precipitation reconstruction model that takes into account the seasonal and trend terms of CSR RL06 Mascon products, by integrating the GRACE data from CSR GRACE/GRACE-FO RL06 Mascon solutions (CSR-M), China Gauge-based Daily Precipitation Analysis (CGDPA, version 1.0) data, and CN05.1 temperature dataset. To address missing data in certain months and ensure continuous time series, an interpolation method was applied, using the average values of adjacent months to fill in the gaps.

2.2.3. Evapotranspiration

Ten ET datasets (Table 2) were selected and evaluated in this study, as follows:

(1) ERA5-ET is derived from the ERA5-Land reanalysis data published by the European Centre for Medium Range Weather Forecasts (ECMWF), featuring improvement of land ET estimation and higher resolution [12].

(2) GLEAM-ET is the GLEAM product generated by applying the Priestley-Taylor equation and introducing the key stress factors of soil moisture [9].

(3) MERRA2-ET is the reanalysis dataset from the Modern-Era Retrospective analysis for Research and Applications-Version 2 (MERRA2) [29,30], which is a global-scale estimations using the data assimilation system combined with numerical models and multiple source observational data.

(4) MOD16-ET is a MODIS standard product that includes surface ET, latent heat flux, potential ET, and potential latent heat flux, calculated using the Penman-Monteith equation along with daily meteorological data and remote sensing inputs such as vegetation coverage and albedo [8].

(5) GLASS-ET is the Global Land Surface Satellite (GLASS) products derived from the National Earth System Science Data Center, National Science & Technology Infrastructure of China (https://www.geodata.cn/main/). It is produced by Yao et al. [31], integrating five traditional ET algorithms based on multi-source remote sensing data and ground observations;

(6) PML-ET is a product of the PML-V2 model [10], which integrates the photosynthesis model with the improved canopy stomatal conductance model based on the Penman-

Monteith equation, and utilizes meteorological data from GLDAS 2.1 and other inputs from MODIS (e.g., leaf area index, albedo, and emissivity) [32].

(7) Ma-ET is derived from the terrestrial evapotranspiration dataset across China [33], generated using the complementary-relationship method with inputs of CMFD and multi-source remote sensing data such as GLASS albedo and longwave emissivity, ERA5-Land surface temperature and humidity [34].

(8) Han-ET is the monthly mean evapotranspiration data set of the Tibet Plateau [35], calculated using the surface energy balance system model (SEBS) with the MODIS and CMFD inputs [36].

(9) Chen-ET is the daily evapotranspiration data for the southwest source region of China, based on the surface flux equilibrium (SFE) estimates of the Bowen ratio, and combined with satellite observations of surface net radiation, near-surface air temperature and humidity [37].

(10) Jung-ET is a global-scale land surface flux dataset that integrates FLUXNET eddy-covariance observations with remote sensing and meteorological data using machine learning methods [38].

For comparative analysis, ET data during 2003–2015 were selected for each source, processed into monthly and annual values, and standardized to units of mm.

Table 2. The original information of multi-source data in the study.

Variable	Data	Spatial Resolution	Temporal Resolution	Time Span	Reference
Precipitation	China meteorological forcing dataset	0.1°	3 h	1951–2024	He et al. [27]
	Observation data at meteorological station	/	daily	1961–2015	/
Runoff	Observation data at hydrological station	/	daily	1961–2015	/
Terrestrial water storage	Dataset of reconstructed terrestrial water storage in China based on precipitation	0.25°	monthly	2002–2019	Zhong et al. [28]
Evapotranspiration	ERA5	0.1°	monthly	1981–2018	Muñoz Sabater [12]
	GLEAM	0.25°	monthly	1980–2018	Martens et al. [9]
	MERRA2	0.5°	monthly	2000–2018	GMAO [30]
	MOD16	0.05°	8 d	2000–2018	Running et al. [8]
	GLASS	0.05°	8 d	2000–2018	Yao et al. [31]
	PML	0.05°	8 d	2002–2019	Zhang and He [10]
	Ma	0.1°	monthly	1982–2015	Ma et al. [33]
	Han	0.1°	monthly	2001–2018	Han et al. [35]
	Chen	0.1°	daily	1979–2018	Chen et al. [37]
	Jung	0.5°	monthly	2001–2015	Jung et al. [38]

2.3. Method

2.3.1. Terrestrial Water Balance and Uncertainty

The terrestrial water balance (TWB) is one of the most direct and effective methods for estimating actual ET at basin scales. The equation is as follows:

$$ET = P - R - \Delta S/\Delta t \quad (1)$$

where P is monthly precipitation (mm), R is monthly runoff depth (mm), and $\Delta S/\Delta t$ is the monthly change of TWS (mm). The time interval Δt is one month, which is consistent with the temporal resolution of GRACE data. This study first estimates $\Delta S/\Delta t$ using GRACE data, and then estimates the monthly ET for the YZB according to Equation (1).

To determine the confidence level of ET estimation, it is essential to quantify the uncertainty associated with the ET estimates derived from the TWB (TWB-ET). In the TWB framework, all components (P, R, and $\Delta S/\Delta t$) are considered independent observations. The uncertainties in TWB-ET can be determined by calculating the root sum square of the uncertainties in each variable, in accordance with the principles of error propagation (Equation (2)):

$$\sigma_{ET} = \sqrt{{\sigma_P}^2 + {\sigma_R}^2 + \sigma^2_{\Delta S/\Delta t}} \tag{2}$$

where σ is the estimated uncertainty of the corresponding variable.

Uncertainty in precipitation of the CMFD product can be estimated from the evaluation results of precipitation products in the Tibetan Plateau by Cheng et al. [39]. A basic relative uncertainty of 14% was assigned to precipitation in the YZB in this study. Runoff observation data from stations are generally reliable and were assumed to have the least uncertainty, with 5% of observational errors [4,11,40]. The uncertainty in GRACE-derived TWS changes was estimated according to the results of Li et al. [22], who quantified the total uncertainty in GRACE-derived TWS changes as 0.12 mm/month in the YZB. More details on uncertainty estimation can be found in the Supporting Information.

2.3.2. Mann-Kendall Trend Test

The Mann-Kendall (MK) trend test was utilized to assess the multi-year (2003–2015) change trends and their significance for each grid of different ET datasets as well as for TWB-ET. The rate of change was estimated using Sen's slope method. The detailed description of the MK test and Sen's slope method refers to Jiang et al. [26]. In this study, integers ranging from -3 to 3 were used to represent the trend of ET changes at a given confidence level α, where -3 indicates strong significant downtrend ($\alpha = 0.01$), -2 indicates significant downtrend ($\alpha = 0.05$), -1, 0, and 1 indicates no significant trend (non-significant decrease, no trend, non-significant increase), 2 indicates significant uptrend ($\alpha = 0.05$), and 3 indicates strong significant uptrend ($\alpha = 0.01$).

2.3.3. Evaluation Method

To quantify the performance of different ET datasets in the YZB, a quantitative evaluation was conducted for the selected ten ET datasets and TWB-ET using the linear weighted method, with considering the aspects of spatial variation range, basin-average values, spatial distribution, and spatio-temporal variation. Three evaluation metrics were employed for the comparison: the correlation coefficient (r), root mean square error ($RMSE$), and relative bias ($Rbias$), as detailed in the following equations.

$$r = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2}} \tag{3}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(x_i - y_i)^2}{n}} \tag{4}$$

$$Rbias = \frac{\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} y_i} \times 100\% \tag{5}$$

where x_i represents the variable to be evaluated, y_i represents the variable taken as the baseline, an overbar represents the mean value, and n is number of data pairs.

Specifically, the various ET datasets were compared in each aspect, and their agreement with the TWB-ET was quantically evaluated. Based on the comparison results, characteristic metrics of ET data were quantified using a standardized scale from 0 to 1: for the basin average and variation range, datasets within this range were assigned a value of 1.0, while those outside this range were assigned a value of 0; Regarding the spatial distribution and temporal variation, the actual evaluation values for each dataset were normalized between 0 and 1 based on their respective maximum and minimum values. Finally, a comprehensive score was computed to evaluate the performance of different ET datasets. The comprehensive score was calculated using linear weighted method, as follows:

$$Q = \sum_{i=1}^{n} w_i a_i \tag{6}$$

where Q is comprehensive score, w_i is weight of the ith characteristic item, a_i is value of the ith characteristic item.

3. Results Analysis

3.1. TWB-Based ET Estimation and Uncertainty

The monthly TWB-ET changes in each sub-basin of the YZB during 2003–2015 are illustrated in Figure 2. Due to the lack of runoff data, TWB-ET estimates for sub-basins B8 and B9 are unavailable. The maximum monthly TWB-ET in the YZB typically occurs from June to September, while the minimum values are mainly observed from November to February of the following year. The basin-average monthly TWB-ET across sub-basins ranges from 25.0 to 34.9 mm/month, with the highest value of 34.9 mm/month recorded in sub-basin B7 and the lowest value of 25.0 mm/month in sub-basin B1 (Table 3). In addition, negative TWB-ET values (as low as −8 mm/month in February) exist during certain winter months, mainly due to large underestimation in precipitation within such a complex and narrow region with large snow cover area [41]. This issue has also been reported in previous studies [22,40]. Whereases, these negative values mainly occur in winter in specific sub-basins (B2, B4 and B5), and their absolute values are relatively small. To minimize the impact, average monthly TWB-ET or annual values were used in the comparisons. Generally, TWB-ET exhibits a decreasing trend from downstream to upstream sub-basins. Sub-basins B4 and B5 have relatively lower average monthly TWB-ET due to more frequent negative values in January and February.

The uncertainties in monthly TWB-ET for each sub-basin, calculated according to Equation (2), are also presented in Figure 2 and Table 3. The uncertainty ranges from 3.25 to 7.65 mm/month, with the lowest observed in B1 and the highest in B7. The relative uncertainty (σ_{ET}/TWB-ET) is approximately 12–22% with an average of around 16% (Table 3), well within the acceptable range (<30%) as defined by Li et al. [22]. Over the period from 2003 to 2015, the average annual TWB-ET in the YZB was approximately 450 mm. Spatially, the maximum annual TWB-ET of 419 ± 69 mm/year was recorded in B7, while the minimum value of 249 ± 41 mm/year was observed in B1. For other sub-basins, the annual TWB-ET was approximately 300 mm.

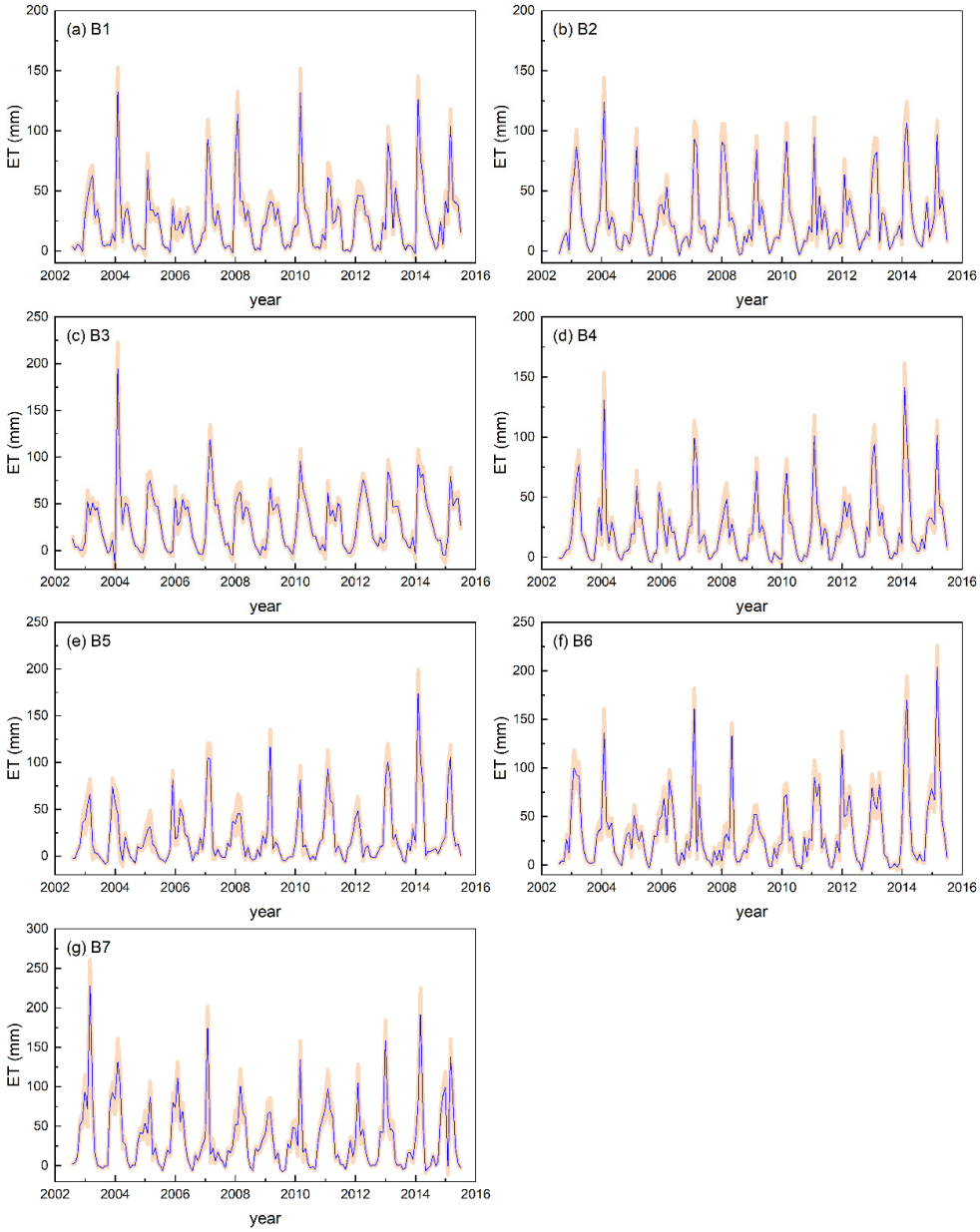


Figure 2. Evapotranspiration estimated by terrestrial water balance (TWB-ET) in the Yarlung Zangbo River basin (The blue line indicates ET, and the shadow shows uncertainty ranges).

Table 3. Mean monthly uncertainty estimates of each component in terrestrial water balance as well as mean monthly TWB-ET in each sub-basins of the Yarlung Zangbo River basin.

Variable	B1	B2	B3	B4	B5	B6	B7
σ_P (mm/month)	3.22	3.72	3.47	3.79	4.78	5.62	7.37
σ_R (mm/month)	0.39	0.60	0.30	0.81	1.38	1.26	2.03
$\sigma_{\Delta S/\Delta t}$ (mm/month)	0.12	0.12	0.12	0.12	0.12	0.12	0.12
σ_{ET} (mm/month)	3.25	3.77	3.49	3.88	4.98	5.76	7.65
TWB-ET (mm/month)	25.01	25.88	29.76	27.07	25.45	33.09	34.89
σ_{ET}/TWB-ET (%)	13.0	14.6	11.7	14.3	19.5	17.4	21.9

Note: B1–B7 is the sub-basins of the Yarlung Zangbo River basin as shown in Figure 1 and Table 1; TWB-ET is evapotranspiration estimated by terrestrial water balance.

3.2. Correlation of Multi-Source ET with TWB-ET

The ten ET datasets were individually compared against TWB-ET in each sub-basin and the entire basin scale in the term of annual variation, with the results presented in Figure 3 and Table 4. Figure 3 shows the comparisons of multi-year average monthly ET from these ten ET datasets with TWB-ET. The R^2 indicates the similarity of spatio-temporal distribution between the two sets of ET data. Overall, the ten ET datasets exhibit good correlations with TWB-ET, with *r* values ranging from 0.78 to 0.90, and six of the datasets achieving an *r* above 0.85, despite some biases (Figure 3). When looking at the comparisons on sub-basin scale, differences become obvious among different ET datasets. For example, in the B2 sub-basin, GLEAM-ET exhibits the smallest *Rbias* (<5%) and *RMSE* (<9 mm/month) compared to TWB-ET, with an *r* of 0.88. In contrast, MOD16-ET show relatively poor correlations with TWB-ET, with *r* values of 0.78 and large bias in this region (Table 4). Overall, for annual variation, GLEAM-ET and MERRA2-ET show better correlations with TWB-ET, with higher average *r* (0.88 and 0.90), the lowest *RMSE* (<15 mm/month), and lower *Rbias* (18.5% and 29.9%) (Figure 3).

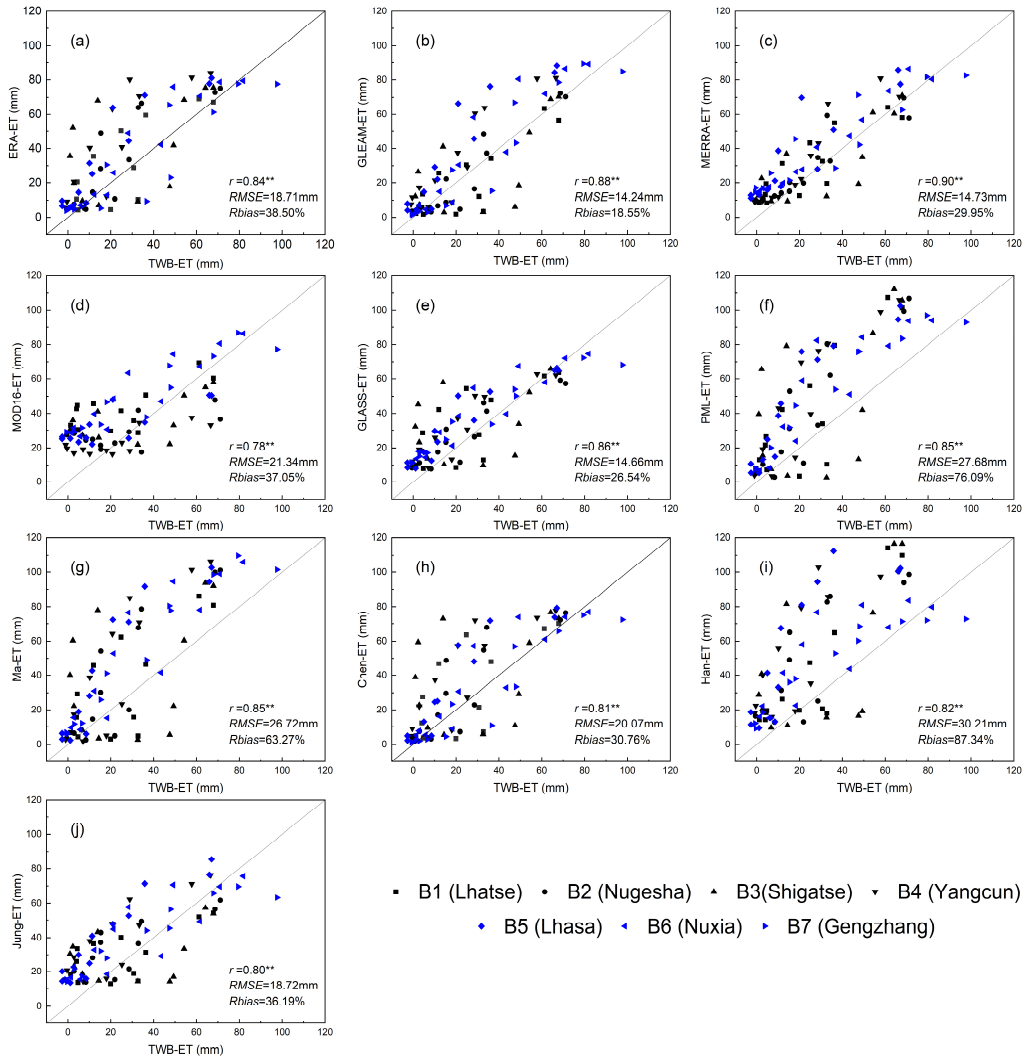


Figure 3. Scatter plot of ET estimates from 10 ET datasets versus TWB-ET in the Yarlung Zangbo River basin ((**a**–**j**) respectively represent ten types of ET datasets, the number of data pairs is 84 for each dataset, the meaning of ** is the same as that in Table 4).

Table 4. Evaluation metrics for comparison of different ET datasets with TWB-ET.

Dataset	B1	B2	B3	B4	B5	B6	B7	Average
				r				
ERA5-ET	0.80 **	0.85 **	0.51	0.85 **	0.91 **	0.95 **	0.90 **	0.84 **
GLEAM-ET	0.83 **	0.94 **	0.65 *	0.92 **	0.92 **	0.94 **	0.95 **	0.88 **
MERRA2-ET	0.86 **	0.92 **	0.80 **	0.94 **	0.89 **	0.99 **	0.92 **	0.90 **
MOD16-ET	0.69 *	0.66 *	0.60 *	0.79 **	0.82 **	0.93 **	0.94 **	0.78 **
GLASS-ET	0.73 **	0.92 **	0.47	0.92 **	0.94 **	0.93 **	0.97 **	0.86 **
PML-ET	0.83 **	0.91 **	0.54	0.91 **	0.92 **	0.89 **	0.96 **	0.85 **
Ma-ET	0.73 **	0.88 **	0.39	0.88 **	0.91 **	0.90 **	0.97 **	0.85 **
Chen-ET	0.70 *	0.85 **	0.35	0.84 **	0.92 **	0.91 **	0.94 **	0.81 **
Han-ET	0.89 **	0.81 **	0.60 *	0.80 **	0.84 **	0.83 **	0.93 **	0.82 **
Jung-ET	0.65 *	0.82 **	0.39	0.87 **	0.94 **	0.84 **	0.95 **	0.80 **
				RMSE (mm/month)				
ERA5-ET	15.47	17.18	27.16	26.32	19.75	10.81	14.28	18.71
GLEAM-ET	12.84	8.87	20.80	17.29	21.43	14.54	10.92	14.24
MERRA2-ET	12.15	9.76	14.91	13.56	20.02	10.42	15.28	14.73
MOD16-ET	23.60	17.99	20.75	15.78	19.84	21.75	29.70	21.34
GLASS-ET	16.69	9.59	24.20	12.41	13.73	13.39	12.59	14.66
PML-ET	26.99	24.46	37.57	31.36	30.99	24.44	17.94	27.68
Ma-ET	21.45	24.50	34.63	30.51	30.61	25.22	20.13	26.72
Chen-ET	19.43	16.58	30.76	31.36	17.77	12.36	12.26	20.07
Han-ET	24.34	30.52	35.34	39.70	42.78	23.21	15.60	30.21
Jung-ET	16.78	14.79	23.31	19.39	24.85	16.90	15.04	18.72
				Rbias (%)				
ERA5-ET	22.22	38.97	34.62	90.74	76.65	18.01	−11.68	38.50
GLEAM-ET	−16.10	−3.70	−13.05	57.97	79.12	24.77	0.86	18.55
MERRA2-ET	22.02	14.87	1.17	44.35	77.45	28.13	21.67	29.95
MOD16-ET	70.77	9.77	20.30	2.30	56.56	57.31	42.37	37.05
GLASS-ET	27.66	9.45	18.99	40.95	53.04	26.31	9.37	26.54
PML-ET	67.44	64.94	60.60	109.09	128.49	57.93	44.18	76.09
Ma-ET	33.67	55.72	36.82	93.97	116.39	58.15	48.17	63.27
Chen-ET	25.28	32.44	29.83	67.29	63.67	13.31	−16.53	30.76
Han-ET	62.38	90.71	54.98	149.81	174.77	53.15	25.58	87.34
Jung-ET	17.36	28.05	−0.64	71.26	97.82	26.59	12.87	36.19

Note: B1–B7 is the sub-basins of the Yarlung Zangbo River basin as shown in Figure 1 and Table 1. The parameter n in Equations (3)–(5) is 12 for each sub-basin and 84 for entire basin. The degree of freedom (df) for r is 10 for each sub-basin and 82 for entire basin, respectively. The * and ** indicate that the p value is less than 0.05 and 0.01, respectively.

3.3. Comparison of Spatial Distribution of Multi-Source ET

The multi-year mean ET values within the grids of each dataset were extracted and calculated, and the variation ranges of these datasets in the YZB are shown in Figure 4. The multi-year mean ET values for most datasets in the YZB fall within the range of 400–510 mm, with only two exceptions: GLEAM-ET is below 400 mm, while Han-ET exceeds 510 mm. This result is also consistent with the result demonstrated by Fan et al. [42] and Yuan et al. [43]. The spatial maximum values (99%) of most ET datasets in the YZB are approximately 900–1100 mm, and the spatial minimum values (1%) are mainly between 190–300 mm. However, ERA5-ET, PML-ET, and Chen-ET exhibit lower maximum ET values, while GLASS-ET, PML-ET, and Jung-ET show higher minimum ET values. Notably, MERRA2-ET has the largest difference between its maximum and minimum values, ranging from 235 to 1136 mm, whereas Chen-ET has the smallest difference (216–590 mm). In term of the main variation range (25–75%), PML-ET exhibits the smallest spatial variabil-

ity, spanning only 470–543 mm, while MOD16A-ET shows the largest spatial variability, ranging from 347 to 687 mm.

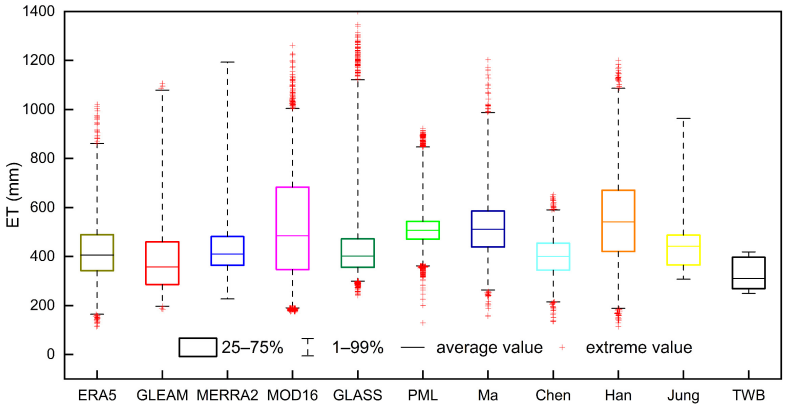


Figure 4. Spatial variation range of multi-year average value from TWB-ET and different ET datasets in the Yarlung Zangbo River basin.

As shown in the spatial distribution of multi-year mean ET from different datasets (Figure 5), the overall spatial patterns of ET across different dataset in the YZB are largely consistent, showing a general trend of decreasing from southeast to northwest. The highest ET values are predominantly observed in the southeastern part of the YZB (B8), while the lowest values are mainly concentrated in the upper reaches of the basin (B1). However, there are notable discrepancies among datasets in the specific variations within individual sub-basins. For example, in sub-basin B2, GLEAM-ET exhibits lower annual ET with an average value of 300 mm, while Han-ET shows significantly higher annual ET with an average of 592 mm. In sub-basin B8, Chen-ET estimates annual ET at 453 mm, significantly lower than MOD16-ET (828 mm) (Table 5). Compared to TWB-ET in sub-basins B1–B7, the correlations of most ET datasets are larger than 0.5, PML-ET and Ma-ET show the largest correlation, whereas Chen-ET exhibits the weakest correlation (Table 5).

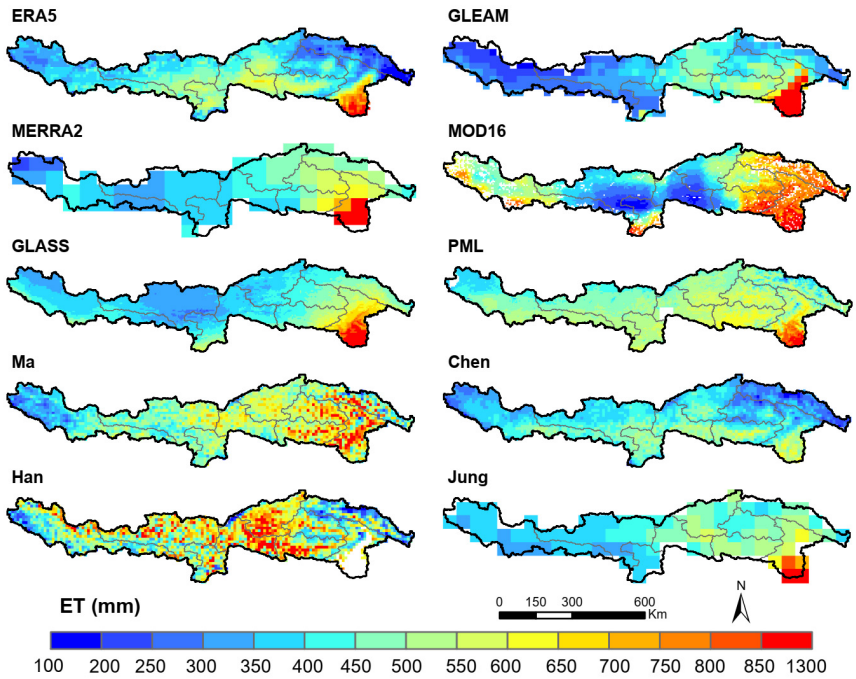


Figure 5. Spatial distribution of multi-year average ET from different sources in the Yarlung Zangbo River basin.

Table 5. Average annual ET from different datasets and TWB-ET in each sub-basin of the YZB, the unit is mm.

Dataset	B1	B2	B3	B4	B5	B6	B7	B8	B9	*r*
TWB	300	310	357	369	349	397	418			
ERA5	366	431	480	513	440	468	369	563	257	0.57
GLEAM	251	299	310	425	446	495	422	767	357	0.78
MERRA2	366	356	361	388	542	508	509	780	491	0.64
MOD16	512	340	429	275	390	624	596	828	720	0.48
GLASS	383	339	424	379	381	501	457	750	474	0.79
PML	502	512	573	563	569	627	603	663	531	0.94
Ma	401	483	488	522	539	627	620	690	544	0.92
Chen	375	411	463	450	407	449	349	453	287	0
Han	487	592	553	672	684	608	525	494	394	0.14
Jung	352	397	354	461	492	502	472	685	448	0.68

3.4. Comparison of Temporal Variation of Multi-Source ET

Whether the ET dataset can reliably reflect the temporal variations significantly impacts ET analysis in the YZB. Therefore, it is essential to compare the changing trends of ET from different sources. The multi-year (2003–2015) ET trends within each grid of the ET dataset were first calculated using the MK test, and the results are presented in Figure 6. Most ET datasets indicates an increasing trend in the region east of Yangcun station (i.e., B6–B9 sub-basins) (Figure 6). However, there are notable spatial differences in the specific regions of changes and their significance levels, as shown in Figure 6. For instance, ERA5-ET and Chen-ET show significant increases in B5 sub-basins and surrounding areas, whereas Ma-ET shows a decrease in this region. GLASS-ET and PML-ET indicate that ET is primarily increasing across the entire YZB, with significant increases mainly concentrated in the middle-lower reaches. Conversely, MERRA2-ET demonstrates a decreasing trend in the middle-lower reaches, which contrasts with the distribution observed in other datasets.

Table 6 summarizes the interannual changing trends of ET among ten ET datasets and TWB-ET. TWB-ET shows a non-significant upward trend in all sub-basins except sub-basin B7. GLASS-ET and PML-ET show consistent trends with TWB-ET in most sub-basins (B1–B6), but their significance levels differ, with PML-ET displaying particularly strong significance in these regions. ERA5-ET and MOD16-ET also perform relatively well, matching both the changing trend and significance level of TWB-ET in four sub-basins. In contrast, MERRA2-ET, Ma-ET and Chen-ET exhibit different trends compared to TWB-ET across most regions. For instance, Chen-ET shows a decreasing trend in sub-basins B1–B4, whereas TWB-ET indicates an insignificant upward trend. For quantitative evaluation, the mean bias relative to TWB-ET was calculated for each dataset according to the trend level (i.e., integers ranging from −3 to 3, as defined in Section 2.3.2) in each sub-basin (Table 6). Overall, MOD16-ET has the smallest mean bias, indicating better consistency with TWB-ET in terms of changing trend, while Chen-ET shows a larger discrepancy from TWB-ET in this regard.

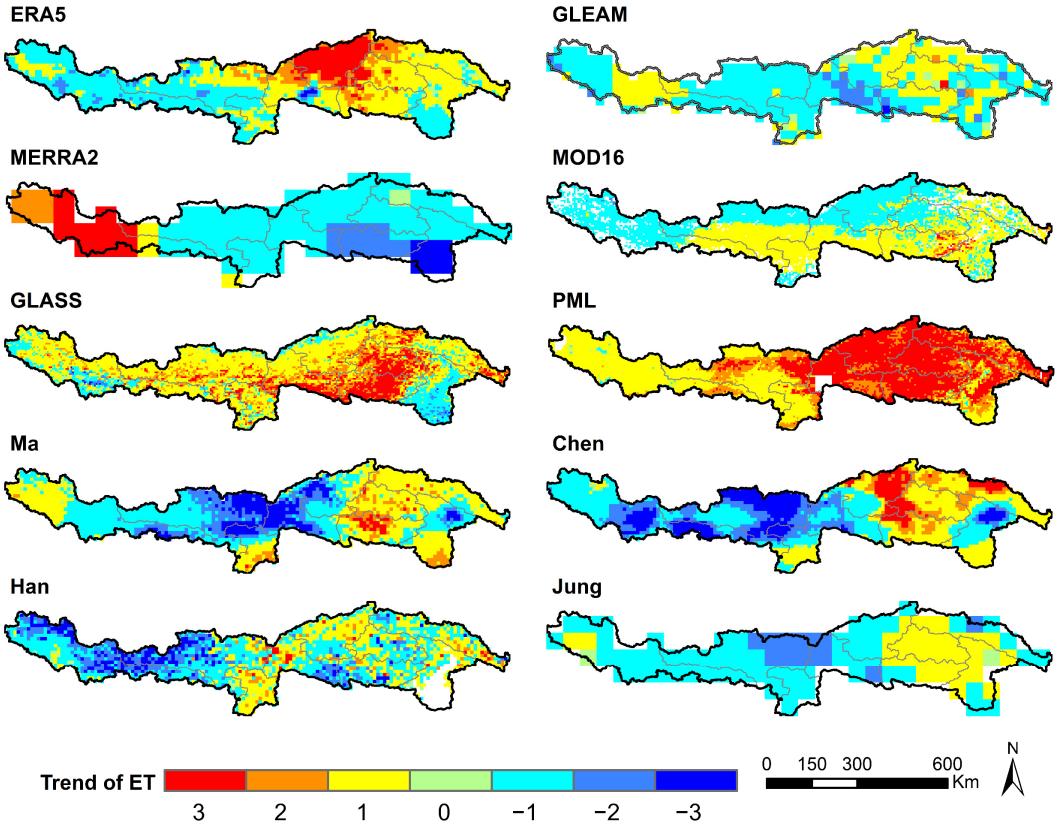


Figure 6. Spatial distribution of changing trend of annual ET from different sources in the Yarlung Zangbo River basin during 2003–2015. *Note*: −3 represents strong significant downtrend, −2 represents significant downtrend, −1 represents non-significant downtrend, 0 represents no trend, 1 represents non-significant uptrend, 2 represents significant uptrend, 3 represents strong significant uptrend.

Table 6. Multi-year changing rate of ET from different datasets and TWB-ET using Sen' slope method, the unit is mm/year.

Dataset	B1	B2	B3	B4	B5	B6	B7	B8	B9	Mean Bias
TWB	5.88	2.32	2.97	9.15	11.80	5.28	−6.45	-	-	-
ERA5	−1.08	−0.51	0.26	0.76	5.63 **	1.75	3.26 *	−0.03	1.39	1.28
GLEAM	−1.54	−4.69	−2.61	−3.81 *	0.39	0.01	1.32	−1.08	0.47	1.71
MERRA2	12.34 **	−3.41	−4.30	−4.13 *	−4.51	−4.31 *	−2.79	−5.73 **	−1.24	2.00
MOD16	−3.28	−0.28	1.05	0.73	−1.66	2.07	−1.72	1.83	1.36	0.86
GLASS	0.53	1.63	1.32	2.22 *	1.17	3.36 **	2.81 **	−1.94	1.34 *	1.00
PML	3.87	3.95 *	4.54 **	8.87 **	12.21 **	11.93 **	12.86 **	10.54 *	10.50 **	1.86
Ma	−2.92	−9.45 *	−1.03	−8.05 **	−3.61 *	2.85	3.90	4.06	3.45	2.28
Chen	−2.89 **	−3.88 *	−1.42	−2.52 *	2.08 **	2.65	2.91 *	−0.18	1.55 *	2.43
Han	−15.98 *	−8.01 **	7.01	0.48	0.44	−4.19 *	−0.65	0.38	−1.04	1.43
Jung	−0.64	−1.79	−1.29	−1.38 *	−0.15	0.31	0.90	0.40	0.01	1.86

Note: * indicates significant trend at a 0.05 confidence level, ** indicates significant trend at a 0.01 confidence level, MB is mean bias.

3.5. Evaluation of Multi-Source ET Datasets

To quantitatively evaluate the performance of ET datasets in the YZB, we selected several characteristic items including spatial variation range, basin average, spatial distribution, interannual variation, and correlation with TWB-ET. These items were quantified using a standardized scale of 0 to 1 according to their statistical values obtained by comparing with the TWB-ET. According the method described in the Section 2.3.3, the quantified

values of these characteristic items, along with the comprehensive scores for each dataset, are presented in Table 7.

Table 7. Comprehensive evaluation of different ET datasets in Yarlung Zangbo River basin.

Characteristic Term	Weight	ERA5	GLEAM	MERRA2	MOD16	GLASS	PML	Ma	Chen	Han	Jung
Correlation with TWB-ET	0.3	0.63	0.93	0.92	0.42	0.83	0.45	0.38	0.56	0.11	0.53
Changing trend	0.3	0.73	0.45	0.27	1.0	0.91	0.36	0.09	0	0.64	0.36
Spatial distribution	0.3	0.61	0.84	0.69	0.51	0.85	1.0	0.95	0	0.15	0.73
Basin average	0.05	1.0	0	1.0	1.0	1.0	1.0	0	1.0	0	1.0
Variation range	0.05	0.0	1.0	0	1.0	0	0	1.0	0	1.0	0.0
Comprehensive score	1	0.64	0.72	0.61	0.68	0.83	0.59	0.47	0.22	0.32	0.53

GLEAM-ET and MERRA2-ET in the YZB show better correlations with TWB-ET in terms of annual variation, while MERRA2-ET perform weaker interannual changing trend. In contrast, MOD16-ET and GLASS-ET achieve higher scores regarding changing trend, indicating better consistency with TWB-ET in the YZB. For spatial distribution, GLEAM-ET and GLASS-ET perform relatively well and rank highest. Whereas, GLEAM-ET has the lowest basin-average ET value, while GLASS-ET shows larger differences in variation range. According to the comprehensive scores, GLASS-ET and GLEAM-ET perform relatively superior overall performance in the YZB when comprehensively considering multiple aspects, whereas Han-ET and Chen-ET shows greater differences from other datasets across various aspects within the basin.

The result is consistent with that of Meng et al. [21], which ranked GLASS, GLEAM, and CLM-BGCDV as the top three among eight ET products based on data from 13 stations, whereas the ITP (i.e., Han-ET) significantly overestimated and MOD16 underestimated the seasonal amplitude of ET. Li et al. [22], by combining TWB and atmospheric water balance methods, found that in the YZB, all four ET products exhibited substantial deviations from the water balance baseline. Nevertheless, GLEAM had the highest Kling-Gupta efficiency (*KGE*) value, and MOD16 showed the highest *r*, which are also largely consistent with our findings. Additional studies across the TP [23,24] have similarly supported the relative superior performance of GLEAM. Compared with previous studies that mainly evaluated ET products at the overall scale of the YZB, our study performs a sub-basin-scale evaluation through a more refined analysis, thereby offering further verification and supplementation of earlier researches.

4. Discussion

4.1. Performance Differences

The comparisons indicates that notable differences exist in the performance of the ten datasets in the YZB. These differences primarily stem from the essential distinctions in model algorithms, input data, and the parameterization of key processes. Particularly, such differences become more pronounced in high-altitude regions with complex terrain, such as the YZB.

Specifically, the reanalysis products (ERA5-ET, MERRA2-ET) exhibit moderate overall performance, but perform relatively poorly in the high-altitude western regions. This may be attributed to inherent characteristics of the reanalysis product themselves and the complex terrain and climate conditions of the YZB. Their ET estimations rely on the assimilated meteorological fields. However, due to sparse ground-based observations and limited satellite data in the YZB, combined with complex terrain, the assimilation constraints are insufficient, potentially resulting in possible large errors in the meteorological forcing data. In comparison, ERA5-ET performs better, whereas the changing trend of MERRA2-ET

differed significantly from other datasets. This may be because MERRA2 simulates and outputs ET through its land surface process module (Catchment CN) within the reanalysis system. This model is developed based on large-scale and uniform underlying surfaces. Its coarser spatial resolution (0.5°) along with uniform parameterization schemes may not accurately describe the land surface processes in complex terrain areas, thereby further increasing ET deviations. Overall, reanalysis products show limited accuracy in high-altitude mountainous regions, which is consistent with those reported by Liu et al. [44] and Qian et al. [45].

Remote sensing-based datasets, such as GLEAM-ET, MOD16-ET and GLASS-ET, generally perform well, while variations remain in the identification of change regions and the significance levels. GLEAM-ET estimates potential ET (PET) based on the Priestley-Taylor equation and reduce PET to ET using soil moisture stress factors in a simple water balance framework. With relatively fewer parameters, it better reflects the relationship between energy and water stress [45]. This may be an important reason for its better correlation with TWB-ET. Many studies have also demonstrated that GLEAM-ET performs relatively well in the TP [21–24]. MOD16-ET exhibits a more accurate interannual changing trend but shows relatively weaker correlation in annual variation, consistent with findings in Meng et al. [21] and Li et al. [22]. MOD16-ET is derived from the Penman-Monteith (PM) equation, which effectively reflect long-term ET changes under climate-vegetation coupling. However, canopy-related parameters (e.g., LAI, FPAR) may have greater uncertainties at short-time scales [46]. PML-ET display a significant upward trend across the entire basin, markedly differing from other datasets. This dataset introduces the dynamic coupling of photosynthesis and stomatal conductance into the PM equation. Although PML enables more accurate transpiration estimation, its parameterization scheme may be inadequately adapted to the low-pressure condition in the YZB, and strong radiation may lead to systematic overestimation of stomatal conductance, thereby causing inaccurate dynamic change estimates [47]. GLASS-ET integrates five process-based algorithms for ET estimation, integrating the strengths of different approaches in energy and vegetation processes. This may be the reason for its superior overall performance.

In contrast, the overall performance of Ma-ET, Chen-ET and Han-ET was slightly poorer. Chen-ET estimates ET by deriving the Bowen ratio from the balance of surface temperature and humidity. These variables are susceptible to local microclimate influences, making it difficult to capture their long-term trends, thereby leading to significant deviations. Ma-ET is based on the nonlinear complementary relationship, which assumes dynamic feedbacks between ET and PET under drought-wet transition conditions. This approach performed well in humid regions but was prone to failure in high-cold region [48]. Han-ET is based on the surface energy balance system (SEBS) model, which is sensitive to surface temperature, surface roughness and radiation distribution. However, MODIS-derived surface temperature and roughness parameters exhibit systematic biases in high-altitudes regions, leading to its overall poor performance in the YZB.

Jung-ET uses FLUXNET observations to establish a nonlinear mapping relationship, achieving higher accuracy in data-rich areas. However, due to the scarcity of flux towers in high-altitude regions, the representativeness of training samples is limited, thereby restricting its performance. Overall, it performs relatively well in the eastern YZB, but the deviations become significantly larger in the central and western regions, indicating that purely data-driven models still lack reliability in high-altitude regions.

In addition, the impacts of resolution differences on ET estimation across each dataset are also reflected in the evaluation. Generally, high-resolution datasets perform better than low-resolution datasets, as they can capture finer spatial variation, whereas low-resolution datasets tend to overlook detailed spatial features during the estimation process.

For example, both MERRA2-ET (0.5°) and Jung-ET (0.5°) exhibited poorer performance compared to GLASS-ET and MOD16-ET.

In summary, potential weaknesses exist in all ET datasets. The performance of ET datasets is highly dependent on the regional characteristics, algorithms and forcing data accuracy. In the future, it is recommended to develop fusion products that integrate the advantages of both reanalysis and remote sensing data, with particular emphasis on improving the parameterization of physical mechanisms for high-altitude cold river basins.

4.2. TWB-ET Estimation

(1) The sources and impacts of TWB-ET uncertainty

Uncertainties exist in TWB-ET estimates in the YZB, which is a key limitation in our research. First, TWB-ET estimates are highly sensitive to precipitation, runoff and TWS. In high-altitude mountainous areas with sparse observational stations, input data uncertainty increases. Among all sources of uncertainty, precipitation uncertainty plays a dominant role. According to Miao et al. [41], observed precipitation data may be significantly underestimated in the high-altitude environments, potentially leading to ET underestimation. Secondly, the impact of glacier mass balance has not been explicitly accounted for in TWB-ET estimates. Given the large glaciers areas in the YZB, glaciers melt contributions to runoff result in systematic deviations in TWB calculations, which may lead to an underestimation of TWB-ET. According to relevant studies [49,50], glacier melt contributes approximately 5% to annual runoff on average in the YZB, mainly occurs from June to September. Therefore, TWB-ET may be underestimated during these months. Considering this 5% contribution suggests that the overall underestimation of TWB-ET in areas with glaciers cover is approximately 4.1%. Thirdly, TWB-ET estimates rely on the assumption of a closed watershed, neglecting potential groundwater lateral flow. This may increase the deviations of TWB-ET. Due to limited observation data in the YZB, quantifying this uncertainty remains challenging. Li et al. [22] similarly identified glacier effects, input data uncertainty, and unmonitored groundwater outflow as key contributors to uncertainty in TWB-ET estimates over the TP.

(2) The causes of negative TWB-ET estimates

The direct reason for the negative TWB-ET in winter may be the significant underestimation of precipitation in this basin, especially the underestimation of snowfall. Research by Miao et al. [41] indicates that in the TP, observed precipitation has been significantly underestimated due to measurement errors caused by strong winds and representativeness errors resulting from sparse stations. This leads to a lower precipitation input in the TWB equation, especially in sub-basins with abundant snowfall. In contrast, GRACE-based TWS changes can effectively capture total water storage changes including snow accumulation. As a result, the TWB calculation derived negative ET in winter. The root cause lies in the fact that the solid (e.g., snow) and liquid (e.g., soil moisture) water storage changes was not effectively distinguished in the current TWB method. The uncertainty analysis of TWB-ET (Table 3) also indicates that precipitation exhibits greater uncertainties (3–7 mm/month) and is a dominant source of TWB-ET uncertainty, compared to the uncertainty of GRACE-based TWS (0.12 mm/month). Moreover, in areas with more glacier and snow cover (e.g., B5, B7), the uncertainty is even greater, which is highly consistent with the seasons and regions where negative values occur. Li et al. [22] also observed negative values in the TWB-ET estimation in the YZB, and attributed them to the same reasons.

(3) The reliability of TWB-ET estimates

TWB-ET showed negative values in some sub-basins during the winter months, indicating the unreliability of these values. Therefore, we mainly conducted evaluation on

the annual scale. On the annual scale, the TWS changes approached 0, and the annual ET estimates was less affected by this. Additionally, this study mainly compared and evaluated the consistency of each ET datasets with TWB-ET, rather than comparing absolute quantities, thereby effectively reducing the impact of TWB-ET uncertainties on the evaluation conclusions. Meanwhile, the uncertainty analysis of TWB-ET also indicated that although there were individual negative values, the relative uncertainty of TWB-ET (σ_{ET}/TWB-ET) is mostly below 20%, reaching only 21.9% in B7 (Table 3), well within the acceptable range (<30%) as defined by Li et al. [22]. Therefore, compared to ET datasets that involve more uncertainty sources, TWB-ET estimates remain certain reliability for use as a baseline. For the YZB with scarce data, the correction of the calculation may bring new and greater uncertainties. Therefore, future research should combine atmospheric water balances and ground observations, and use multi-method integration to build a more reliable regional ET baseline.

4.3. Limitations

Several limitations should also be acknowledged. First, uncertainties in TWB-ET estimates limit a more detailed evaluation of ET with respect to spatio-temporal variations. Future studies should combine ground flux observations (such as eddy covariance data) to improve the evaluation. Secondly, the lack of runoff observation data in sub-basins B8 and B9 make it unable to calculate TWB-ET in these regions, thereby restricting quantitative evaluation of ET products in the lower reaches of the basin. The comprehensive score is restricted to the upper-middle reaches of the YZB. Additionally, the comprehensive score is calculated by assigning equal weights to the main aspects and lower weights to the remaining aspects. To identify datasets that perform better in specific aspects, the corresponding aspect scores can be directly examined. If different emphases are required, the weights can be adjusted accordingly, and the overall scores can be recomputed. Meanwhile, the comprehensive scores are based on relative performance across 10 datasets, implying that even top-performing datasets still exhibit notable limitations. For example, the GLASS-ET and GLEAM-ET, ranking among the top, still exhibit significant systematic deviation in ET estimation during winter in high-altitude and glacierized areas. GLEAM-ET also performs limited capability in capturing interannual variability. Therefore, practical applications should carefully consider specific regional characteristics, seasonal dynamics, and research objectives. Moreover, due to space limitations, certain datasets were not included in the analysis, such as data fusion products and land surface model outputs. Future studies should summarize and evaluate the applicability of a broader range ET datasets.

5. Conclusions

This paper estimated the actual evapotranspiration (ET) in the YZB using the terrestrial water balance approach (TWB-ET), incorporating GRACE-based terrestrial water storage change data, precipitation, and observed runoff data. The TWB-ET was then used to compare and evaluate the performance of ten different ET datasets in the YZB. The main conclusions are as follows: (1) Across the entire basin, all ten ET datasets exhibit a good correlation with TWB-ET in annual variation, with *r* values ranging from 0.78 to 0.90. Six of these datasets achieve an *r* value exceeding 0.85, although biases exist. The correlations between multi-source ET datasets and TWB-ET vary across the sub-basins. Overall, GLEAM-ET demonstrates the highest correlation with TWB-ET (r = 0.88) and the smallest bias (*RMSE* = 14.24 mm/month, *Rbias* = 18.55%) in annual variation. (2) The spatial distribution of the ten ET datasets is generally similar, showing a decreasing trend from southeast to northwest. However, significant differences are observed in the primary variation range and specific changes within sub-basins. Temporally, there are evident

discrepancies in the specific regions and significance level of changing trends among multi-source ET datasets. (3) By comparing the ten ET datasets with TWB-ET in terms of annual variation, spatio-temporal variation in the YZB, it is found that GLASS-ET and GLEAM-ET perform relatively well, whereas Han-ET and Chen-ET exhibits larger differences in these aspects.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/rs18010162/s1, Table S1: Evaluation metrics in each evaluated item across ten ET datasets under TWB-ET estimates uncertainties; Table S2: Comprehensive evaluation of different ET datasets under TWB-ET estimates uncertainties. References [4,11,22,23,27,40,51] are cited in the supplementary materials.

Author Contributions: Methodology, Y.J.; Formal analysis, Z.X. (Zihao Xia); Writing—review & editing, L.X.; Supervision, Z.X. (Zongxue Xu). All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Natural Science Foundation of China (Grant number: 52269012) and the Natural Science Foundation of Jiangxi province (Grant number: 20232BAB214084, 20232BAB214087).

Data Availability Statement: The original contributions presented in this study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Acknowledgments: The authors especially acknowledge the authors of datasets cited in the text, as well as the platforms and organizations for providing data support.

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

References

1. Guo, N.; Chen, H.; Han, Q.; Wang, T. Evaluating data-driven and hybrid modeling of terrestrial actual evapotranspiration based on an automatic machine learning approach. *J. Hydrol.* **2024**, *628*, 130594. [CrossRef]
2. Wang, W.; Li, J.; Yu, Z.; Ding, Y.; Xing, W.; Lu, W. Satellite retrieval of actual evapotranspiration in the Tibetan Plateau: Components partitioning, multidecadal trends and dominated factors identifying. *J. Hydrol.* **2018**, *559*, 471–485. [CrossRef]
3. Long, D.; Singh, V. Two-source trapezoid model for evapotranspiration (TTME) from satellite imagery. *Remote Sens. Environ.* **2012**, *121*, 370–388. [CrossRef]
4. Zhong, Y.; Zhong, M.; Mao, Y.; Ji, B. Evaluation of evapotranspiration for exorheic catchments of China during the GRACE Era: From a water balance perspective. *Remote Sens.* **2020**, *12*, 511. [CrossRef]
5. Li, S.; Wang, G.; Zhu, C.; Lu, J.; Ullah, W.; Hagan, D.; Kattel, G.; Peng, J. Attributing of global evapotranspiration trends based on the Budyko framework. *Hydrol. Earth Syst. Sci.* **2022**, *26*, 3691–3707. [CrossRef]
6. Zhou, Y.; Marshall, L.; Li, D.; Sharma, A. Revisiting evapotranspiration inputs in eco-hydrological modeling for climate change assessment. *J. Hydrol.* **2024**, *642*, 131888. [CrossRef]
7. Liu, D.; Wang, Z.; Wang, L.; Chen, J.; Li, C.; Shi, Y. Improved remote sensing reference evapotranspiration estimation using simple satellite data and machine learning. *Sci. Total Environ.* **2024**, *947*, 174480. [CrossRef] [PubMed]
8. Running, S.; Mu, Q.; Zhao, M.; Moreno, A. MOD16A2GF MODIS/Terra Net Evapotranspiration Gap-Filled 8-Day L4 Global 500 m SIN Grid V006 [Dataset]. NASA EOSDIS Land Processes DAAC; 2019. Available online: https://www.earthdata.nasa.gov/data/catalog/lpcloud-mod16a2gf-006 (accessed on 20 December 2023).
9. Martens, B.; Miralles, D.; Lievens, H.; Van Der Schalie, R.; De Jeu, R.A.; 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]
10. Zhang, Y.; He, S. PML-V2(China): Evapotranspiration and Gross Primary Production Dataset (2000.02.26–2020.12.31). National Tibetan Plateau/Third Pole Environment Data Center. 2022. Available online: https://data.tpdc.ac.cn/en/data/40f57c67-33a6-402d-bd37-6ede91919f23/ (accessed on 20 December 2023).
11. Rodell, M.; Famiglietti, J.S.; Chen, J.; Seneviratne, S.; Viterbo, P.; Holl, S.; Wilson, C.R. Basin scale estimate of evapotranspiration using GRACE and other observations. *Geophys. Res. Lett.* **2004**, *31*, L20504. [CrossRef]

12. Muñoz Sabater, J. ERA5-Land Monthly Averaged Data from 1950 to Present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). 2019. Available online: https://cds.climate.copernicus.eu/datasets/reanalysis-era5-land-monthly-means?tab=overview (accessed on 29 December 2023).
13. Cai, Y.; Xu, Q.; Bai, F.; Cao, X.; Wei, Z.; Lu, X.; Wei, N.; Yuan, H.; Zhang, S.; Liu, S.; et al. Reconciling global terrestrial evapotranspiration estimates from multi-product intercomparison and evaluation. *Water Resour. Res. Water Resour. Res.* **2024**, *60*, e2024WR037608. [CrossRef]
14. Li, S.; Wang, G.; Sun, S.; Chen, H.; Bai, P.; Zhou, S.; Huang, Y.; Wang, J.; Deng, P. Assessment of Multi-Source Evapotranspiration Products over China Using Eddy Covariance Observations. *Remote Sens.* **2018**, *10*, 1962. [CrossRef]
15. Zuo, L.; Zou, L.; Xia, J.; Zhang, L.; Cao, H.; She, D. Multi-scale analysis of six evapotranspiration products across China: Accuracy, uncertainty and spatiotemporal pattern. *J. Hydrol.* **2025**, *650*, 132516. [CrossRef]
16. Li, X.; Zhang, W.; Vermeulen, A.; Dong, J.; Duan, Z. Triple collocation-based merging of multi-source gridded evapotranspiration data in the Nordic Region. *Agric. For. Meteorol.* **2023**, *335*, 109451. [CrossRef]
17. de Ferreira, L.; da Paz, A. Enhanced calibration of a distributed hydrological model in the Brazilian Semi-Arid: Integrating spatiotemporal evapotranspiration and streamflow data. *Environ. Earth Sci.* **2024**, *83*, 345. [CrossRef]
18. Nkiaka, E.; Bryant, R.G.; Ntajal, J.; Biao, E.I. Evaluating the accuracy of gridded water resources reanalysis and evapotranspiration products for assessing water security in poorly gauged basins. *Hydrol. Earth Syst. Sci.* **2022**, *26*, 5899–5916. [CrossRef]
19. Yu, X.; Qian, L.; Wang, W.; Hu, X.; Dong, J.; Pi, Y.; Fan, K. Comprehensive evaluation of terrestrial evapotranspiration from different models under extreme condition over conterminous United States. *Agric. Water Manag.* **2023**, *289*, 108555. [CrossRef]
20. 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]
21. Meng, X.; Deng, M.; Shu, L.; Chen, H.; Wang, S.; Li, Z.; Zhao, L.; Shang, L. An evaluation of evapotranspiration products over the Tibetan plateau. *J. Hydrometeorol.* **2024**, *25*, 1665–1677. [CrossRef]
22. Li, X.; Long, D.; Han, Z.; Scanlon, B.; Sun, Z.; Han, P.; Hou, A. Evapotranspiration estimation for Tibetan Plateau headwaters using conjoint terrestrial and atmospheric water balances and multisource remote sensing. *Water Resour. Res.* **2019**, *55*, 8608–8630. [CrossRef]
23. Cheng, M.; Zhong, L.; Ma, Y.; Ma, H.; Chang, Y.; Li, P.; Cheng, M.; Wang, X.; Ge, N. A study on the assessment and integration of multi-source evapotranspiration products over the Tibetan Plateau. *Adv. Atmos. Sci.* **2024**, *41*, 435–448. [CrossRef]
24. Liu, W. Evaluating remotely sensed monthly evapotranspiration against water balance estimates at basin scale in the Tibetan Plateau. *Hydrol. Res.* **2018**, *49*, 1977–1990. [CrossRef]
25. Li, X.; Wang, L.; Chen, D.; Yang, K.; Wang, A. Seasonal evapotranspiration changes (1983-2006) of four large basins on the Tibetan Plateau. *J. Geophys. Res. Atmos.* **2014**, *19*, 13079–13095. [CrossRef]
26. Jiang, Y.; Xu, Z.; Xiong, L. Runoff variation and response to precipitation on multi-spatial and temporal scales in the southern Tibetan Plateau. *J. Hydrol. Reg. Stud.* **2022**, *42*, 101157. [CrossRef]
27. He, J.; Yang, K.; Li, X.; Tang, W.; Shao, C.; Jiang, Y.; Ding, B. China Meteorological Forcing Dataset v2.0 (1951–2024). National Tibetan Plateau/Third Pole Environment Data Center. 2024. Available online: https://data.tpdc.ac.cn/en/data/e60dfd96-5fd8-493f-beae-e8e5d24dece4 (accessed on 29 December 2024).
28. Zhong, Y.; Feng, W.; Zhong, M.; Ming, Z. Dataset of Reconstructed Terrestrial Water Storage in Mainland China Based on Precipitation (2002–2019). National Tibetan Plateau/Third Pole Environment Data Center. 2020. Available online: https://data.tpdc.ac.cn/en/data/71cf70ec-0858-499d-b7f2-63319e1087fc/ (accessed on 29 December 2023).
29. Gelaro, R.; McCarty, W.; Suarez, M.J.; Todling, R.; Molod, A.; Takacs, L.; Zhao, B. The Modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *J. Clim.* **2017**, *30*, 5419–5454. [CrossRef] [PubMed]
30. Global Modeling and Assimilation Office (GMAO). MERRA-2 tavgU_2d_lnd_Nx: 2d, Diurnal, Time-Averaged, Single-Level, Assimilation, Land Surface Diagnostics V5.12.4. Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC); 2015. Available online: https://disc.gsfc.nasa.gov/datasets/M2TUNXLND_5.12.4/summary (accessed on 29 December 2023).
31. Yao, Y.; Liang, S.; Li, X.; Hong, Y.; Fisher, J.; Zhang, N.; Chen, J.; Cheng, J.; Zhao, S.; Zhang, X.; et al. Bayesian multimodel estimation of global terrestrial latent heat flux from eddy covariance, meteorological, and satellite observations. *J. Geophys. Res. Atmos.* **2014**, *119*, 4521–4545. [CrossRef]
32. Zhang, Y.; Kong, D.; Gan, R.; Chiew, F.H.S.; McVicar, T.R.; Zhang, Q.; Yang, Y. Coupled estimation of 500 m and 8-day resolution global evapotranspiration and gross primary production in 2002–2017. *Remote Sens. Environ.* **2019**, *222*, 165–182. [CrossRef]
33. Ma, N.; Jozsef, S.; Zhang, Y.; Liu, W. Terrestrial Evapotranspiration Dataset Across China (1982–2017). National Tibetan Plateau/Third Pole Environment Data Center. 2019. Available online: https://data.tpdc.ac.cn/en/data/b6d9f525-5b76-48b0-82db-bb2963465cac (accessed on 29 December 2023).
34. Ma, N.; Szilagyi, J.; Zhang, Y.; Liu, W. Complementary-relationship-based modeling of terrestrial evapotranspiration across China during 1982-2012: Validations and spatiotemporal analyses. *J. Geophys. Res. Atmos.* **2019**, *124*, 4326–4351. [CrossRef]

35. Han, C.; Ma, Y.; Wang, B.; Zhong, L.; Ma, W.; Chen, X.; Su, Z. Monthly Mean Evapotranspiration Data Set of the Tibet Plateau (2001–2018). National Tibetan Plateau/Third Pole Environment Data Center. 2020. Available online: https://data.tpdc.ac.cn/en/data/5a0d2e28-ebc6-4ea4-8ce4-a7f2897c8ee6/ (accessed on 29 December 2023).
36. Han, C.; Ma, Y.; Wang, B.; Zhong, L.; Ma, W.; Chen, X.; Su, Z. Long term variations of actual evapotranspiration over the Tibetan Plateau. *Earth Syst. Sci. Data* **2021**, *13*, 3513–3524. [CrossRef]
37. Chen, S.; McColl, K.A.; Berg, A.; Huang, Y. Surface flux equilibrium estimates of evapotranspiration at large spatial scales. *J. Hydrometeorol.* **2021**, *22*, 769–779. [CrossRef]
38. Jung, M.; Koirala, S.; Weber, U.; Ichii, K.; Gans, F.; Camps-Valls, G.; Papale, D.; Schwalm, C.; Tramontana, G.; Reichstein, M. The FLUXCOM ensemble of global land-atmosphere energy fluxes. *Sci. Data* **2019**, *6*, 74. [CrossRef] [PubMed]
39. Cheng, Y.; Zhang, X.; Wang, K.; Zhang, Y.; Guo, Y.; Shen, Y. Multidimensional evaluation of satellite-based and reanalysis-based precipitation datasets in the Tibetan Plateau. *J. Hydrol.* **2025**, *660*, 133364. [CrossRef]
40. Pan, Y.; Zhang, C.; Gong, H.; Yeh, P.; Shen, Y.; Guo, Y.; Huang, Z.; Li, X. Detection of human-induced evapotranspiration using GRACE satellite observations in the Haihe River basin of China. *Geophys. Res. Lett.* **2017**, *44*, 190–199. [CrossRef]
41. Miao, C.; Immerzeel, W.W.; Xu, B.; Yang, K.; Duan, Q.; Li, X. Understanding the Asian water tower requires a redesigned precipitation observation strategy. *Proc. Natl. Acad. Sci. USA* **2024**, *121*, e1891410175. [CrossRef]
42. Fan, X.; Wang, L.; Liu, H.; Chen, L.; Wang, Y.; Qi, J.; Chai, C.; Liu, R.; Li, X.; Zhou, J.; et al. Tibetan Plateau runoff and evapotranspiration dataset by an observation-constrained cryosphere-hydrology model. *Sci. Data* **2024**, *11*, 773. [CrossRef]
43. Yuan, L.; Chen, X.; Ma, Y.; Han, C.; Wang, B.; Ma, W. Long-term monthly 0.05° terrestrial evapotranspiration dataset (1982-2018) for the Tibetan Plateau. *Earth Syst. Sci. Data* **2024**, *16*, 775–801. [CrossRef]
44. Liu, Y.; Wang, W.; Zhao, T.; Huo, Z. Performance evaluation and spatiotemporal dynamics of nine reanalysis and remote sensing evapotranspiration products in China. *Remote Sens.* **2025**, *17*, 1881. [CrossRef]
45. Qian, L.; Zhang, Z.; Wu, L.; Fan, S.; Yu, X.; Liu, X.; Ba, Y.; Ma, H.; Wang, Y. High uncertainty of evapotranspiration products under extreme climatic conditions. *J. Hydrol.* **2023**, *626*, 130332. [CrossRef]
46. Mu, Q.; Zhao, M.; Running, S.W. Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Remote Sens. Environ.* **2011**, *115*, 1781–1800. [CrossRef]
47. Guo, L.; Wu, Y.; Zheng, H.; Zhang, B.; Fan, L.; Chi, H.; Yan, B.; Wang, X. Consistency and uncertainty of gridded terrestrial evapotranspiration estimations over China. *J. Hydrol.* **2022**, *612*, 128245. [CrossRef]
48. Yu, W.; Xie, Y.; Li, Y.; Kumar, A.; Shao, W.; Zhao, Y. Complementary relationship-based validation and analysis of evapotranspiration in the permafrost region of the Qinghai-Tibetan Plateau. *Atmosphere* **2025**, *16*, 932. [CrossRef]
49. Zhao, D.; He, K.; Xiong, D.; Lu, X.; Qin, X.; Wang, X.; Zhang, W. Dam-induced alternations of flow and sediment regimes in the Tibetan Plateau: An example of the Yarlung Tsangpo river. *Water Resour. Res.* **2025**, *61*, e2024WR039016. [CrossRef]
50. Nan, Y.; Tian, F.; Mcdonnell, J.; Ni, G.; Tian, L.; Li, Z.; Yan, D.; Xia, X.; Wang, T.; Han, S.; et al. Glacier meltwater has limited contributions to the total runoff in the major rivers draining the Tibetan Plateau. *Npj Clim. Atmos. Sci.* **2025**, *8*, 155. [CrossRef]
51. Scanlon, B.R.; Zhang, Z.; Save, H.; Sun, A.Y.; Müller Schmied, H.; Van Beek, L.P.; Wiese, D.N.; Wada, Y.; Long, D.; Reedy, R.C.; et al. Global models underestimate large decadal declining and rising water storage trends relative to GRACE satellite data. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E1080–E1089. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.