

## Article

# Intercomparison of Runoff and River Discharge Reanalysis Datasets at the Upper Jinsha River, an Alpine River on the Eastern Edge of the Tibetan Plateau

Shuanglong Chen <sup>1,†</sup>, Heng Yang <sup>2,†</sup>, and Hui Zheng <sup>3,\*</sup>

<sup>1</sup> Baihetan Hydropower Plant, China Three Gorges Corporation, Sichuan 615421, China; [chen\\_shuanglong@ctg.com.cn](mailto:chen_shuanglong@ctg.com.cn)

<sup>2</sup> Science and Technology Research Institute, China Three Gorges Corporation, Beijing 100038, China; [yang\\_heng2@ctg.com.cn](mailto:yang_heng2@ctg.com.cn)

<sup>3</sup> Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China; [hzheng\\_iap@outlook.com](mailto:hzheng_iap@outlook.com)

\* Correspondence: [hzheng\\_iap@outlook.com](mailto:hzheng_iap@outlook.com)

† These authors contributed equally to this work and should be considered co-first authors.

**Abstract:** This study assesses the effectiveness and limitations of publicly accessible runoff and river discharge reanalysis datasets through an intercomparison in the Upper Jinsha River, an alpine region with substantial hydropower potential on the eastern edge of the Tibetan Plateau. The examined datasets are the European Centre for Medium-Range Weather Forecast Reanalysis version 5 (ERA5-Land), the Global Flood Awareness System (GloFAS), the Global Reach-Level Flood Reanalysis (GRFR), and the China Natural Runoff Dataset (CNRD). These datasets are created using various meteorological forcing, runoff generation models, river routing models, and calibration methods. To determine the causes of discrepancies, additional simulations were carried out. One simulation, driven by meteorological forcing similar to that of ERA5-Land and GloFAS but utilizing the uncalibrated NoahMP land surface model at a higher spatial resolution, was included to evaluate the effects of meteorological inputs, spatial resolution, and calibration on runoff estimation. Runoff from all datasets was rerouted on a high-resolution river network derived from the 3-arcsecond Multi-Error-Removed Improved-Terrain Hydrography (MERIT-Hydro) dataset, allowing for a comparison between vector- and grid-based river routing models for discharge estimates. The intercomparison is grounded in observations from three gauging stations—Zhimenda, Gangtuo, and Benzilan—at monthly, daily, and hourly scales. The results suggest that model calibration has a more significant influence on runoff and discharge estimates than meteorological data. Calibrated datasets, such as GloFAS and GRFR, perform better than others, despite variations in the forcing data. The runoff characteristics-based calibration method used in GRFR exhibits superior performance at Zhimenda and Benzilan. However, at Gangtuo, GRFR's performance is unsatisfactory, highlighting the limitation of the machine-learning-based method in regions with rugged terrain and limited observations. Vector-based river routing models demonstrate advantages over grid-based models. GloFAS, which uses a grid-based routing model, encounters difficulties in simultaneously producing accurate runoff and discharge estimates. The intercomparison shows that GRFR's river routing is sub-optimally configured. However, when GRFR's runoff rerouted, the performance of discharge improves substantially, attaining a Kling–Gupta Efficiency of approximately 0.9. These findings offer valuable insights for the further development of reanalysis datasets in this region.

Received:

Revised:

Accepted:

Published:

**Citation:** Chen, S.; Yang, H.; Zheng, H. Intercomparison of Runoff and River Discharge Reanalysis Datasets at the Upper Jinsha River. *Water* **2025**, *1*, 0. <https://doi.org/>

**Copyright:** © 2025 by the authors. Submitted to *Water* for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** runoff; river discharge; model intercomparison; Upper Jinsha River

## 1. Introduction

The rapid growth of global- and continental-scale runoff and discharge reanalysis datasets marks a significant advancement in hydrological research. Datasets such as the European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalysis version 5 (ERA5) [1], ERA5-Land [2], the Global Flood Awareness System (GloFAS) [3,4], and the Global Reach-Level Flood Reanalysis (GRFR) [5,6] provide continuous, spatially comprehensive runoff and river discharge estimates. Additionally, regional datasets like the National Water Model (NWM) of the United States [7], the European Flood Early Warning Systems [8], and the China Natural Runoff Dataset (CNRD) [9,10] provide runoff and discharge estimates tailored to specific regions. These datasets have greatly improved our understanding of river systems [11,12] and their interactions with other components of the Earth system [13]. Furthermore, they have been instrumental in supporting decision-making processes, particularly in the area of water-resources utilization and flood protection [7,14].

Reanalysis datasets are generated through a complex chain of methodologies that include meteorological forcing data, land surface models or hydrological models, river routing models, river network delineations, and calibrations. The diversity among these components can significantly influence the estimation of runoff and river discharge. Studies have demonstrated that meteorological forcing data can differ remarkably [15,16]. The use of different meteorological forcing data [17,18], in combination with varying model structure [19], spatial resolutions [20–22], parameterizations [23,24], and parameters [25] of land surface models or hydrological models, can lead to disparate estimates of runoff [26,27]. Furthermore, these differences in runoff estimation can be further amplified by the methods employed in river routing [28–30].

The diversity in estimation methodologies necessitates careful intercomparison before using reanalysis datasets. In the past decades, a variety of intercomparison studies have been conducted with different objectives. Land surface model comparisons such as the Project for Intercomparison of Land-surface Parameterization Schemes [31] and the Land Surface, Snow and Soil Moisture Model Intercomparison Project [32] conducted with the same meteorological forcing data greatly help to reveal the weakness of the participant models and the interactions of land surface processes. The Distributed Model Intercomparison Project [33,34] compared distributed models with lumped models. This project revealed the superiority of distributed routing models in reproducing flood peaks and reported the importance of calibration on distributed rainfall-runoff models. Beck et al. [27] compared ten hydrological models at the global scale. Their results further confirm that proper parameter calibration is vital for accurate runoff estimation. The Continental Hydrological Intercomparison Project [35] compared two high-resolution, physics-based models at the continental scale. The intercomparison promotes high-resolution modeling to enhance topographic processing and to improve topographic representation over complex topography. There are also intercomparisons with regional focuses. The Great Lakes Runoff Intercomparison Project [29] and the Nelson Model Intercomparison Project under the Integrated Modelling Program of Canada initiative [30] are two examples. These intercomparison studies aim to select the best-performing estimation methodologies for the region and guide further improvements of operations.

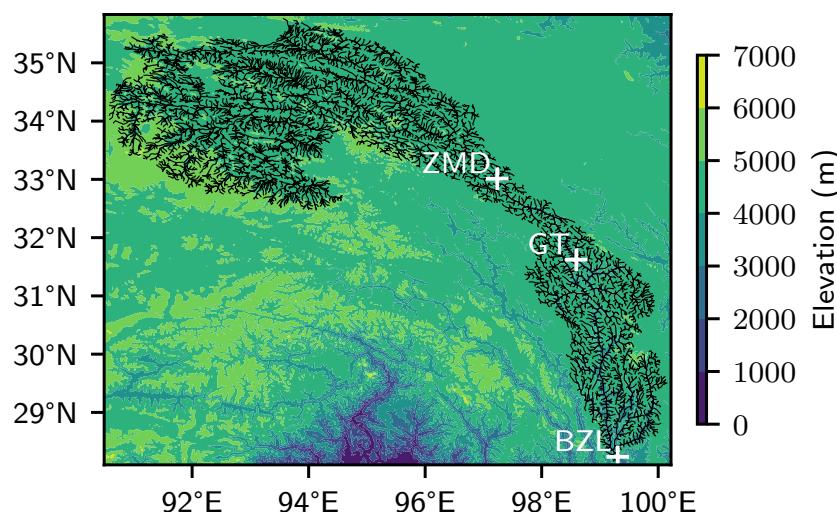
The Upper Jinsha River, the source of the Changjiang River, is notable for its steep gradient and high hydroelectric potential. The river is located on the eastern edge of the Tibetan Plateau. Thirteen hydroelectric facilities have been built or planned along its main stem, collectively possessing an installed capacity of approximately 14.6 gigawatts. The river originates from the Tibetan Plateau. The source area of the river consists of a major part of the Sanjiangyuan National Nature Reserve, the highest and largest nature reserve in China yet. The operations of hydroelectric facilities and natural reserves require

accurate estimation of water resources. However, the scarcity of in-situ hydrometeorological observations in the region is a notable challenge. In this context, reanalysis datasets of runoff and discharge emerge as an indispensable data source for the area. Simultaneously, the Upper Jinsha River's intricate topography, high-alpine environmental conditions, and the pressing need for accurate runoff and river discharge estimations render it an exemplary testing ground for these reanalysis datasets.

This study intercompares runoff and river discharge reanalysis data obtained from ERA5-Land, GloFAS, GRFR, and CNRD for the Upper Jinsha River as a pilot analysis. The primary objective is to assess the effectiveness and limitations of the methodologies employed in these reanalyses, providing insights to guide future improvements in runoff and discharge estimation.

## 2. Materials and Methods

Figure 1 illustrates the terrain and river network of the upper Jinsha River. The study area encompasses the region upstream of the Benzilan gauging station. The elevation ranges from approximately 6600 m at its highest point to 2000 m at its lowest. The river network was extracted from the Multi-Error-Removed Improved-Terrain Hydrography (MERIT-Hydro) dataset [36,37]. MERIT-Hydro is a globally recognized dataset providing hydrologically conditioned elevation, flow direction, and flow accumulation area data, with a spatial resolution of 3 arc-seconds. The river network was extracted using a flow accumulation area threshold of 25 km<sup>2</sup>, resulting in 5,787 river reaches.



**Figure 1.** Study domain and river network of the source region of the upper Jinsha River. White crosses mark the gauging stations: ZMD for Zhimenda, GT for Gangtuo, and BZL for Benzilan.

### 2.1. Reanalysis Datasets

Table 1 presents an overview of the datasets employed in this study. ERA5-Land [2] is an enhanced global dataset, an improvement over ERA5 [1]. This dataset features a significant improvement in spatial resolution, increasing from ERA5's 31 km to 0.1° globally. A variety of land surface variables, including runoff, are simulated using the Carbon Hydrology-Tiled ECMWF Scheme for Surface Exchanges over Land (CHTESSEL) model [38,39], operating at an hourly temporal resolution.

The GloFAS historical reanalysis version 4.0 [3,4] is another dataset produced by the European Centre for Medium-Range Weather Forecasts (ECMWF), with a focus on hydrological variables. GloFAS shares the same spatial resolution of 0.1° as ERA5-Land. It is driven by ERA5 meteorological data, which is highly similar to ERA5-Land. The primary distinctions between GloFAS and ERA5-Land lie in the hydrological model employed, the

temporal resolution, and the output variables. GloFAS utilizes the LISFLOOD model [40] at a daily time step. Unlike ERA5-Land, the parameters of LISFLOOD are calibrated against daily river discharge observations using an evolutionary algorithm [41]. LISFLOOD simulates both runoff and river discharge, whereas ERA5-Land provides only runoff estimates.

The GRFR dataset [5] is a comprehensive global dataset providing runoff and river discharge data at a 3-hourly temporal resolution. The runoff is simulated using the Variable Infiltration Capacity (VIC) model [42] at a spatial resolution of  $0.05^\circ$ , which is driven by precipitation data from the Multi-Source Weighted-Ensemble Precipitation (MSWEP) dataset [43] and other meteorological variables from ERA5. The VIC model's parameters are calibrated [44] at each  $0.05^\circ$  grid cell against a machine-learning-derived runoff characteristics dataset produced by Beck et al. [45]. The bias in the VIC-simulated runoff is further corrected using the same runoff characteristics dataset, following the methodology of Lin et al. [6]. The calibrated and bias-corrected runoff data are then utilized to drive the Routing Application for Parallel computation of Discharge (RAPID) river routing model [46] for simulating river discharge. The parameters of RAPID are estimated using a statistical relationship that correlates discharge with river width, water depth, channel slope, and celerity [6]. The river network employed in the routing process is also derived from the MERIT Hydro dataset [47].

The China Natural Runoff Dataset (CNRD) [9,10] comprises runoff data at a  $0.25^\circ$  spatial resolution. This dataset is simulated using the VIC model, driven by the China Meteorological Forcing Dataset [48]. The parameters of the VIC model were calibrated against river discharge observations from approximately 200 catchments across China and then extrapolated to ungauged river basins using the multiscale regionalization method [49].

As presented in Table 1, a new dataset was generated in this study for comparative purposes. This dataset was driven by ERA5-Land meteorological data and simulated via a distinct land surface model and river routing model. Given that the new dataset shares a similar meteorological forcing dataset with GloFAS and ERA5-Land, this comparison can effectively isolate the impacts of model selection, spatial resolution, and calibration on runoff and discharge estimations. The model employed in this study emulates the National Water Model (NWM). The Noah with Multi - Parameterization (NoahMP) land surface model [50] was utilized, operating at a spatial resolution of  $0.01^\circ$  from 2008 to 2016. The static parameters for NoahMP are derived from the Weather Research and Forecasting geographic static data ([https://www2.mmm.ucar.edu/wrf/users/download/get\\_sources\\_wps\\_geog.html](https://www2.mmm.ucar.edu/wrf/users/download/get_sources_wps_geog.html)). The initial condition for 2008 is sourced from ERA5-Land. The first year is allocated for model spin-up, with the simulated runoff from 2009 to 2016 used for analysis. NoahMP offers a variety of parameterization schemes for land surface processes, and we adopted the same schemes as those used in the NWM of the United States [7]. The runoff is routed using the variable-parameter Muskingum–Cunge (MC) method on the river network depicted in Figure 1 to estimate river discharge. The routing algorithm mirrors that of the NWM, assuming a trapezoidal channel cross-section. Channel cross-section geometry and roughness parameters are sourced from the default parameter lookup table of the NWM [51] and the Strahler stream order for each river reach within the network. In theory, the MC routing method would outperform the kinematic wave routing method used in GloFAS by also additionally considering diffusive wave dynamics, while the Muskingum routing method used in GRFR would be the least performing among the three. It is important to note that this comparison is theoretical, and actual performance can vary significantly with model implementation and parameter settings

## 2.2. River discharge observations

We utilized river discharge observations from three gauging stations along the mainstream of the upper Jinsha River: Zhimenda (ZMD), Gangtuo (GT), and Benzilan (BZL), covering the period from 2009 to 2016. Zhimenda serves as the demarcation point between the source region of the Changjiang River and the Upper Jinsha River. All the planned hydroelectric power stations are located on the river segment stretching between Zhimenda and Benzilan. The inclusion of Gangtuo provides additional data coverage along the river's mainstream. By choosing these stations, the performance of reanalysis datasets in a region of high hydropower potential and complex hydrological processes can be assessed. The selection of this timeframe is justified by the minimal presence of hydroelectric facilities, which ensures that the river discharge data are largely unaffected by human interventions.

The dataset from Zhimenda is compiled at a daily resolution. The dataset comprises 1,477 records, with 49% of the days missing between 2009 and 2016. In contrast, Gangtuo and Benzilan provide data at an hourly resolution, accumulating 52,246 hours and 69,011 hours, respectively. The respective missing data rates for these stations are 25% for Gangtuo and 1.6% for Benzilan. Given the nature of our analysis, we have chosen not to impute the missing values, opting instead to focus solely on the complete data records. This choice was made with the aim of maximizing the utilization of available observations in this region where data are scarce. Additionally, by avoiding the use of gap-filling methods, we sought to minimize the potential biases that could be introduced, especially given the proportion of missing data is substantial.

To our knowledge, the daily river discharge observations at Zhimenda have been used to calibrate GloFAS and GRFR, while the observations at Gangtuo and Benzilan have not been incorporated into the datasets analyzed in this study.

**Table 1.** The datasets used in this study.

Dataset	Time period	Temporal Resolution	Runoff Resolution	River Routing	References
<i>Publically available datasets</i>					
ERA5-Land	1950–Present	Hourly	0.10°	N/A	Muñoz-Sabater et al. [2]
GloFAS v4.0	1979–Present	Daily	0.10°	LISFLOOD	Harrigan et al. [4]
GRFR v1.0	1980–2019	3-hourly	0.05°	RAPID	Yang et al. [5]
CNRD v1.0	1961–2018	Monthly	0.25°	N/A	Miao et al. [10]
<i>Datasets produced in this study</i>					
ERA5-Land/ NoahMP	2009–2016	Hourly	0.01°	MC	This study

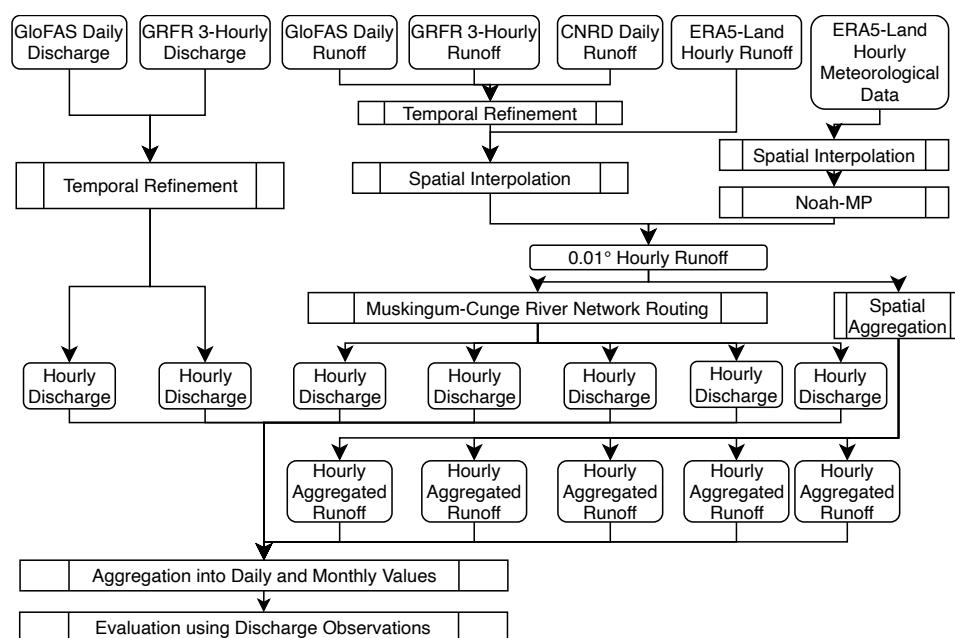
## 2.3. Evaluation methods

As shown in Table 1, the datasets vary in spatial and temporal resolution. River discharge estimates are not available in every dataset examined. To address these inconsistencies, we remapped the runoff onto a common 0.01° grid and refined both daily and monthly values to hourly values, as depicted in Figure 2. The spatial remapping was conducted using a first-order conservative method. For temporal refinement, we used the fraction of hourly ERA5-Land runoff to the accumulated runoff within a day or a month to redistribute the daily or monthly values. Subsequently, the runoff from all datasets was routed using the same Muskingum–Cunge (MC) method, as previously described, to estimate river discharge. For GloFAS and GRFR, which provide river discharge estimates,

we also refined the 3-hourly or daily values to hourly values using the same refinement method as the runoff.

The river discharge estimates derived from the Muskingum–Cunge method, as well as the temporally refined river discharge data from GRFR and GloFAS, are then compared with the observations at the gauging stations. This evaluation is conducted across three temporal scales—monthly, daily, and hourly—to comprehensively assess the accuracy and consistency of the estimates with the actual observations.

The structure and parameters of the river routing model significantly influence the river discharge estimates. To isolate the impact of river routing on these estimates, we not only compared the model-derived river discharge with observations at gauging stations but also aggregated the runoff above these stations for a more direct comparison. This aggregation helps to isolate the effects of routing by focusing on the input runoff rather than the routing process itself. The aggregated runoff is calculated by summing the runoff within the upstream area of each gauging station. The delineation of the upstream area was based on the flow direction data from the MERIT-Hydro dataset.



**Figure 2.** Workflow of this study.

#### 2.4. Evaluation metrics

We employed the Kling-Gupta Efficiency (KGE) [52] to assess the performance of the reanalysis datasets. The KGE shares similarities with the commonly used Mean-Square Error (MSE) and Nash–Sutcliffe Efficiency (NSE). It incorporates the correlation coefficient, the ratio of the standard deviations, and the ratio of the mean values of both the observed and simulated data. Unlike MSE or NSE, in the KGE, these three components are combined

in an independent manner. This enables a straightforward separation of their relative contributions to the overall metric. The KGE is formulated as follows:

$$\text{KGE} = 1 - \sqrt{G_1^2 + G_2^2 + G_3^2}, \quad (1)$$

$$G_1 = r - 1, \quad (2)$$

$$G_2 = \frac{\sigma_s}{\sigma_o} - 1, \quad (3)$$

$$G_3 = \frac{\mu_s}{\mu_o} - 1, \quad (4)$$

where  $r$  is the correlation coefficient,  $\sigma_s$  and  $\sigma_o$  are the standard deviations of the simulated and observed data, respectively, and  $\mu_s$  and  $\mu_o$  are the mean values of the simulated and observed data, respectively. The KGE ranges from  $-\infty$  to 1, with a value of 1 indicating a perfect match between the simulated and observed data. A KGE value significantly less than 1 indicates a poor match.

The relative contribution of the three terms in the KGE can be calculated using the following formula [52]:

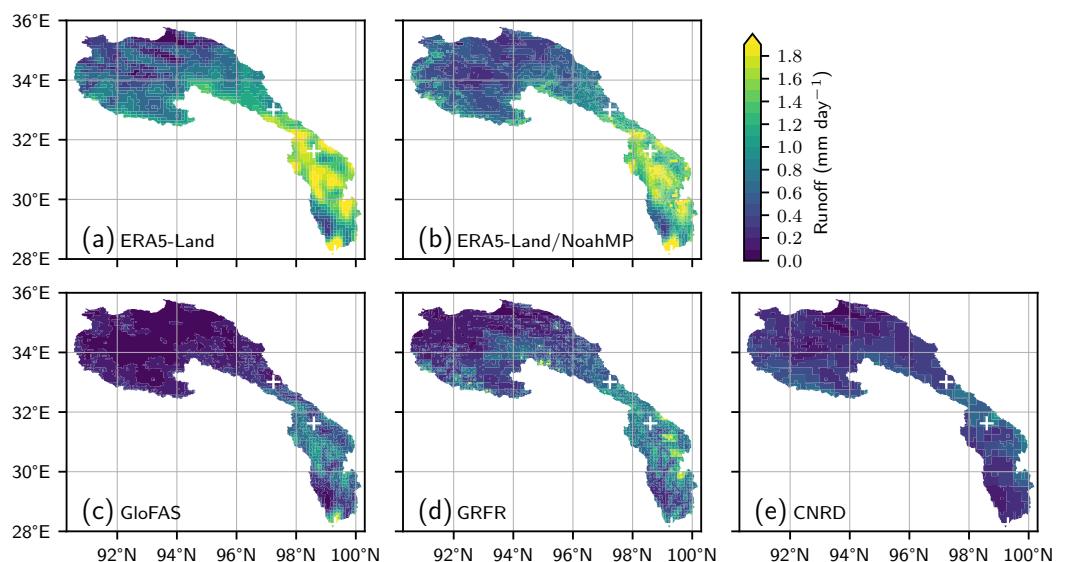
$$g_i = \frac{G_i^2}{G_1^2 + G_2^2 + G_3^2}, \quad \text{for } i = 1, 2, 3, \quad (5)$$

where  $g_1$ ,  $g_2$ , and  $g_3$  are the relative contributions of correlation coefficient, variability, and bias, respectively. A higher value indicates a greater contribution of the corresponding term to the deviation of the KGE from 1.

### 3. Results and discussion

#### 3.1. Runoff

Figure 3 geographically compares the runoff estimates within the Benizian catchment. The five datasets fall into three categories: ERA5-Land and ERA5-Land/NoahMP; GloFAS and GRFR; and CNRD. The first group utilizes uncalibrated land surface models, the second group employs extensively calibrated hydrological models, and the third group features a hydrological model with parameters that are regionalized rather than individually calibrated. Notably, ERA5-Land, ERA5-Land/NoahMP, and GloFAS exhibit significant differences despite similar meteorological forcing data. On the other hand, GloFAS and GRFR, which are driven by different meteorological forcing data, show close spatial patterns. This comparison indicates that the calibration of land surface models or hydrological models has a profound impact on runoff estimation, more than the meteorological data.

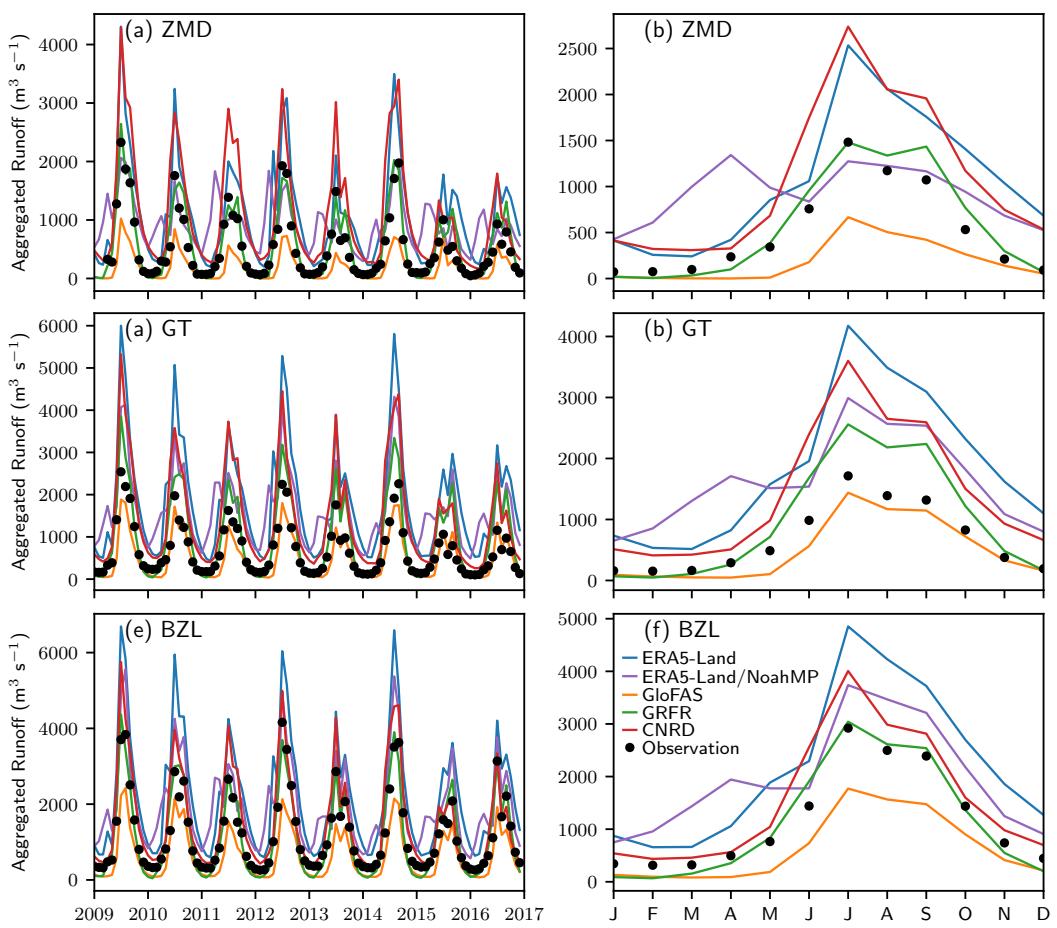


**Figure 3.** Intercomparison of the spatial distribution of multiyear averaged runoff from 2009 to 2016.

Figure 4 intercompares the monthly variations and annual cycle of the aggregated runoff above Zhimenda, Gangtuo, and Benzilan. Since the river flow travel time is typically smaller than a month [53], it is generally safe to neglect the impact of river routing at the monthly time scale. Both ERA5 and CNRD overestimate the runoff at all examined stations, but the reasons for this overestimation differ. The overestimation in ERA5 is consistent with previous studies, which attribute it to overestimated precipitation in high-altitude regions [54–57]. In contrast, the overestimation of CNRD is primarily observed in the upstream area of Zhimenda. Given that CMFD, the meteorological data that drive CNRD, has shown closeness to in-situ precipitation observations in this region [57], the overestimation of the CNRD runoff is likely due to less constrained hydrological model parameters. CNRD uses parameters calibrated at gauged basins and extrapolates these parameters to ungauged basins using multiscale regionalization functions [9,10]. The scarcity of river discharge gauges on the Tibetan Plateau means that this extrapolation can lead to bias.

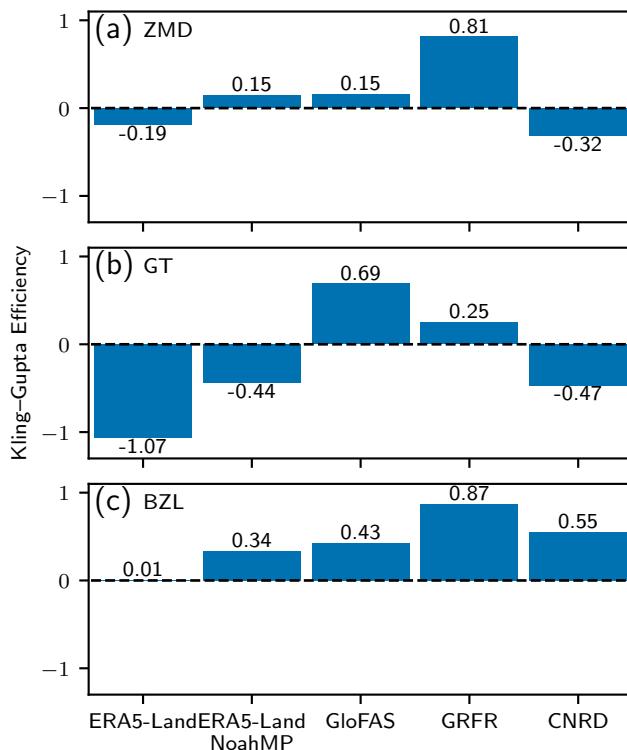
The runoff from ERA5-Land/NoahMP exhibits two peaks, particularly in the high-altitude regions above Zhimenda. In contrast, ERA5-Land, which utilizes the same meteorological data, shows an insignificant rise in runoff during Spring. The discrepancy between ERA5-Land/NoahMP and ERA5-Land is likely due to differences in the spatial resolution of the land surface model. The higher resolution of ERA5-Land/NoahMP allows for a better capture of the vertical gradient of land surface temperature. The combination of overestimated precipitation in the ERA5-Land meteorological data and increased snowmelt in low-altitude model grids [21] results in an artificial spring runoff peak not observed in actual data.

The two calibrated datasets, GloFAS and GRFR, exhibit distinct behaviors. This difference is even evident at Zhimenda, where observations are publicly available through the Global Runoff Data Centre and have been utilized in both datasets. The distinct calibration approaches of the two datasets likely account for this difference. GloFAS is calibrated against river discharge observations [41], while its runoff is not directly calibrated. Since river discharge can be calculated by multiplying the catchment area by runoff, different estimates of the catchment area will lead to different runoff values after the calibration of river discharge. In contrast, GRFR's runoff is directly calibrated against a machine-learning-derived runoff characteristics dataset [44]. The results suggest that the calibration method employed by GRFR is highly effective, yielding the best-performing runoff among the examined datasets.



**Figure 4.** Intercomparison of the aggregated runoff at the gauging stations. (a) monthly aggregated runoff at Zhimenda; (b) annual cycle of the aggregated runoff at Zhimenda; (c) monthly aggregated runoff at Gangtuo; (d) annual cycle of the aggregated runoff at Gangtuo; (e) monthly aggregated runoff at Benzilan; and (f) annual cycle of the aggregated runoff at Benzilan.

Figure 5 presents the performance of the aggregated runoff in comparison with discharge observations at the monthly time scale. ERA5-Land consistently underperforms across all stations and time scales, corresponding to the overestimation shown in Figure 4. CNRD also performs poorly in high-altitude areas above Gangtuo due to overestimation, but its performance improves at Benzilan. The performance of CNRD slightly outperforms the high-resolution ERA5-Land/NoahMP. Among the examined datasets, the two calibrated datasets, GloFAS and GRFR, perform the best at almost every station examined, highlighting the importance of calibration for land surface models or hydrological models. Notably, GRFR's performance is exceptionally good, with a KGE above 0.8 at Zhimenda and Benzilan. GRFR's high performance underscores the effectiveness of the characteristics-based calibration method. Meanwhile, GRFR shows subpar performance at Gangtuo. This underperformance can be attributed to the underestimation of both the variability and the long-term mean, as illustrated in Figure 9c. Since both variability and the long-term mean are the characteristics to which GRFR is calibrated, the poor performance of GRFR at Gangtuo implies that the machine-learning-derived runoff characteristics dataset does not accurately represent the runoff characteristics at Gangtuo. This is likely because the terrain at Gangtuo is steep, and the observations from this location were not incorporated into the training dataset.



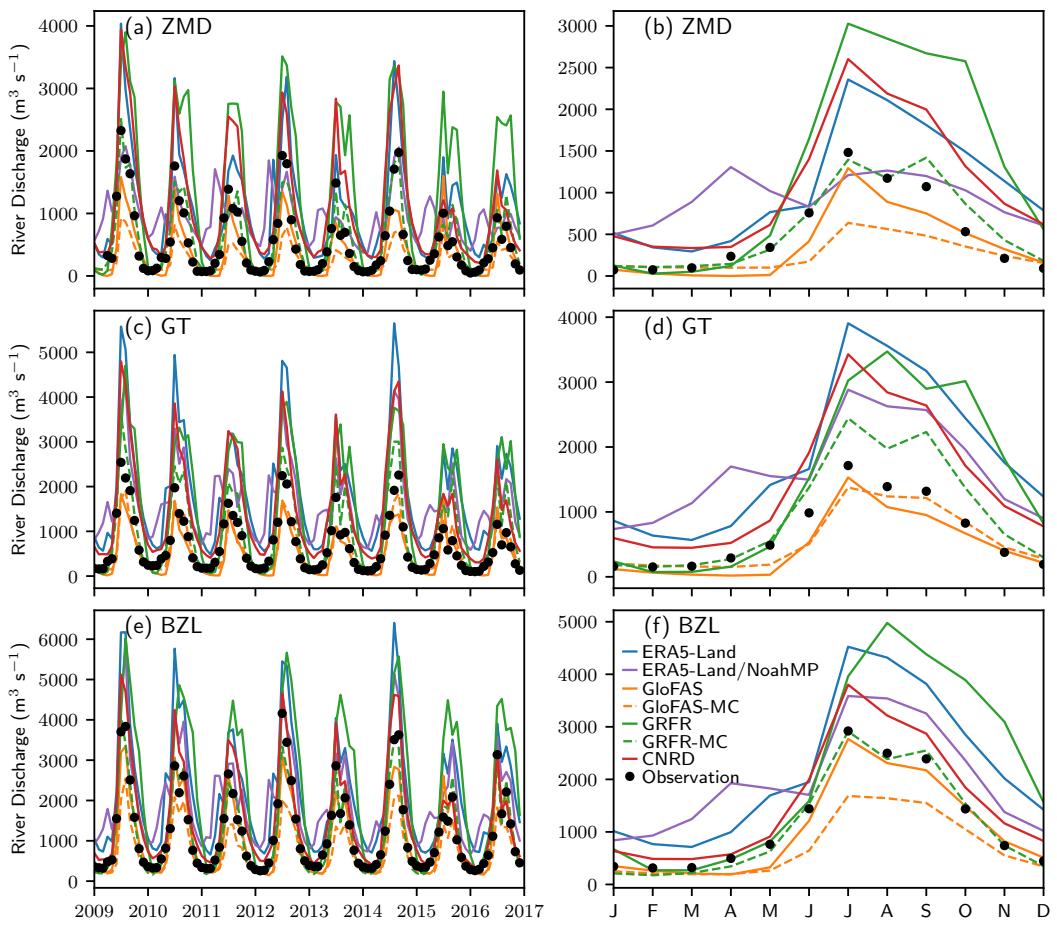
**Figure 5.** Kling–Gupta Efficiency of the monthly aggregated runoff. (a) the measures at Zhimenda, (b) the measures at Gangtuo, and (c) the measures at Benzilan.

### 3.2. Streamflow

Figure 6 compares the monthly discharge estimations with observations. The patterns generally follow the runoff pattern shown in Figure 4. The overestimation observed in ERA5-Land and CNRD closely mirrors their runoff overestimation. The artificial Spring peak, which is also present in the ERA5-Land/NoahMP discharge, aligns with the runoff pattern.

It is interesting to note that the GloFAS and GRFR discharge behave differently from the runoff. Our rerouted GloFAS runoff (GloFAS-MC in Figure 6) can reproduce the underestimated patterns observed for the runoff (Figure 4), confirming that GloFAS underestimates runoff. However, the underestimation of runoff is no longer presented in the river discharge. GloFAS closely reproduced the observed river discharge (Figure 6). The contradiction between the runoff underestimation and the unbiased discharge estimation suggests that the spatial resolution of GloFAS’s routing grid plays a role. GloFAS is routed on a  $\frac{1}{10}^\circ$  grid cell basis, whereas the river basins delineated in this study (as shown in Figure 1) are based on  $\frac{1}{1200}^\circ$  grids. The resolution of GloFAS may be too coarse to accurately delineate the river basins. However, this error is mitigated in the discharge due to the calibration, and the impact is transferred to errors in the runoff.

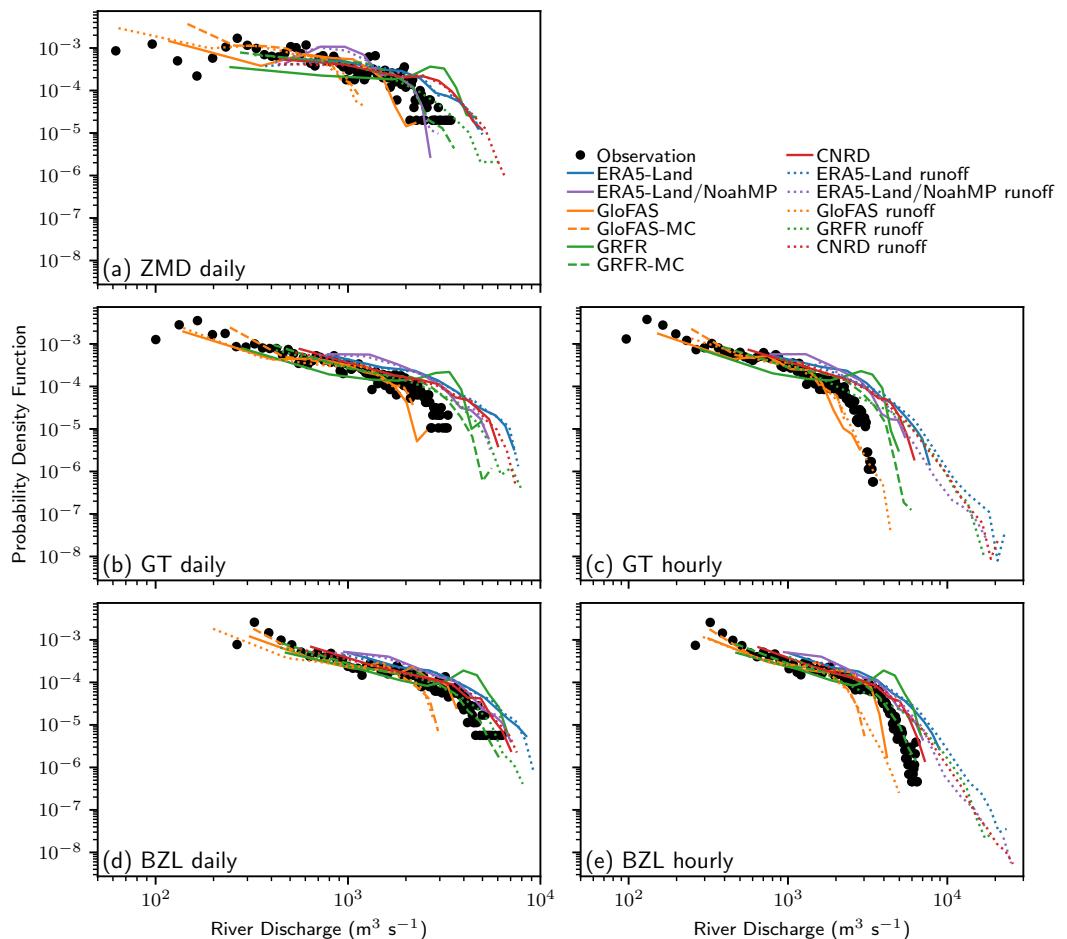
Despite the high skill of GRFR’s runoff demonstrated in Figure 4, the skill of GRFR’s discharge is unexpectedly unsatisfactory. GRFR remarkably overestimates discharge at all gauging stations. The overestimation is also reported in literature [58,59]. We argue that neither the dynamics nor the parameters of the river routing model should have such significant impacts on monthly discharge, the unexpected overestimation in GRFR is likely attributable to some inadequate configurations of the river routing model. Our rerouted discharge (GRFR-MC) using the GRFR runoff and the Muskingum–Cunge routing method can closely reproduce the observed discharge patterns. This confirms that the overestimation in GRFR is not due to the runoff but rather the routing process.



**Figure 6.** Same as Figure 4, but for river discharge.

Figure 7 presents the river discharge estimates compared with both daily and hourly observations. The aggregated runoff is also displayed for comparison. The difference between aggregated runoff and river discharge highlights the impact of river routing. The comparison shows that river routing has a marginal impact in low flow conditions and at the monthly time scale. The impact increases with the flow rate and on shorter time scales. Significant impacts of the river routing are found on the hourly high flow estimations at all gauging stations. This finding suggests the indispensable role of river routing in flood forecasting.

The differences among the examined river discharge estimates are mainly exhibited in high flow conditions. However, these differences do not vary significantly with the time scale of the estimates (i.e., daily or hourly). The GRFR river discharge estimation suspiciously indicates a high probability of discharge at approximately  $4000 \text{ m}^3 \text{ s}^{-1}$  at all the examined stations, whereas such a high probability is not found in the other estimates including GRFR-MC. This suspicious pattern is another strong evident of inadequate configurations within GRFR's routing process. The GloFAS estimate can well reproduce the observed discharge probability patterns. Our rerouted river discharge using the GloFAS runoff and the Muskingum–Cunge method deviates from the observed pattern of discharge probability. The discrepancy arises from the calibration of river discharge, which compensates for imprecise river basin area estimation.



**Figure 7.** Intercomparison of the probability distribution of river discharge and aggregated runoff at the river gauging stations. (a) daily estimates at Zhimenda; (b) hourly estimates at Zhimenda; (c) daily estimates at Gangtuo; (d) hourly estimates at Gangtuo; (e) daily estimates at Benzilan; (f) hourly estimates at Benzilan. Dotted lines denote the aggregated runoff. Solid lines denote the river discharge estimates provided by the examined datasets. Dashed lines denote the river discharge estimates using the runoff and the Muskingum–Cunge routing method. Zhimenda is not shown in the hourly estimates due to the lack of hourly observations.

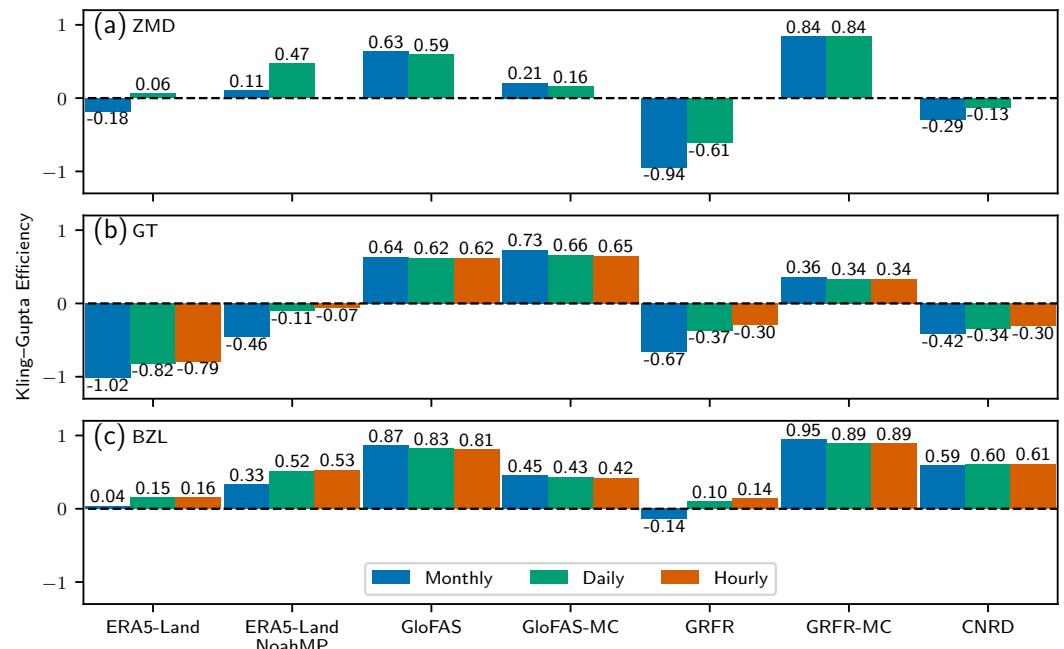
Figure 8 illustrates the skill scores for the estimation of river discharge across monthly, 338 daily, and hourly temporal scales. At the monthly scale, the river discharge estimation's 339 skill closely mirrors that of the runoff, as depicted in Figure 5. The coincidence hints 340 that the primary determinant of river discharge variability within the study area is the runoff 341 estimation, rather than the river routing process. An exception to this pattern is observed 342 with GloFAS, where the divergence in discharge skill scores from those of runoff is likely 343 due to the coarse and imprecise river routing grid employed, as previously discussed. 344

Among the examined dataset, GloFAS, GloFAS-MC, and GRFR-MC perform the best 345 at all gauging stations and across all time scales. Proper calibration of the land surface 346 model and the hydrological model is essential for accurate river discharge estimation. 347 As revealed in Figure 9, the primary reasons for the poor performance of the examined 348 datasets, except for GRFR-MC, are errors in reproducing the observed mean and variability 349 of river discharge. GRFR-MC can effectively reproduce the observed mean and variability 350 of river discharge. The KGE of GRFR-MC's hourly discharge can reach a high score of 351 approximately 0.9 at Benzilan, which again confirms the effectiveness of the characteristics- 352 based calibration method. The relatively lower performance of GRFR at Gangtuo implies 353 that the machine-learning-based calibration method employed by GRFR needs to incorporate 354 more local observations to enhance its applicability in specific regions. At all three 355

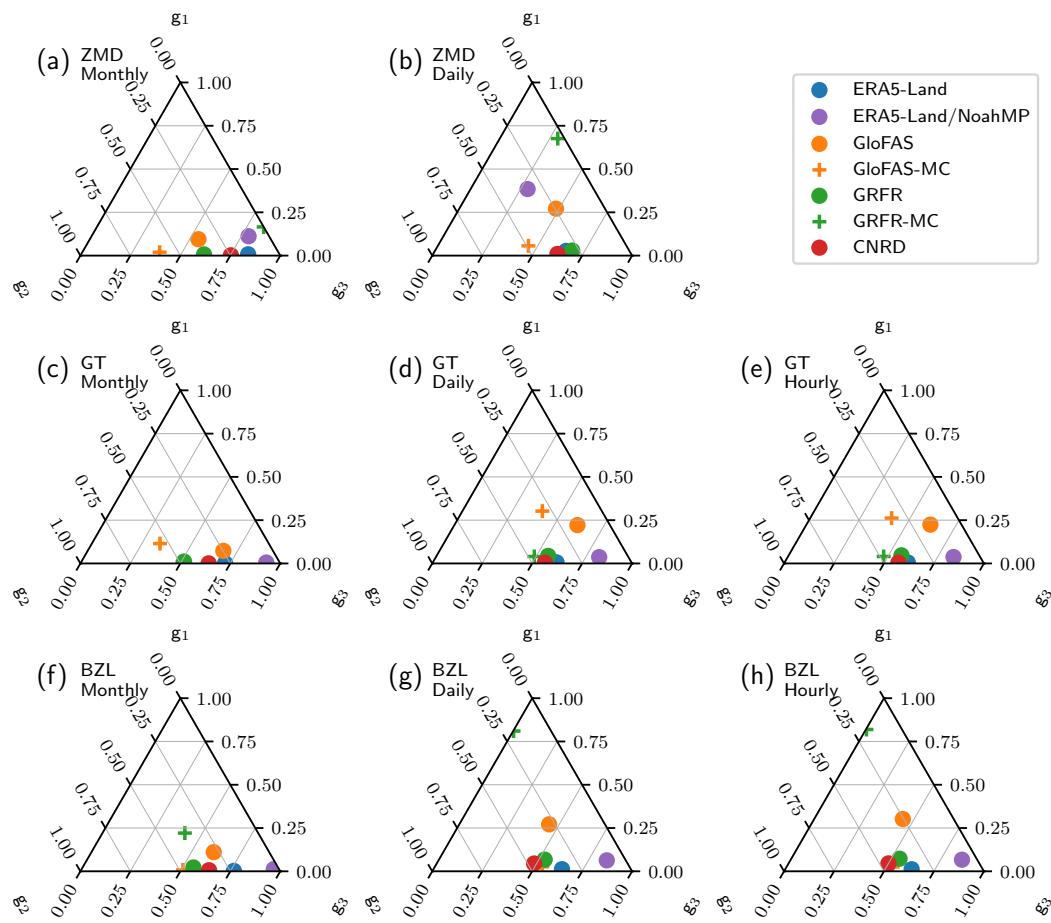
gauging stations, the deviation of the KGE from 1 for GRFR-MC is mainly attributed to the correlation coefficient, which can be further improved through the calibration of the river routing parameters. The exceptionally high skill of GRFR-MC again confirms the effectiveness of the characteristics-based calibration method.

It is worth noting that the KGE (Figure 8) and the relative contribution within the KGE (Figure 9) are consistent for both hourly and daily estimates. This consistency suggests that calibration at the daily time scale could potentially be applied to the hourly time scale. The relative contributions of the correlation coefficient, variability, and bias, differ, however, between the monthly, daily, and hourly time scales, as illustrated in Figure 9. At the monthly scale, the datasets under examination mainly exhibit differences in the variability and bias of river discharge estimations. The correlation coefficient plays a relatively minor role, indicating that all the examined datasets are capable of reasonably capturing the timing of the peak discharge, as evidenced by Figure 6. In contrast, at the daily and hourly scales, the correlation coefficient emerges as a crucial factor. This suggests that the examined datasets face challenges in accurately reproducing the observed timing of floods.

The KGE of river discharge is also consistent between daily and monthly estimates for most datasets, with the exception of the high-resolution ERA5-Land/NoahMP estimation (Figure 8). The increase in KGE from the monthly to daily time scale indicates that a more accurate representation of the spatial distribution of runoff and river routing could improve the temporal variations in river discharge estimates, aligning with the concept of the time of concentration [60]. However, these findings are based on a limited number of gauging stations and limited time periods. Further studies are needed to validate these preliminary observations.



**Figure 8.** Kling–Gupta Efficiency of the river discharge estimation. (a) the measures at the Zhimenda station, (b) the measures at the Gangtuo station, and (c) the measures at the Benzilan station.



**Figure 9.** Ternary plot showing the relative contribution of correlation coefficient ( $g_1$ ), variability ( $g_2$ ), and bias ( $g_3$ ) to the deviation of the Kling–Gupta Efficiency of simulated river discharge from one. The three rows of the plot represent the estimation at Zhimenda, Gangtuo, and Benzilan, respectively. The three columns of the plot represent the results of the monthly, daily, and hourly estimates, respectively.

#### 4. Conclusions

This study presents a pilot intercomparison of several publicly available large-domain reanalysis datasets of runoff and river discharge in the Upper Jinsha River basin. The datasets under scrutiny include ERA5-Land, GloFAS, GRFR, and CNRD. To enhance the understanding of the effects of calibration, spatial resolution, and meteorological forcing on runoff estimation, a high-resolution runoff simulation has been incorporated into the comparison. The runoff from the examined datasets is rerouted on a high-resolution river network delineated from the 3-arcsecond MERIT-Hydro dataset. This rerouting aids in identifying the impacts of different river routing models and river network delineations on river discharge estimation. The estimated runoff and river discharge are compared with observations at three gauging stations—Zhimenda, Gangtuo, and Benzilan—spanning monthly, daily, and hourly time scales. The principal findings of this study are as follows:

- The examined runoff and river discharge reanalysis datasets exhibit significant discrepancies. When compared with observations, the Kling-Gupta Efficiency of river discharge closely aligns with that of runoff. The primary driver of river discharge variability in the study area is runoff estimation, rather than the river routing process. The data and models used in runoff estimation need further scrutinization.
- Variations in runoff estimates are predominantly attributed to the calibration processes of land surface and hydrological models. Among the datasets, GloFAS and GRFR,

which have been calibrated, demonstrate superior performance in runoff estimation. GloFAS and GRFR show similar runoff patterns, despite being driven by different meteorological forcing data. In contrast, the runoff estimates from GloFAS markedly diverge from those of ERA5-Land, even though the meteorological forcing data are similar.

- The runoff characteristics-based calibration method is demonstrated to be highly effective. GRFR, which utilizes this calibration approach, outperforms all the other examined datasets in terms of runoff estimation at Zhimenda and Benzilan. However, this machine-learning-based method is sensitive to the training dataset used. At Gangtuo, GRFR's performance is subpar. This poor performance suggests that the runoff characteristics relevant to Gangtuo are not adequately represented in the calibration dataset. To improve the applicability of this method in specific regions like Gangtuo, more local observations are needed.
- Intercomparison proves to be a valuable approach for uncovering the shortcomings within reanalysis datasets. Although the GRFR demonstrates a high level of proficiency in runoff estimation, its performance in discharge estimation shows an unexpected decline. Notably, an uncalibrated re-routing process that employs the Muskingum-Cunge method, utilizing the GRFR runoff data, is able to achieve the most favorable performance in estimating river discharge when compared to the other examined datasets. This outcome strongly implies that the river routing model employed by GRFR is insufficient and requires improvement.
- The river routing process is especially important during high-flow conditions and at sub-monthly time scales. This finding underscores the importance of incorporating a river routing model in flood forecasting. Conversely, the impact of river routing is minimal in low flow conditions and at the monthly time scale. The similarity between monthly aggregated runoff and river discharge estimates allows examination of the adequacy of the river routing model and the river network delineation.
- A high-resolution river network for flow routing is crucial for concurrent runoff and river discharge estimation. The intercomparison shows a contradiction between GloFAS's high skill in river discharge estimation and relatively low skill in runoff estimation. The contradiction is likely due to the coarse river routing grid. The coarse grid introduces errors in catchment area estimation, which are transferred to the runoff estimates during the calibration of river discharge.
- High-resolution runoff simulations exhibit distinct runoff characteristics compared with their low-resolution counterparts. Calibrating high-resolution runoff models can mitigate these discrepancies, but inherent differences in parameter settings may still lead to persistent divergences in other aspects of hydrological cycle.

Given the scarcity of in-situ observations in the Upper Jinsha River, reanalysis datasets are indeed irreplaceable for these regions. The findings of this study have implications for both the utilization of existing reanalysis datasets and future development of new ones. For users of reanalysis datasets, the datasets investigated in this study display notable variances in the estimation of runoff and river discharge. Generally, GloFAS has been found to be relatively more reliable for estimating river discharge, while GRFR shows more promise for runoff estimation. For users with prior modeling experience, rerouting the runoff data from GRFR can be beneficial.

For those involved in developing reanalysis datasets, calibrating land surface and hydrological models is a critical step in obtaining accurate runoff and river discharge estimates. The calibration process should be carefully constructed to incorporate as many relevant observations as possible within the target region, to better capture the distinct runoff characteristics. In terms of river routing methods, the vector-based approach demonstrates

advantages over the grid-based method. The vector-based method is better represents the detailed geometry of the river network and define the catchment area accurately, which are essential factors for achieving reliable runoff and river discharge estimations.

**Author Contributions:** Conceptualization, H.Z.; methodology, H.Z.; software, H.Z.; validation, S.C. and H.Y.; formal analysis, S.C. and H.Y.; investigation, S.C. and H.Y.; resources, S.C., H.Y. and H.Z.; data curation, S.C. and H.Y.; writing—original draft preparation, S.C. and H.Y.; writing—review and editing, S.C., H.Y., and H.Z.; visualization, S.C. and H.Y.; supervision, H.Z.; project administration, H.Z.; funding acquisition, H.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study is funded by China Three Gorges Corporation (contract Z532302035) and the Natural Science Foundation of China (grants 42075165 and 42275178).

**Data Availability Statement:** The MERIT-Hydro dataset was obtained from [https://hydro.iis.u-tokyo.ac.jp/~yamadai/MERIT\\_Hydro/](https://hydro.iis.u-tokyo.ac.jp/~yamadai/MERIT_Hydro/) (accessed on 2024-11-10). The ERA5-Land dataset was obtained from <https://cds.climate.copernicus.eu/datasets/reanalysis-era5-land>. The GloFAS v4.0 dataset was obtained from <https://ewds.climate.copernicus.eu/datasets/cems-glofas-historical> (accessed on 2024-11-10). The GRFR v1.0 dataset was obtained from <https://www.reachhydro.org/home/records/grfr> (accessed on 2024-11-10). The COMIDs used to extract the river discharge from GRFR are 43037713 for Zhimenda, 43046824 for Gangtuo, and 43066085 for Benzilan, respectively. The CNRD v1.0 dataset was obtained from <https://doi.org/10.11888/Atmos.tpdc.272864> (accessed on 2024-11-10). The river discharge observations were obtained from the China Three Gorges Corporation; however, the authors do not have permission to share the data. The delineated river network, the aggregated runoff and simulated river discharge at the gauging stations, along with the scripts for figure generation used throughout this paper, can be found at the GitHub repository <https://github.com/hzheng88/paper-2025-upperjsj-streamflow-intercomp> (accessed on 2025-02-03).

**Acknowledgments:** We thank for the technical support of the National Large Scientific Infrastructure “Earth System Numerical Simulation Facility” (<https://cstr.cn/31134.02.EL>)

**Conflicts of Interest:** Authors Shuanglong Chen and Heng Yang were employed by the China Three Gorges Corporation. The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

1. Hersbach, H.; Bell, B.; Berrisford, P.; Hirahara, S.; Horányi, A.; Muñoz-Sabater, J.; Nicolas, J.; Peubey, C.; Radu, R.; Schepers, D.; et al. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society* **2020**, *146*, 1999–2049. <https://doi.org/10.1002/qj.3803>.
2. Muñoz-Sabater, J.; Dutra, E.; Agustí-Panareda, A.; Albergel, C.; Arduini, G.; Balsamo, G.; Boussetta, S.; Choulga, M.; Harrigan, S.; Hersbach, H.; et al. ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data* **2021**, *13*, 4349–4383. <https://doi.org/10.5194/essd-13-4349-2021>.
3. Alfieri, L.; Lorini, V.; Hirpa, F.A.; Harrigan, S.; Zsoter, E.; Prudhomme, C.; Salamon, P. A global streamflow reanalysis for 1980–2018. *Journal of Hydrology X* **2020**, *6*, 100049. <https://doi.org/10.1016/j.hydrona.2019.100049>.
4. Harrigan, S.; Zsoter, E.; Alfieri, L.; Prudhomme, C.; Salamon, P.; Wetterhall, F.; Barnard, C.; Cloke, H.; Pappenberger, F. GloFAS-ERA5 operational global river discharge reanalysis 1979–present. *Earth System Science Data* **2020**, *12*, 2043–2060. <https://doi.org/10.5194/essd-12-2043-2020>.
5. Yang, Y.; Pan, M.; Lin, P.; Beck, H.E.; Zeng, Z.; Yamazaki, D.; David, C.H.; Lu, H.; Yang, K.; Hong, Y.; et al. Global reach-level 3-hourly river flood reanalysis (1980–2019). *Bulletin of the American Meteorological Society* **2021**, *102*, E2086–E2105. <https://doi.org/10.1175/BAMS-D-20-0057.1>.
6. Lin, P.; Pan, M.; Beck, H.E.; Yang, Y.; Yamazaki, D.; Frasson, R.; David, C.H.; Durand, M.; Pavelsky, T.M.; Allen, G.H.; et al. Global reconstruction of naturalized river flows at 2.94 million reaches. *Water Resources Research* **2019**, *55*, 6499–6516. <https://doi.org/10.1029/2019WR025287>.
7. Cosgrove, B.; Gochis, D.; Flowers, T.; Dugger, A.; Ogden, F.; Graziano, T.; Clark, E.; Cabell, R.; Casiday, N.; Cui, Z.; et al. NOAA’s National Water Model: Advancing operational hydrology through continental-scale modeling. *Journal of the American Water Resources Association* **2024**, *60*, 247–272. <https://doi.org/10.1111/1752-1688.13184>.

8. Najafi, H.; Shrestha, P.K.; Rakovec, O.; Apel, H.; Vorogushyn, S.; Kumar, R.; Thober, S.; Merz, B.; Samaniego, L. High-resolution impact-based early warning system for riverine flooding. *Nature Communications* **2024**, *15*, 3726. <https://doi.org/10.1038/s41467-024-48065-y>. 499  
500
9. 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. *Bulletin of the American Meteorological Society* **2021**, *102*, E929–E947. <https://doi.org/10.1175/BAMS-D-20-0094.1>. 501  
502  
503
10. 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. *Science Bulletin* **2022**, *67*, 547–556. <https://doi.org/10.1016/j.scib.2021.09.022>. 504  
505
11. Feng, D.; Gleason, C.J.; Lin, P.; Yang, X.; Pan, M.; Ishitsuka, Y. Recent changes to Arctic river discharge. *Nature Communications* **2021**, *12*, 6917. <https://doi.org/10.1038/s41467-021-27228-1>. 506  
507
12. Collins, E.L.; David, C.H.; Riggs, R.; Allen, G.H.; Pavelsky, T.M.; Lin, P.; Pan, M.; Yamazaki, D.; Meentemeyer, R.K.; Sanchez, G.M. Global patterns in river water storage dependent on residence time. *Nature Geoscience* **2024**, *17*, 433–439. <https://doi.org/10.1038/s41561-024-01421-5>. 508  
509  
510
13. Liu, M.; Raymond, P.A.; Lauerwald, R.; Zhang, Q.; Trapp-Müller, G.; Davis, K.L.; Moosdorf, N.; Xiao, C.; Middelburg, J.J.; Bouwman, A.F.; et al. Global riverine land-to-ocean carbon export constrained by observations and multi-model assessment. *Nature Geoscience* **2024**, *17*, 896–904. <https://doi.org/10.1038/s41561-024-01524-z>. 511  
512  
513
14. Coughlan de Perez, E.; van den Hurk, B.; van Aalst, M.K.; Amuron, I.; Bamanya, D.; Hauser, T.; Jongma, B.; Lopez, A.; Mason, S.; Mendler de Suarez, J.; et al. Action-based flood forecasting for triggering humanitarian action. *Hydrology and Earth System Sciences* **2016**, *20*, 3549–3560. <https://doi.org/10.5194/hess-20-3549-2016>. 514  
515  
516
15. Sun, Q.; Miao, C.; Duan, Q.; Ashouri, H.; Sorooshian, S.; Hsu, K.L. A review of global precipitation data sets: data sources, estimation, and intercomparisons. *Reviews of Geophysics* **2018**, *56*, 79–107. <https://doi.org/10.1002/2017RG000574>. 517  
518
16. Henn, B.; Newman, A.J.; Livneh, B.; Daly, C.; Lundquist, J.D. An assessment of differences in gridded precipitation datasets in complex terrain. *Journal of Hydrology* **2018**, *556*, 1205–1219. <https://doi.org/10.1016/j.jhydrol.2017.03.008>. 519  
520
17. Renard, B.; Kavetski, D.; Kuczera, G.; Thyre, M.; Franks, S.W. Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors. *Water Resources Research* **2010**, *46*, W05521. <https://doi.org/10.1029/2009WR008328>. 521  
522  
523
18. Bai, F.; Wei, Z.; Wei, N.; Lu, X.; Yuan, H.; Zhang, S.; Liu, S.; Zhang, Y.; Li, X.; Dai, Y. Global assessment of atmospheric forcing uncertainties in the Common Land Model 2024 simulations. *Journal of Geophysical Research: Atmospheres* **2024**, *129*, e2024JD041520. <https://doi.org/10.1029/2024JD041520>. 524  
525  
526
19. Butts, M.B.; Payne, J.T.; Kristensen, M.; Madsen, H. An evaluation of the impact of model structure on hydrological modelling uncertainty for streamflow simulation. *Journal of Hydrology* **2004**, *298*, 242–266. <https://doi.org/10.1016/j.jhydrol.2004.03.042>. 527  
528
20. Haddeland, I.; Matheussen, B.V.; Lettenmaier, D.P. Influence of spatial resolution on simulated streamflow in a macroscale hydrologic model. *Water Resources Research* **2002**, *38*, 29–1–29–10. <https://doi.org/10.1029/2001WR000854>. 529  
530
21. Barnhart, T.B.; Putman, A.L.; Heldmyer, A.J.; Rey, D.M.; Hammond, J.C.; Driscoll, J.M.; Sexstone, G.A. Evaluating distributed snow model resolution and meteorology parameterizations against streamflow observations: finer is not always better. *Water Resources Research* **2024**, *60*, e2023WR035982. <https://doi.org/10.1029/2023WR035982>. 531  
532  
533
22. Li, L.; Bisht, G.; Leung, L.R. Spatial heterogeneity effects on land surface modeling of water and energy partitioning. *Geoscientific Model Development* **2022**, *15*, 5489–5510. <https://doi.org/10.5194/gmd-15-5489-2022>. 534  
535
23. Zheng, H.; Yang, Z.L.; Lin, P.; Wei, J.; Wu, W.Y.; Li, L.; Zhao, L.; Wang, S. On the sensitivity of the precipitation partitioning into evapotranspiration and runoff in land surface parameterizations. *Water Resources Research* **2019**, *55*, 95–111. <https://doi.org/10.1029/2017WR022236>. 536  
537  
538
24. Zheng, H.; Fei, W.; Yang, Z.L.; Wei, J.; Zhao, L.; Li, L.; Wang, S. An ensemble of 48 physically perturbed model estimates of the 1/8° terrestrial water budget over the conterminous United States, 1980–2015. *Earth System Science Data* **2023**, *15*, 2755–2780. <https://doi.org/10.5194/essd-15-2755-2023>. 539  
540  
541
25. Li, J.; Chen, F.; Lu, X.; Gong, W.; Zhang, G.; Gan, Y. Quantifying contributions of uncertainties in physical parameterization schemes and model parameters to overall errors in noah-mp dynamic vegetation modeling. *Journal of Advances in Modeling Earth Systems* **2020**, *12*, e2019MS001914. <https://doi.org/10.1029/2019MS001914>. 542  
543  
544
26. Saxe, S.; Farmer, W.; Driscoll, J.; Hogue, T.S. Implications of model selection: a comparison of publicly available, conterminous US-extent hydrologic component estimates. *Hydrology and Earth System Sciences* **2021**, *25*, 1529–1568. <https://doi.org/10.5194/hess-25-1529-2021>. 545  
546  
547
27. Beck, H.E.; van Dijk, A.I.J.M.; de Roo, A.; Dutra, E.; Fink, G.; Orth, R.; Schellekens, J. Global evaluation of runoff from 10 state-of-the-art hydrological models. *Hydrology and Earth System Sciences* **2017**, *21*, 2881–2903. <https://doi.org/10.5194/hess-21-2881-2017>. 548  
549
28. Fry, L.M.; Gronewold, A.D.; Fortin, V.; Buan, S.; Clites, A.H.; Luukkonen, C.; Holtschlag, D.; Diamond, L.; Hunter, T.; Seglenieks, F.; et al. The Great Lakes Runoff Intercomparison Project Phase 1: Lake Michigan (GRIP-M). *Journal of Hydrology* **2014**, *519*, 3448–3465. <https://doi.org/10.1016/j.jhydrol.2014.07.021>. 550  
551  
552  
553

29. Mai, J.; Shen, H.; Tolson, B.A.; Gaborit, E.; Arsenault, R.; Craig, J.R.; Fortin, V.; Fry, L.M.; Gauch, M.; Klotz, D.; et al. The Great Lakes Runoff Intercomparison Project Phase 4: the Great Lakes (GRIP-GL). *Hydrology and Earth System Sciences* **2022**, *26*, 3537–3572. <https://doi.org/10.5194/hess-26-3537-2022>. 553
30. Ahmed, M.I.; Stadnyk, T.; Pietroniro, A.; Awoye, H.; Bajracharya, A.; Mai, J.; Tolson, B.A.; Shen, H.; Craig, J.R.; Gervais, M.; et al. Learning from hydrological models' challenges: A case study from the Nelson basin model intercomparison project. *Journal of Hydrology* **2023**, *623*, 129820. <https://doi.org/10.1016/j.jhydrol.2023.129820>. 555
31. Henderson-Sellers, A.; Yang, Z.L.; Dickinson, R.E. The Project for Intercomparison of Land-surface Parameterization Schemes. *Bulletin of the American Meteorological Society* **1993**, *74*, 1335–1349. [https://doi.org/10.1175/1520-0477\(1993\)074<1335:TPFIOL>2.0.CO;2](https://doi.org/10.1175/1520-0477(1993)074<1335:TPFIOL>2.0.CO;2). 559
32. van den Hurk, B.; Kim, H.; Krinner, G.; Seneviratne, S.I.; Derksen, C.; Oki, T.; Douville, H.; Colin, J.; Ducharme, A.; Cheruy, F.; et al. LS3MIP (v1.0) contribution to CMIP6: the Land Surface, Snow and Soil moisture Model Intercomparison Project – aims, setup and expected outcome. *Geoscientific Model Development* **2016**, *9*, 2809–2832. <https://doi.org/10.5194/gmd-9-2809-2016>. 561
33. Reed, S.; Koren, V.; Smith, M.; Zhang, Z.; Moreda, F.; Seo, D.J.; DMIP Participants. Overall Distributed Model Intercomparison Project results. *Journal of Hydrology* **2004**, *298*, 27–60. <https://doi.org/10.1016/j.jhydrol.2004.03.031>. 565
34. Smith, M.B.; Seo, D.J.; Koren, V.I.; Reed, S.M.; Zhang, Z.; Duan, Q.; Moreda, F.; Cong, S. The Distributed Model Intercomparison Project (DMIP): motivation and experiment design. *Journal of Hydrology* **2004**, *298*, 4–26. <https://doi.org/10.1016/j.jhydrol.2004.03.040>. 566
35. Tijerina, D.; Condon, L.; FitzGerald, K.; Dugger, A.; O'Neill, M.M.; Sampson, K.; Gochis, D.; Maxwell, R. Continental hydrologic intercomparison project, phase 1: A large-scale hydrologic model comparison over the continental united states. *Water Resources Research* **2021**, *57*, e2020WR028931. <https://doi.org/10.1029/2020WR028931>. 571
36. Yamazaki, D.; Ikeshima, D.; Sosa, J.; Bates, P.D.; Allen, G.H.; Pavelsky, T.M. MERIT Hydro: A high-resolution global hydrography map based on latest topography dataset. *Water Resources Research* **2019**, *55*, 5053–5073. <https://doi.org/10.1029/2019wr024873>. 573
37. Yamazaki, D.; Ikeshima, D.; Tawatari, R.; Yamaguchi, T.; O'Loughlin, F.; Neal, J.C.; Sampson, C.C.; Kanae, S.; Bates, P.D. A high-accuracy map of global terrain elevations. *Geophysical Research Letters* **2017**, *44*, 5844–5853. <https://doi.org/10.1002/2017GL072874>. 575
38. Boussetta, S.; Balsamo, G.; Arduini, G.; Dutra, E.; McNorton, J.; Choulga, M.; Agustí-Panareda, A.; Beljaars, A.; Wedi, N.; Munoz-Sabater, J.; et al. ECLand: The ECMWF land surface modelling system. *Atmosphere* **2021**, *12*, 723. <https://doi.org/10.3390/atmos12060723>. 578
39. Balsamo, G.; Beljaars, A.; Scipal, K.; Viterbo, P.; van den Hurk, B.; Hirschi, M.; Betts, A.K. A revised hydrology for the ECMWF model: verification from field site to terrestrial water storage and impact in the Integrated Forecast System. *Journal of Hydrometeorology* **2009**, *10*, 623–643. <https://doi.org/10.1175/2008JHM1068.1>. 581
40. Knijff, J.M.V.D.; Younis, J.; Roo, A.P.J.D. LISFLOOD: a GIS-based distributed model for river basin scale water balance and flood simulation. *International Journal of Geographical Information Science* **2010**, *24*, 189–212. <https://doi.org/10.1080/13658810802549154>. 584
41. Hirpa, F.A.; Salamon, P.; Beck, H.E.; Lorini, V.; Alfieri, L.; Zsoter, E.; Dadson, S.J. Calibration of the Global Flood Awareness System (GloFAS) using daily streamflow data. *Journal of Hydrology* **2018**, *566*, 595–606. <https://doi.org/10.1016/j.jhydrol.2018.09.052>. 586
42. Liang, X.; Lettenmaier, D.P.; Wood, E.F.; Burges, S.J. A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research: Atmospheres* **1994**, *99*, 14415–14428. <https://doi.org/10.1029/94JD00483>. 588
43. Beck, H.E.; Wood, E.F.; Pan, M.; Fisher, C.K.; Miralles, D.G.; van Dijk, A.I.J.M.; McVicar, T.R.; Adler, R.F. MSWEP V2 global 3-hourly 0.1° precipitation: Methodology and quantitative assessment. *Bulletin of the American Meteorological Society* **2019**, *100*, 473–500. <https://doi.org/10.1175/BAMS-D-17-0138.1>. 591
44. Yang, Y.; Pan, M.; Beck, H.E.; Fisher, C.K.; Beighley, R.E.; Kao, S.C.; Hong, Y.; Wood, E.F. In quest of calibration density and consistency in hydrologic modeling: distributed parameter calibration against streamflow characteristics. *Water Resources Research* **2019**, *55*, 7784–7803. <https://doi.org/10.1029/2018wr024178>. 594
45. Beck, H.E.; de Roo, A.; van Dijk, A.I.J.M. Global maps of streamflow characteristics based on observations from several thousand catchments. *Journal of Hydrometeorology* **2015**, *16*, 1478–1501. <https://doi.org/10.1175/JHM-D-14-0155.1>. 597
46. David, C.H.; Maidment, D.R.; Niu, G.Y.; Yang, Z.L.; Habets, F.; Eikhout, V. River network routing on the NHDPlus dataset. *Journal of Hydrometeorology* **2011**, *12*, 913–934. <https://doi.org/10.1175/2011JHM1345.1>. 599
47. Lin, P.; Pan, M.; Wood, E.F.; Yamazaki, D.; Allen, G.H. A new vector-based global river network dataset accounting for variable drainage density. *Scientific Data* **2021**, *8*, 28. <https://doi.org/10.1038/s41597-021-00819-9>. 601
48. 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. *Scientific Data* **2020**, *7*, 25. <https://doi.org/10.1038/s41597-020-0369-y>. 603
49. Samaniego, L.; Bárdossy, A.; Kumar, R. Streamflow prediction in ungauged catchments using copula-based dissimilarity measures. *Water Resources Research* **2010**, *46*, W02506. <https://doi.org/10.1029/2008WR007695>. 605

50. Niu, G.Y.; Yang, Z.L.; Mitchell, K.E.; Chen, F.; Ek, M.B.; Barlage, M.; Kumar, A.; Manning, K.; Niyogi, D.; Rosero, E.; et al. The  
community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with  
local-scale measurements. *Journal of Geophysical Research: Atmospheres* **2011**, *116*, D12109. <https://doi.org/10.1029/2010JD015139>.  
608  
609
51. Read, L.K.; Yates, D.N.; McCreight, J.M.; Rafieeinab, A.; Sampson, K.; Gochis, D.J. Development and evaluation of the channel  
routing model and parameters within the National Water Model. *JAWRA Journal of the American Water Resources Association* **2023**,  
610  
59, 1051–1066. <https://doi.org/10.1111/1752-1688.13134>.  
611  
612
52. Gupta, H.V.; Kling, H.; Yilmaz, K.K.; Martinez, G.F. Decomposition of the mean squared error and NSE performance criteria:  
613  
Implications for improving hydrological modelling. *Journal of Hydrology* **2009**, *377*, 80–91. <https://doi.org/10.1016/j.jhydrol.2009.08.003>.  
614  
615
53. Allen, G.H.; David, C.H.; Andreadis, K.M.; Hossain, F.; Famiglietti, J.S. Global estimates of river flow wave travel times and  
616  
implications for low-latency satellite data. *Geophysical Research Letters* **2018**, *45*, 7551–7560. <https://doi.org/10.1029/2018GL077914>.  
617  
618
54. Bian, Q.; Xu, Z.; Zhao, L.; Zhang, Y.F.; Zheng, H.; Shi, C.; Zhang, S.; Xie, C.; Yang, Z.L. Evaluation and intercomparison of  
619  
multiple snow water equivalent products over the Tibetan Plateau. *Journal of Hydrometeorology* **2019**, *20*, 2043–2055. <https://doi.org/10.1175/JHM-D-19-0011.1>.  
620  
621
55. Bian, Q.; Xu, Z.; Zheng, H.; Li, K.; Liang, J.; Fei, W.; Shi, C.; Zhang, S.; Yang, Z.L. Multiscale changes in snow over the Tibetan  
622  
Plateau during 1980–2018 represented by reanalysis data sets and satellite observations. *Journal of Geophysical Research: Atmospheres*  
623  
**2020**, *125*, e2019JD031914. <https://doi.org/10.1029/2019JD031914>.  
624  
625
56. Sun, H.; Su, F.; Yao, T.; He, Z.; Tang, G.; Huang, J.; Zheng, B.; Meng, F.; Ou, T.; Chen, D. General overestimation of ERA5  
626  
precipitation in flow simulations for High Mountain Asia basins. *Environmental Research Communications* **2021**, *3*, 121003.  
627  
<https://doi.org/10.1088/2515-7620/ac40f0>.
57. Wang, X.; Zhou, J.; Ma, J.; Luo, P.; Fu, X.; Feng, X.; Zhang, X.; Jia, Z.; Wang, X.; Huang, X. Evaluation and comparison of  
628  
reanalysis data for runoff simulation in the data-scarce watersheds of alpine regions. *Remote Sensing* **2024**, *16*, 751. <https://doi.org/10.3390/rs16050751>.  
629  
630
58. Tu, T.; Wang, J.; Zhao, G.; Zhao, T.; Dong, X. Scaling from global to regional river flow with global hydrological models: Choice  
631  
matters. *Journal of Hydrology* **2024**, *633*, 130960. <https://doi.org/10.1016/j.jhydrol.2024.130960>.  
632
59. Cui, H.; Huang, C. Accuracy evaluation of multiple runoff products: a case study of the middle reaches of the Yellow River.  
633  
*Water* **2025**, *17*, 461. <https://doi.org/10.3390/w17030461>.  
634
60. Beven, K.J. A history of the concept of time of concentration. *Hydrology and Earth System Sciences* **2020**, *24*, 2655–2670.  
635  
<https://doi.org/10.5194/hess-24-2655-2020>.  
636

**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.

637  
638  
639