



Evaluations and improvements of GLDAS 2.0 and 2.1 forcing data's applicability for basin scale hydrological simulations in the Tibetan Plateau

Wei Qi^{a, b, c}, Junguo Liu^{a, b, c*} and Deliang Chen^{d, e}

^a School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China

^b Guangdong Provincial Key Laboratory of Soil and Groundwater Pollution Control, School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China

^c State Environmental Protection Key Laboratory of Integrated Surface Water-Groundwater Pollution Control, School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China

^d Regional Climate Group, Department of Earth Sciences, University of Gothenburg, Gothenburg, Box 460 S-405 30, Sweden

^e CAS Center for Excellence in Tibetan Plateau Earth Sciences, Beijing, 100101, China

*Email: junguo.liu@gmail.com; liujg@sustc.edu.cn

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1029/2018JD029116

Abstract:

Hydro-climatic data are of importance to understand the water cycle and therefore for water resource assessment. Such data are of paramount importance for the Tibetan Plateau (TP) which is the source region of several large rivers in Asia. The Global Land Data Assimilation System (GLDAS) 2.0 and 2.1 provide abundant fine resolution hydro-climatic data. However, evaluations on their applicability have not been carried out for the TP. This study aims to evaluate and improve their applicability in basin scale hydrological applications in the TP. Gauge-based data, a hydrological model including biosphere, and seven state-of-the-art global precipitation products are utilized to carry out the study in four large basins in the TP. We find GLDAS2.1 shows significant warming trends from 2001 to 2010, whereas GLDAS2.0 shows cooling trends, although only significant in the Upper Yellow River basin. The contrasting trends imply caution should be taken when using them to analyze climate change impacts. On a monthly scale, GLDAS2.1 precipitation on average is closer to the gauge-based data than GLDAS2.0, but both of them have high uncertainty. Therefore, further quality improvements in precipitation are of importance. We also find CMORPH-BLD has better performance than other products in terms of Nash-Sutcliffe Efficiency (NSE), Relative Bias (RB) and root mean square error. Combining CMORPH-BLD with GLDAS2.0 forcing data generates more realistic runoff simulation than GLDAS2.1, with NSE and RB being 0.85 and 16% on average. The results provide unique insights into the studied data, and are beneficial for water resource assessment in the TP.

Keywords: GLDAS; Hydrology; Meteorology; Precipitation; Tibetan Plateau; WEB-DHM

1 Introduction

Hydro-climatic data are of importance to understand the water cycle and therefore for water research [Qi et al., 2015; McColl et al., 2017; Wang et al., 2017a]. Nevertheless, in situ observations are rather sparse in mountainous regions, rural areas, and developing countries; e.g., some regions in the Tibetan Plateau (TP).

The TP is the source region of several major Asian rivers, including the Yangtze River, Yarlung Tsangpo - Brahmaputra River, Lancang - Mekong River, Yellow River, etc. Therefore, knowledge about hydro-climatic change in the TP is very important. Most in situ stations are in the southern and eastern parts of the TP, and the stations are sparse in its northern and western parts [Gao and Liu, 2013; Yang et al., 2014; Su et al., 2016]. Therefore, to facilitate hydro-climatic change studies in the TP, other data sets should be implemented to compensate for the gauge data deficits, for example, data sets of the Global Land Data Assimilation System (GLDAS) [Rodell et al., 2004].

Understanding uncertainty and investigating possible improvements are two key issues in GLDAS data applications. Many studies have evaluated GLDAS 1.0 and uncertainty correction approaches have been developed [Kato et al., 2007; Zaitchik et al., 2010; Wang et al., 2011; Chen et al., 2013; Huang et al., 2013; Zhou et al., 2013; Qi et al., 2015; Bai et al., 2016; Wang et al., 2016; Yang et al., 2017]. For example, Gottschalck et al. [2005] investigated combinations of GLDAS 1.0 forcing data with four precipitation products seeking an optimal combination. Meanwhile, Qi et al. [2015] studied the applicability of GLDAS 1.0 data in a coastal region and proposed correction approaches. For GLDAS 2.0, only a few studies have investigated its applicability (e.g., Wang et al. [2016]). Regarding

GLDAS 2.1 which is the latest version of GLDAS, no studies have been conducted to investigate its applicability in hydrological simulations in the TP. In addition, to the best of the authors' knowledge, no studies have been carried out to inter-compare GLDAS 2.0 and 2.1 and investigate approaches to improve their applicability for the TP.

It is necessary to appraise the applicability of GLDAS 2.0 and 2.1 forcing data because new forcing data sets have been utilized for them compared to GLDAS 1.0 [Matthew and Hiroko Kato, 2015; Matthew and Hiroko Kato, 2016]. In addition, GLDAS 2.0 has been utilized to study changes and variations in hydrological processes on both regional and global scales. For example, Cheng et al. [2015] investigated the variability and trends of soil moisture over Asia based on GLDAS 2.0 data; Cheng and Huang [2016] studied soil moisture trends on a global scale utilizing GLDAS 2.0 data; and, Chen et al. [2017] investigated total water storage changes in China based on GLDAS 2.0 data. Given the importance of these studies, it is essential to have an awareness of the data quality problems since this may help understand how reliable the results are.

The overall objective of this study is to evaluate and improve the applicability of GLDAS 2.0 and 2.1 forcing data in basin scale hydrological simulations in the TP. The GLDAS data studied are GLDAS_NOAH025_3H_2.0 product [Matthew and Hiroko Kato, 2015] and GLDAS_NOAH025_3H_2.1 product [Matthew and Hiroko Kato, 2016]. A Water and Energy Budget-based Distributed biosphere Hydrological Model (WEB-DHM) [Wang et al., 2009a; Wang et al., 2009b; Wang et al., 2009c] was implemented. WEB-DHM was calibrated and validated using observed runoff in four large rivers in the TP. Seven state-of-the-art global fine resolution precipitation products were evaluated and combined with GLDAS 2.0 and 2.1

forcing data to seek the best coalescence in hydrological simulations in the TP. This paper is unique in that, for the first time, it evaluates and inter-compares GLDAS 2.0 and 2.1 (termed GLDAS2.0 and GLDAS2.1 hereafter) forcing data on large river basin scales in the TP, and approaches are also provided to improve their applicability in hydrological simulations for the TP.

2 Study area, data materials, and criteria for evaluation

2.1 The studied river basins in the TP

Four large river basins (Figure 1) in the southern and eastern TP were utilized to carry out this study because most gauge stations in the TP are in these regions [Gao and Liu, 2013; Yang et al., 2014; Xu et al., 2016; Wang et al., 2017b]. The four basins cover a total area of 57,1471 km², which is almost one quarter of the total area of the TP. Details about the river basins studied are shown in Table 1.

< Figure 1 here please >

< Table 1 here please >

2.2 Data sets

Precipitation data from China Gauge-based Daily Precipitation Analysis (CGDPA) is utilized in this study. CGDPA is based on over 2400 daily-scale in situ rain gauges and generated by the National Meteorological Information Center of China with a resolution of $0.5^0 \times 0.5^0$ [Zhao and Zhu, 2015; Shen and Xiong, 2016]. The CGDPA data have been utilized in many studies showing good performance [Miao et al., 2016; Gao et al., 2017; Liu et al., 2018]. Other forcing data are from the China Meteorological Forcing Dataset (CMFD), including

downward shortwave radiation, downward longwave radiation, specific humidity, wind speed, air pressure and air temperature. CMFC was developed by the Institute of Tibetan Plateau Research, Chinese Academy of Sciences. CMFC is based on China Meteorological Administration station measurements [He and Yang, 2011], and has been implemented in many studies showing general good performance [Xue et al., 2013; Zhou et al., 2015; Yang et al., 2017].

Digital Elevation Model (DEM) data were obtained from NASA Shuttle Radar Topographic Mission with a resolution of $30\text{ m} \times 30\text{ m}$ [Rabus et al., 2003]. Land-use data were obtained from the USGS (<http://edc2.usgs.gov/glcc/glcc.php>), and slope data were calculated based on the DEM (Supplementary Information Figure S1). Soil data were obtained from the Harmonized World Soil Database, and Future Water's Global Maps of Soil Hydraulic Property product (<http://www.futurewater.eu/>). The one-km, eight-day Moderate Resolution Imaging Spectroradiometer (MODIS) global Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation (FPAR) products (MOD15A2 1-km 8-day) [Myneni et al., 1997] were utilized to represent vegetation dynamic.

Downward solar and longwave radiation data of GLDAS2.0 are based on the Princeton Global meteorological Forcing (PGF) dataset. PGF dataset are generated from the National Centers for Environmental Prediction–National Center for Atmospheric Research reanalysis (NCEP–NCAR) [Kalnay et al., 1996] and NASA Langley surface radiation budget data [Stackhouse et al., 2011; Cox et al., 2017]. GLDAS2.0 air temperature is produced using Climatic Research Unit (CRU) TS2.0 and NCEP–NCAR reanalysis. For specific humidity, GLDAS2.0 uses data from NCEP–NCAR reanalysis without bias-corrections because

observation-based data are not available on a global scale [Sheffield et al., 2006]. In 2000 and 2001, GLDAS2.1 utilizes Global Data Assimilation System (GDAS) downward solar and longwave radiation data, and, after 2002, GLDAS2.1 uses downward solar and longwave radiation data from the air force weather agency's agricultural meteorological modeling system. Air temperature and Specific Humidity (SH) data of GLDAS2.1 are from the National Oceanic and Atmospheric Administration (NOAA) atmospheric analysis datasets [Derber et al., 1991].

2.3 Global precipitation data

The studied precipitation products are shown in Table 2. The data sources of the precipitation products mainly include satellite-based data, reanalysis data and gauge-based data. There are a few approaches to estimate precipitation from satellite, such as microwave-based approach, infrared-based approach and precipitation radar. Microwave estimation could miss convective rainfall and typhoon rain because of its sparse time interval resolution; infrared estimation has a higher time interval resolution, but it cannot penetrate thick clouds; precipitation radar could offer better precipitation estimation than microwave- and infrared-based estimations, but its data have smaller spatial coverage than microwave- and infrared-based data. Reanalysis data could capture large scale precipitation events, but may have high uncertainty on small scales because of coarse resolutions of parameterization on small scales [Ebert et al., 2007; Kidd et al., 2013]. Global scale gauge-based data are limited by the number of in situ gauges and gauge data quality.

< Table 2 here please >

The Tropical Rainfall Measuring Mission (TRMM) is a joint mission between NASA and Japan Aerospace Exploration Agency [Huffman et al., 2007]. Three instruments, i.e., Visible Infrared Radiometer (VIR), TRMM Microwave Imager (TMI), and precipitation radar, are employed to estimate precipitation. TRMM3B42 V7 estimates rainfall based on VIR, TMI and precipitation radar, also utilizing Global Precipitation Climatology Center (GPCP) monthly data [Huffman et al., 2001] and the Climate Assessment and Monitoring System monthly gauge data to improve its accuracy. The GLDAS2.0 precipitation data originate from disaggregated GPCP data [Huffman et al., 2001] based on the TRMM3B42 real time product (which used microwave- and infrared-based data). In addition, CRU TS2.0 and National Center for Atmospheric Research precipitation data are also used to produce GLDAS2.0 precipitation [Kalnay and Cai, 2003; Sheffield et al., 2006]. Different from GLDAS2.0, GLDAS2.1 utilizes GPCP and a disaggregation approach developed using GDAS precipitation estimation [Derber et al., 1991].

The Climate Prediction Center (CPC) MORPHing technique gauge-satellite blended precipitation (CMORPH-BLD) estimates precipitation based on passive microwaves aboard the Defense Meteorological Satellite Program, NOAA advanced microwave estimation, TMI, and GPCP data [Joyce et al., 2004]. Gauge adjusted Global Satellite Mapping of Precipitation (GSMaP-Gauge) utilizes a Kalman filter approach to estimate rainfall rates based on microwave-based data, and NOAA CPC gauge data and GPCP data are also utilized to improve its accuracy [Mega et al., 2014]. Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks - Climate Data Record (PERSIANN-CDR) estimates precipitation on the basis of GPCP, TMI and NOAA microwave-based precipitation estimation using an artificial neural network approach [Ashouri et al., 2015]. The Climate

Hazards Group InfraRed Precipitation with Station data (CHIRPS) is produced based on a few satellite based data, such as CMORPH and TRMM3B42, and is bias-corrected based on gauge data from Food and Agriculture Organization of the United Nations, the Global Historical Climate Network and CPC [Funk et al., 2015a; Funk et al., 2015b]. Multi-Source Weighted-Ensemble Precipitation (MSWEP) is the latest global precipitation product developed by Beck et al. [2017]. MSWEP merges several satellite based precipitation products, such as CMORPH, TRMM3B42RT, and GSMaP-MVK, and is bias-corrected according to GPCP and CPC gauge data. The Watch Forcing Data methodology applied to ERA-Interim (WFDEI-CRU) is a gauge and reanalysis data combined global product which utilizes the CRU data to correct bias of ERA-Interim data [Weedon et al., 2014]. More details about the precipitation products used are shown in Table 2.

2.4 Criteria for evaluations

Uncertainty evaluations are based on basin average values. Four criteria are utilized: correlation coefficient (CC), Nash-Sutcliffe Efficiency (NSE), Relative Bias (RB), and Root Mean Square Error (RMSE):

$$NSE = 1 - \frac{\sum_{i=1}^n (X_{si} - X_{oi})^2}{\sum_{i=1}^n (X_{si} - \overline{X_o})^2} \quad (1)$$

$$RB = \frac{\sum_{i=1}^n X_{si} - \sum_{i=1}^n X_{oi}}{\sum_{i=1}^n X_{oi}} \times 100\% \quad (2)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (X_{si} - X_{oi})^2}{n}} \quad (3)$$

where X_{si} represents GLDAS data or simulations using GLDAS data; X_{oi} represents observed

data or simulations using observations; i represents time; n represents total number of data points; \bar{X}_o represents average of the observed data or simulated data using observations.

3 Methodology

3.1 Overview

The approach used in this study includes four steps. First, WEB-DHM is validated using observed runoff in the four large river basins in the TP. Second, downward solar radiation, downward longwave radiation, specific humidity and air temperature of GLDAS2.0 and 2.1 are compared with gauge-based data. Third, precipitation data of GLDAS2.0 and 2.1 are evaluated against gauge-based data and compared with the global precipitation products to reveal their differences. Fourth, the best precipitation products are selected to replace original GLDAS2.0 and 2.1 precipitation to simulate runoff, and the performances of simulated runoff are evaluated based on observed runoff. This study is carried out on the basis of the data from March 2000 to December 2010 because the GSMaP-Gauge precipitation product is not available for January and February 2000 and there is no data for GLDAS2.0 after December 2010.

3.2 WEB-DHM model

WEB-DHM couples a geomorphology-based hydrological model [Yang, 1998] with a simple biosphere scheme (SiB2) [Sellers et al., 1986; Sellers et al., 1996a; Sellers et al., 1996b] to describe basin scale water, energy and CO₂ fluxes. Many evaluations have been carried out [Wang et al., 2010a; Wang et al., 2010b; Wang et al., 2012; Hu et al., 2014; Qi et al., 2015; Qi et al., 2016a], showing that WEB-DHM can simulate water and energy fluxes well on basin scales. The overall structure of WEB-DHM is shown in Supplementary Information Figure

S2. The input data of WEB-DHM includes air temperature, precipitation, downward solar radiation, air pressure, downward longwave radiation, humidity, and wind speed. The data used in this study are introduced in Section 2.2.

One year data were utilized to calibrate WEB-DHM parameters. It took almost two months to finish the calibration in a personal computer with four Inter Core i7-6700 3.40 GHz processors. In the Yarlung Tsangpo, Upper Lancang and Upper Yangtze Rivers, the data in 2001 were used; in the Upper Yellow River, the data in 2005 were used. The selections of the data for calibration considered the magnitudes of observed runoff: neither too low nor too high. The data in the other years were used to validate calibrated parameters. The calibration and validation were carried out on a monthly scale. Model parameter multipliers were calibrated, similar to the studies by Wang et al. [2011] and Qi et al. [2015]. The calibration has two steps. First, all the multiplier values were set to 1 which represents the default parameter values in WEB-DHM. Second, the multiplier values were optimized until an acceptable runoff simulation accuracy is obtained. In this study, the Dynamically Dimensioned Search (DDS) algorithm is used to optimize the multiplier values because DDS shows very good performance compared with other approaches [Tolson and Shoemaker, 2008; Tolson et al., 2009; Qi et al., 2016b]. $|NSE - 1| + |RB|$ is used as the objective function in the calibrations with an evaluation number of the objective function values up to 1000. The lower the objective function values, the better the multiplier values. The multiplier values with the best objective function value after finishing calibration processes are selected. Basin average values of calibrated model parameter values are shown in Table 3.

< Table 3 here please >

4 Results and discussion

4.1 WEB-DHM evaluations

Table 4 shows the monthly NSE and RB values in the calibration and validation periods in the four river basins. It can be seen that WEB-DHM performs well in the calibration period with NSE being up to 0.93. Similarly, the performances of WEB-DHM are also acceptable with NSE being up to 0.87 in the validation period.

< Table 4 here please >

Figure 2 shows the observed and simulated runoff using WEB-DHM from March 2000 to December 2010 in the four river basins on a monthly scale. In the Yarlung Tsangpo River, the overall NSE and RB between 2000 and 2010 are 0.87 and -9%, respectively, indicating the accuracy of the model simulation is acceptable. Similarly, WEB-DHM performs well in other basins with overall NSE being higher than 0.80 and absolute RB being less than 8%. On a multi-year mean monthly scale (Figure 3), the overall NSE values are all higher than 0.88 in the four basins. These results indicate that WEB-DHM can simulate hydrological processes well in the four river basins.

< Figure 2 here please >

< Figure 3 here please >

4.2 Evaluations of downward solar radiation, downward longwave radiation, specific humidity, and air temperature

4.2.1 Downward solar radiation

Figure 4 shows the comparison of downward solar radiation (RSW) among observation based data (CMFD), GLDAS2.0 and GLDAS2.1. In the Yarlung Tsangpo River, the seasonal variations of both GLDAS2.0 and GLDAS2.1 agree well with CMFD with CC being above 0.95 and NSE being above 0.79. However, the estimations of GLDAS2.0 and GLDAS2.1 are smaller than CMFD with RB being -7% and -1%, respectively. In addition, both GLDAS2.0 and GLDAS2.1 have smaller peak estimations than CMFD in most of the summer periods (with the exceptions of 2007 and 2008). In 2007, the peak values of GLDAS2.0 and CMFD in summer agree well, and GLDAS2.1 overestimates a little, whereas the peaks values of GLDAS2.1 and CMFD agree well in summer 2008. At the end of 2007, both GLDAS2.0 and GLDAS2.1 have higher estimations than CMFD. Overall, GLDAS2.1 outperforms GLDAS2.0 in downward solar radiation estimation in terms of NSE, RB, and RMSE.

< Figure 4 here please >

Similarly, the results in other three basins also show GLDAS2.0 and GLDAS2.1 have lower peak values than CMFD in most of the summer periods with exceptions in a few years; e.g., in 2008 and 2009 in the Upper Lancang River and in 2009 in the Upper Yangtze River. The differences between GLDAS2.0 and 2.1 may be due to their different data sources used (as explained in the section 2.2). In addition, similar to the results in the Yarlung Tsangpo River, GLDAS2.1 also outperformed GLDAS2.0 in terms of NSE, RB and RMSE in the Upper Lancang, Upper Yangtze and Upper Yellow Rivers.

4.2.2 Downward longwave radiation

Figure 5 shows the comparison of downward long wave radiation (RLW) among observation based data (CMFD), GLDAS2.0 and GLDAS2.1. In the Yarlung Tsangpo River, both GLDAS2.0 and GLDAS2.1 have smaller estimations than CMFD in most of the winter periods, and slightly overestimate in most of the summer periods. Similarly, in the Upper Lancang River, both GLDAS2.0 and GLDAS2.1 have smaller estimations than CMFD in winter. Comparatively, in the Upper Yangtze River, the three datasets agree relatively well. In the Upper Yellow River, GLDAS2.0 and 2.1 overestimate a little with RB being 6% and 4%, respectively. However, both GLDAS2.0 and 2.1 replicate the seasonal variations well with NSE being above 0.88 in the Upper Yellow River. Similar to the results in Figure 4, the differences between GLDAS2.0 and 2.1 may also result from their different data sources (as explained in the section 2.2). Overall, GLDAS2.0 and 2.1 can reasonably represent the seasonal variations of downward longwave radiation in the four river basins.

< Figure 5 here please >

4.2.3 Specific humidity

Figure 6 shows the comparison of SH among observation-based data (CMFD), GLDAS2.0, and GLDAS2.1. In the Yarlung Tsangpo River, GLDAS2.0 largely overestimates SH (RB equals 46%), especially in the summer, and GLDAS2.1 agrees relatively well with CMFD (NSE is 0.93 and RB is 1%). Similarly, GLDAS2.0 also overestimates SH in the other three regions with RB being up to 36% in the Upper Yellow River. GLDAS2.1 agrees relatively well with CMFD in the Upper Lancang and Upper Yellow Rivers, whereas GLDAS2.1 slightly underestimates in the Upper Yangtze River with RB being -14%. The differences

between GLDAS2.0 and 2.1 may result from their different data sources (as explained in the section 2.2). Overall, GLDAS2.0 has higher SH estimation than GLDAS2.1, especially in the summer.

< Figure 6 here please >

4.2.4 Air temperature

Figure 7 compares air temperature among observation based data (CMFD), GLDAS2.0, and GLDAS2.1. The data in January and February 2000 are not included, and therefore the data in 2000 is not considered when calculating annual averages.

< Figure 7 here please >

In the Yarlung Tsangpo River, GLDAS2.0 agrees well with CMFD, whereas GLDAS2.1 largely underestimates air temperature before 2005 and slightly underestimates air temperature in winter after 2006 (Figure 7a). Similarly, in the Upper Lancang River, GLDAS2.1 has lower estimations than GLDAS2.0 and CFMD in winter before 2005. In addition, it can be seen that GLDAS2.1 has higher estimations than GLDAS2.0 in summer after 2006 in the Upper Lancang River (Figure 7c). Comparatively, Figures 7e and f show the three datasets agree relatively well in the Upper Yangtze and Upper Yellow Rivers (NSE values are above 0.93) though with some exceptions, e.g., in the winter of 2000 and 2001 and the summer of 2006 and 2010. Over the four river basins, the average RMSE of GLDAS2.0 and 2.1 are 1.63 and 3.03, and the average NSE are 0.94 and 0.76. Therefore, GLDAS2.1 has higher overall uncertainty than GLDAS2.0 in the TP.

On a yearly scale, GLDAS2.1 shows a significant warming trend in the four river basins on the basis of a Mann-kendall test at a significance level of 0.05. CMFD also shows a warming trend but it is not significant in the Upper Yangtze and Upper Yellow Rivers. Different from GLDAS2.1 and CMFD, GLDAS2.0 shows a cooling trend in the Yarlung Tsangpo, Upper Yangtze and Upper Yellow Rivers (significant in the Upper Yellow River) and a warming trend in the Upper Lancang River. These differences indicate that caution should be taken when analyzing climate change impacts from 2001 to 2010 in the TP using the different datasets. The differences between GLDAS2.0 and 2.1 may be due to their different data sources (as explained in the section 2.2). The PGF air temperature used in GLDAS2.0 is based on NCEP reanalysis data [Kalnay et al., 1996]. NCEP calculates near surface temperature based on modeled atmosphere variables, and therefore is influenced by model parameterization [Sheffield et al., 2006]. Several studies have suggested that the uncertainty in the parameterization of NCEP could lead to large uncertainty in the trend of air temperature (e.g., Kalnay and Cai [2003]; Simmons et al. [2004]).

4.3 Precipitation evaluations by comparing measuring data and other global precipitation products

4.3.1 Daily scale

Figure 8 shows the evaluation results of GLDAS2.0 and 2.1 precipitation data and comparisons with global precipitation products based on the four criteria on a daily scale. Each row represents the results in one river basin, and each column represents the results of one evaluation criterion. The larger the circles, the better the precipitation products.

< Figure 8 here please >

In terms of CC, GLDAS2.0 has higher CC values than GLDAS2.1 in all four river basins. Regarding NSE, GLDAS2.0 is also better than GLDAS2.1, though the NSE values are not high in the four river basins (the highest value is 0.37 in the Upper Yellow River). With respect to RB, GLDAS2.0 has a better RB value (51%) than GLDAS2.1 (72%) in the Yarlung Tsangpo River, though both of their RB values show large uncertainty. RB values of GLDAS2.1 are better than GLDAS2.0 in the Upper Lancang, Upper Yangtze, and Upper Yellow Rivers. Regarding RMSE, GLDAS2.0 has better RMSE values than GLDAS2.1 in all four river basins. Concerning the average values of the four criteria, GLDAS2.0 is better than GLDAS2.1 in terms of CC, NSE, and RMSE. With regards to absolute values of RB ($|RB|$), GLDAS2.1 is better than GLDAS2.0 though their average $|RB|$ values are high (42% and 29% for GLDAS2.0 and 2.1 respectively). The high bias in rainfall estimation in GLDAS2.0 may result from uncertainties in PGF data and TRMM data because TRMM3B42RT data were used to generate GLDAS2.0 precipitation. For GLDAS2.1, the uncertainties may come from GDAS precipitation data which are used to develop GLDAS2.1 precipitation.

For other precipitation products, CMORPH-BLD is the best in all the four river basins in terms of CC, NSE, and RMSE. Regarding RB, CMORPH-BLD has the best RB value (-16%) in the Yarlung Tsangpo River; in the Upper Lancang River, the RB values of CMORPH-BLD and CHIRPS2 are the same and lower than others. Different from the results in the Yarlung Tsangpo and Upper Lancang Rivers, CHIRPS2 has the best RB values in the Upper Yangtze and Upper Yellow Rivers (-15% and 1% respectively), and CMORPH-BLD has the second best RB value (4%) in the Upper Yellow River. The average values of CC, NSE, $|RB|$, and

RMSE over the four river basins are shown in the bottom row of Figure 8. It can be seen that CMORPH-BLD is the best in terms of the average values of CC, NSE, |RB|, and RMSE. Therefore, CMORPH-BLD is better than other precipitation products overall on the daily scale.

4.3.2 Monthly scale

Figure 9 shows the evaluation results of GLDAS2.0 and 2.1 precipitation data and comparison with global precipitation products on a monthly scale. The larger the circles, the better the precipitation products.

< Figure 9 here please >

In terms of CC, GLDAS2.1 is better than GLDAS2.0 in all four river basins, whereas GLDAS2.0 is better than GLDAS2.1 in the Yarlung Tsangpo River with respect to NSE and RMSE. In the Upper Lancang, Upper Yangtze, and Upper Yellow Rivers, GLDAS2.1 is better than GLDAS2.0 regarding NSE and RMSE. With regard to the average values of CC, NSE, and RMSE, GLDAS2.1 is better than GLDAS2.0 overall.

For other precipitation products, CHIRPS2 has higher CC values than the others in the Upper Lancang and Upper Yangtze Rivers. In the Yarlung Tsangpo River, TRMM3B42 and GLDAS2.1 have the same CC values and are better than the others, while in the Upper Yellow River, TRMM3B42 has better CC values than the others. Regarding NSE and RMSE, CMORPH-BLD is better than the others in the Yarlung Tsangpo and Upper Yellow Rivers, whereas CHIRPS2 is better than the others in the Upper Lancang and Upper Yangtze Rivers.

In the Upper Yellow River, TRMM3B42 and CHIRPS2 have the same NSE and RMSE values as CMORPH-BLD. Concerning the average criteria values, TRMM3B42 is better than the others in terms of CC; regarding NSE, CMORPH-BLD is the best; and CHIRPS2 is better than the others with respect to RMSE. These results appear to indicate that it is difficult to identify the best precipitation products on a monthly scale.

4.3.3 Multi-year average scale

Figure 10 shows the evaluation results of global precipitation products on a multi-year average monthly scale. The bigger the circles, the better the precipitation products.

< Figure 10 here please >

Regarding CC, GLDAS2.0 is the best in the Yarlung Tsangpo, Upper Lancang, and Upper Yellow Rivers. Meanwhile, in the Upper Yangtze River, CHIRPS2 is the best. With respect to NSE and RMSE, CMORPH-BLD is the best in the Yarlung Tsangpo and Upper Yellow Rivers, whereas CHIRPS2 is the best in the Upper Lancang and Upper Yangtze Rivers. In the Upper Yellow River, the NSE value of CHIRPS2 is the same as CMORPH-BLD. The average values of CC, NSE, and RMSE over the four river basins are shown in the bottom row of Figure 10. Regarding CC, all the precipitation products have CC values above 0.98 and TRMM3B42 has the highest CC value. With respect to NSE and RMSE, CMORPH-BLD is the best among all the studied precipitation products. Compared with the results on the daily scale, it can be seen that CMORPH-BLD is better than the other precipitation products in terms of the averages of |RB|, NSE, and RMSE on both daily and multi-year average monthly scales.

Figure 11 compares the precipitation data on a multi-year average monthly scale in time series plots. In the Yarlung Tsangpo River basin, NSWEP has the lowest rainfall intensity estimation in July and August among all the precipitation products investigated. Similarly, MSWEP also has the lowest rainfall intensity estimation from May to August in the Upper Yellow River. In the Upper Lancang and Upper Yangtze Rivers, MSWEP has lower rainfall estimation than the observation-based data, especially in June, July, and August. MSWEP utilizes GPCP and NOAA CPC gauge data to correct bias of satellite and reanalysis precipitation data, whereas a limited number of rain gauges used in CGDPA are included in GPCP and CPC, especially in western China [Chen et al., 2008; Shen and Xiong, 2016; Gao et al., 2017]. As a new global precipitation product developed in 2017, MSWEP has rarely been specifically evaluated in the TP. To the best of the authors' knowledge, this is the first time that it has been found that MSWEP underestimates precipitation in all the four large river basins, especially in summer.

< Figure 11 here please >

4.4 Hydrological evaluations in runoff simulations

Because the results on a multi-year average monthly scale show CMORPH-BLD can outperform the other products, this section will compare the performance of CMORPH-BLD with other precipitation products in terms of runoff simulation. Figure 12 shows evaluations of simulated runoff using GLDAS2.0 forcing data and global precipitation products on a multi-year average monthly scale. Here, the forcing data refer to downward solar radiation, downward longwave radiation, specific humidity and air temperature. The first four rows are

the results of the four evaluation criteria in the four river basins, and the last two rows present average criteria values over the four river basins.

< Figure 12 here please >

Regarding CC, MSWEP and CMORPH-BLD generate the highest and the second highest values in the Yarlung Tsangpo River, and CMORPH-BLD generates the highest CC values in the other three basins. With respect to NSE, RB, and RMSE, the best precipitation products vary as the river basins change. Therefore, the average values of the evaluation criteria over the four river basins are compared. With respect to the average value of CC, CMORPH-BLD generates the highest value (0.99). Similarly, CMORPH-BLD also generates the best average values in terms of NSE, |RB| and RMSE. Thus, CMORPH-BLD could be a good substitution for the precipitation data in GLDAS2.0 in the TP.

Figure 13 shows the evaluations of simulated runoff using GLDAS2.1 forcing data and global precipitation products on a multi-year average monthly scale. In the Yarlung Tsangpo River, CMORPH-BLD is the best in terms of all the four criteria, and PERSIANN-CDR is the best in the Upper Yellow River according to the four criteria. In terms of NSE, RB and RMSE, CHIRPS2 is the best in the Upper Lancang River, and PERSIANN-CDR is the best in the Upper Yangtze River. In terms of the average criterion values, it can be seen that CMORPH-BLD generates the best average criterion values, which is similar to the results in Figure 12. Compared to the average criterion values in Figure 12, all the average criterion values of CMORPH-BLD in Figure 13 are worse. For example, the average NSE is 0.85 for CMORPH-BLD in Figure 12, and 0.79 in Figure 13; the average |RB| is 16% for

CMORPH-BLD in Figure 12, and 22% in Figure 13. These results indicate the coalescence of GLDAS2.0 forcing data and CMORPH-BLD could be better than the coalescence of GLDAS2.1 forcing data and CMORPH-BLD in runoff simulation in the TP.

< Figure 13 here please >

4.5 Discussion

Gao and Liu [2013] evaluated four precipitation products (i.e., TRMM3B42 V6, TRMM3B42RT V6, CMORPH without gauge corrections and PERSIANN) in the TP based on in situ gauge data. Wang et al. [2015] evaluated five global precipitation products (i.e. TRMM3B42 V7, TRMM3B42RT V7, CMORPH gauge adjusted product, CMORPH real time product, and CMORPH-BLD) in two sub-basins (covering a total area of 32,652 km²) of the Yarlung Tsangpo River. Tong et al. [2014] assessed TRMM3B42 V7, TRMM3B42RT V7, CMORPH without gauge corrections, and PERSIANN in two river basins, i.e., the Upper Yellow River basin and the Upper Yangtze River basin. Compared to these prior studies, our study was carried out in four large river basins with a total area of 57,1471 km² (almost one quarter of the total area of the TP) and examined more precipitation products. Wang et al. [2015] concluded that CMORPH-BLD is better than TRMM3B42, TRMM3B42RT, and CMORPH real time product in runoff simulation. This result seems to support our results, but their study area is much smaller. Similarly, Su et al. [2017] also indicated CMORPH-BLD has better performance than CMORPH real time product, PERSIANN-CDR, and TRMM3B42 V7. Nevertheless, their study was carried out only for the Upper Yellow River basin. Compared to the research by Qi et al. [2016a], our study evaluated the new version datasets of GLDAS and more global precipitation products in four larger river basins.

Therefore, the results of our study are much more comprehensive and informative than the prior studies.

The different performance of GLDAS2.0 and 2.1 in runoff simulations may be due to their different uncertainty in the forcing data. Although GLDAS2.1 outperforms GLDAS2.0 in downward solar radiation estimation in terms of NSE, RB and RMSE, the air temperature estimation of GLDAS2.1 has a larger uncertainty than GLDAS2.0. This difference may lead to the higher uncertainty in simulated runoff using GLDAS2.1 forcing data than GLDAS2.0. According to the best of the authors' knowledge, this is the first time that runoff simulations of different combinations of the global precipitation products and GLDAS2.0 and 2.1 forcing data have been investigated in the TP. We find that CMORPH-BLD could be the best substitution for GLDAS2.0 and 2.1 precipitation data in hydrological simulation for the TP.

It should be noted that the performance of global precipitation products may change as study regions and application purposes vary [Qi et al., 2016a; Beck et al., 2017]. In this study, we investigated the performance of the global precipitation products in the TP for runoff simulations. The identified best precipitation product may change when study region and simulated hydrological variables are different. However, our results do show CMORPH-BLD product has better performances on average than other products in runoff simulations on a multi-year average monthly scale. This finding could be beneficial for hydrological simulations in ungauged regions in the TP.

In this study, the initial land surface states on March 1 2000, such as soil moisture in different soil layers, are assumed to be the default values in WEB-DHM in the first model run. After

the first run, the simulated land surface states on March 1 2010 is used as the new initial land surface states on March 1 2000 for the model calibrations. After the model calibrations, the simulated land surface states on March 1 2010 using calibrated parameters are used as new initial land surface states on March 1 2000 in other simulations. The results in Figures 2 and 3 indicate the simulated runoff are satisfactory. Therefore, the approach used to spin the model up is acceptable.

In this study, the WEB-DHM model was calibrated using one year data and 1000 objective function evaluations because of the computational burden of the calibration. Nevertheless, the calibrated parameters show good performance in runoff simulations (as shown in Figures 2 and 3) and the calibration approach used is therefore acceptable. In addition, the DDS algorithm was utilized in the calibration. However, other appropriate optimization algorithms can also be utilized, even though DDS was implemented in this study. The simulated runoff replicates observation well. Therefore, using DDS is acceptable.

5 Conclusions

GLDAS2.0 and 2.1 data utilize new sources compared with GLDAS1.0 and benchmarking is important. However, the two new versions have been validated at very limited locations, and inter-comparisons have not been carried out in the TP. This study aimed to evaluate, inter-compare and improve their applicability in basin scale hydrological studies in the TP. Gauge based meteorological forcing data, seven global fine resolution precipitation products, and a water and energy budget-based distributed biosphere hydrological model were utilized to conduct the study. The results provide unique insights into the similarities and differences of the studied data, and promising coalescence of GLDAS2.0 and 2.1 forcing data and global

precipitation products in hydrological simulation in the TP have also been identified. The results could be beneficial for a broad range of applications in the TP. The following conclusions are presented on the basis of this study.

First, GLDAS2.1 shows significant warming trends in all the four river basins from 2001 to 2010, whereas GLDAS2.0 shows cooling trends which are only significant in the Upper Yellow River basin. Therefore, caution should be taken when using them to analyze and interpret climate change impacts in the TP.

Second, on a monthly scale, GLDAS2.1 precipitation is better than GLDAS2.0 data with regards to average values of CC, NSE, |RB| and RMSE over the four large river basins. Nevertheless, both of them have high uncertainty: averages of |RB| are 42% and 29% for GLDAS2.0 and 2.1. Therefore, improvements in precipitation data are of importance for their practical applications in the TP.

Third, CMORPH-BLD offers the best precipitation on average in terms of CC, NSE, RB, and RMSE on a daily scale and in terms of NSE and RMSE on a multi-year average monthly scale. Combining CMORPH-BLD with GLDAS2.0 forcing data generates better runoff simulations than combining CMORPH-BLD with GLDAS2.1 forcing data. Therefore, it is promising to utilize CMORPH-BLD precipitation and GLDAS2.0 forcing data in hydrological simulation in data sparse regions in the TP.

Acknowledgements:

This study was supported by the National Science Fund for Distinguished Youth Scholars

(41625001), Strategic Priority Research Program of Chinese Academy of Sciences (Grant No. XDA20060402), National Key Research and Development Program (2017YFA0603704), the Young Scientists Fund of the National Natural Science Foundation of China (51809136), China Postdoctoral Science Foundation (2017M622516), and the National Natural Science Foundation of China (41571022). Additional support was provided by Guangdong Provincial Key Laboratory of Soil and Groundwater Pollution Control (2017B030301012), and State Environmental Protection Key Laboratory of Integrated Surface Water-Groundwater Pollution Control. GLDAS 2.0 and 2.1 data were downloaded from <https://hydro1.gesdisc.eosdis.nasa.gov/data/GLDAS/>. MODIS data were downloaded from USGS land processes distributed active archive center (https://lpdaac.usgs.gov/products/modis_products_table/mod15a2). CMORPH-BLD V1.0 was downloaded from ftp://ftp.cpc.ncep.noaa.gov/precip/CMORPH_V1.0/. GSMaP-Gauge was downloaded from ftp://hokusai.eorc.jaxa.jp/standard_gauge/. PERSIANN-CDR was downloaded from <https://www.ncei.noaa.gov/data/>. TRMM3B42 was downloaded from <http://mirador.gsfc.nasa.gov/cgi-bin/mirador/>. WFDEI-CRU was downloaded from <ftp.iiasa.ac.at>. MSWEP V2.01 was downloaded from <https://data.princetonclimate.com/opensap>. CHIRPS V2.0 was downloaded from <ftp://ftp.chg.ucsb.edu/>. The forcing dataset used in this study (i.e., CMFD) was developed by the Data Assimilation and Modeling Center for Tibetan Multi-spheres, the Institute of Tibetan Plateau Research, Chinese Academy of Sciences. The observed discharge data used in this study were collected from hydrology bureaus. The authors are deeply indebted to editors and three anonymous reviewers for their valuable time and constructive suggestions that greatly improved the quality of this paper.

References:

- Ashouri, H., Hsu, K.-L., Sorooshian, S., Braithwaite, D.K., Knapp, K.R., Cecil, D.L., Nelson, B.R., Prat, O.P., (2015). PERSIANN-CDR: Daily Precipitation Climate Data Record from Multisatellite Observations for Hydrological and Climate Studies. *Bulletin of the American Meteorological Society*, 96(1): 69-83. DOI:10.1175/BAMS-D-13-00068.1
- Bai, P., Liu, X., Yang, T., Liang, K., Liu, C., (2016). Evaluation of streamflow simulation results of land surface models in GLDAS on the Tibetan plateau. *Journal of Geophysical Research: Atmospheres*, 121(20). DOI:10.1002/2016JD025501
- Beck, H.E., van Dijk, A., Levizzani, V., Schellekens, J., Miralles, D.G., Martens, B., de Roo, A., (2017). MSWEP: 3-hourly 0.25° global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data. *Hydrology and Earth System Sciences*, 21(1): 589-615. DOI:10.5194/hess-21-589-2017
- Chen, M., Shi, W., Xie, P., Silva, V.B.S., Kousky, V.E., Higgins, W.R., Janowiak, J.E., (2008). Assessing objective techniques for gauge - based analyses of global daily precipitation. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 113(D4). DOI:10.1029/2007JD009132
- Chen, Y., Yang, K., Qin, J., Zhao, L., Tang, W., Han, M., (2013). Evaluation of AMSR - E retrievals and GLDAS simulations against observations of a soil moisture network on the central Tibetan Plateau. *Journal of Geophysical Research: Atmospheres*, 118(10): 4466-4475. DOI:10.1002/jgrd.50301
- Chen, Z., Jiang, W., Wu, J., Chen, K., Deng, Y., Jia, K., Mo, X., (2017). Detection of the spatial patterns of water storage variation over China in recent 70 years. *Scientific Reports*, 7(1): 6423. DOI:10.1038/s41598-017-06558-5
- Cheng, S., Guan, X., Huang, J., Ji, F., Guo, R., (2015). Long - term trend and variability of

-
- soil moisture over East Asia. *Journal of Geophysical Research: Atmospheres*, 120(17): 8658-8670. DOI:10.1002/2015JD023206
- Cheng, S., Huang, J., (2016). Enhanced soil moisture drying in transitional regions under a warming climate. *Journal of Geophysical Research: Atmospheres*, 121(6): 2542-2555. DOI:10.1002/2015JD024559
- Cox, S.J., Stackhouse, P.W., Gupta, S.K., Mikovitz, C.J., Zhang, T., (2017). NASA/GEWEX shortwave surface radiation budget: Integrated data product with reprocessed radiance, cloud, and meteorology inputs, and new surface albedo treatment. *AIP Conference Proceedings 1810*: 090001. DOI:10.1063/1.4975541
- Derber, J.C., Parrish, D.F., Lord, S.J., (1991). The New Global Operational Analysis System at the National Meteorological Center. *Weather and Forecasting*, 6(4): 538-547. DOI:10.1175/1520-0434(1991)006<0538:TNGOAS>2.0.CO;2
- Ebert, E.E., Janowiak, J.E., Kidd, C., (2007). Comparison of near-real-time precipitation estimates from satellite observations and numerical models. *Bulletin of the American Meteorological Society*, 88(1): 47-64.
- Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., Michaelsen, J., (2015a). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data*, 2. DOI:10.1038/sdata.2015.66
- Funk, C., Verdin, A., Michaelsen, J., Peterson, P., Pedreros, D., Husak, G., (2015b). A global satellite-assisted precipitation climatology. *Earth System Science Data*, 7(2): 275-287. DOI:10.5194/essd-7-275-2015
- Gao, Y.C., Liu, M.F., (2013). Evaluation of high-resolution satellite precipitation products using rain gauge observations over the Tibetan Plateau. *Hydrology and Earth System*

Sciences, 17(2): 837-849. DOI:10.5194/hess-17-837-2013

Gao, Z., Long, D., Tang, G., Zeng, C., Huang, J., Hong, Y., (2017). Assessing the potential of satellite-based precipitation estimates for flood frequency analysis in ungauged or poorly gauged tributaries of China's Yangtze River basin. *Journal of Hydrology*, 550: 478-496.

Gottschalck, J., Meng, J., Rodell, M., Houser, P., (2005). Analysis of Multiple Precipitation Products and Preliminary Assessment of Their Impact on Global Land Data Assimilation System Land Surface States. *Journal of Hydrometeorology*, 6(5): 573-598. DOI:10.1175/JHM437.1

He, J., Yang, K., (2011). China Meteorological Forcing Dataset. *Cold and Arid Regions Science Data Center at Lanzhou*. DOI:10.3972/westdc.002.2014.db

Hu, Z., Wang, L., Wang, Z., Hong, Y., Zheng, H., (2014). Quantitative assessment of climate and human impacts on surface water resources in a typical semi-arid watershed in the middle reaches of the Yellow River from 1985 to 2006. *International Journal of Climatology*: 97-113. DOI:10.1002/joc.3965

Huang, Y., Salama, M.S., Krol, M.S., van der Velde, R., Hoekstra, A.Y., Zhou, Y., Su, Z., (2013). Analysis of long-term terrestrial water storage variations in the Yangtze River basin. *Hydrology and Earth System Sciences*, 17(5): 1985-2000. DOI:10.5194/hess-17-1985-2013

Huffman, G.J., Adler, R.F., Morrissey, M.M., Bolvin, D.T., Curtis, S., Joyce, R., McGavock, B., Susskind, J., (2001). Global Precipitation at One-Degree Daily Resolution from Multisatellite Observations. *Journal of Hydrometeorology*, 2(1): 36-50. DOI:10.1175/1525-7541(2001)002<0036:GPAODD>2.0.CO;2

Huffman, G.J., Bolvin, D.T., Nelkin, E.J., Wolff, D.B., Adler, R.F., Gu, G., Hong, Y.,

-
- Bowman, K.P., Stocker, E.F., (2007). The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. *Journal of Hydrometeorology*, 8(1): 38-55. DOI:10.1175/jhm560.1
- Joyce, R.J., Janowiak, J.E., Arkin, P.A., Xie, P., (2004). CMORPH: A Method that Produces Global Precipitation Estimates from Passive Microwave and Infrared Data at High Spatial and Temporal Resolution. *Journal of Hydrometeorology*, 5(3): 487-503. DOI:10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, R., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K.C., Ropelewski, C., Wang, J., Jenne, R., Joseph, D., (1996). The NCEP/NCAR 40-Year Reanalysis Project. *Bulletin of the American Meteorological Society*, 77(3): 437-471. DOI:10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2
- Kalnay, E., Cai, M., (2003). Impact of urbanization and land-use change on climate. *Nature*, 423(6939): 528-531. DOI:10.1038/nature01675
- Kato, H., Rodell, M., Beyrich, F., Cleugh, H., Gorsel, E.v., Liu, H., Meyers, T.P., (2007). Sensitivity of Land Surface Simulations to Model Physics, Land Characteristics, and Forcings, at Four CEOP Sites. *Journal of the Meteorological Society of Japan*, 85A: 187-204. DOI:10.2151/jmsj.85A.187
- Kidd, C., Dawkins, E., of Hydrometeorology, H.-G., (2013). Comparison of precipitation derived from the ECMWF operational forecast model and satellite precipitation datasets. *Journal of Hydrometeorology*, 14: 1463–1482. DOI:10.1175/JHM-D-12-0182.1
- Liu, T., Liang, Z., Chen, Y., Lei, X., Li, B., (2018). Long-duration PMP and PMF estimation

-
- with SWAT model for the sparsely gauged Upper Nujiang River Basin. *Natural Hazards*, 90(2): 735-755. DOI:10.1007/s11069-017-3068-z
- Matthew, R., Hiroko Kato, B., 2015. GLDAS Noah Land Surface Model L4 3 hourly 0.25 x 0.25 degree V2.0. NASA/GSFC/HSL, Goddard Earth Sciences Data and Information Services Center (GES DISC), Greenbelt, Maryland, USA. DOI:10.5067/342OHQM9AK6Q
- Matthew, R., Hiroko Kato, B., 2016. GLDAS Noah Land Surface Model L4 3 hourly 0.25 x 0.25 degree V2.1. NASA/GSFC/HSL, Goddard Earth Sciences Data and Information Services Center (GES DISC), Greenbelt, Maryland, USA. DOI:10.5067/E7TYRXPJKWOQ
- McColl, K.A., Alemohammad, S., Akbar, R., Konings, A.G., Yueh, S., Entekhabi, D., (2017). The global distribution and dynamics of surface soil moisture. *Nature Geoscience*, 10(2): 100-104. DOI:10.1038/ngeo2868
- Mega, T., Ushio, T., Kubota, T., Kachi, M., Aonashi, K., Shige, S., (2014). Gauge Adjusted Global Satellite Mapping of Precipitation (GSMaP_Gauge). *2014 XXXIth URSI General Assembly and Scientific Symposium (URSI GASS)*: 1-4. DOI:10.1109/URSIGASS.2014.6929683
- Miao, C., Sun, Q., Duan, Q., Wang, Y., (2016). Joint analysis of changes in temperature and precipitation on the Loess Plateau during the period 1961–2011. *Climate Dynamics*, 47(9-10): 3221-3234. DOI:10.1007/s00382-016-3022-x
- Myneni, R.B., Ramakrishna, R., Nemani, R., Running, S.W., (1997). Estimation of global leaf area index and absorbed par using radiative transfer models. *Geoscience and Remote Sensing, IEEE Transactions on*, 35(6): 1380-1393. DOI:10.1109/36.649788
- Qi, W., Zhang, C., Fu, G., Zhou, H., (2015). Global Land Data Assimilation System data

assessment using a distributed biosphere hydrological model. *Journal of Hydrology*, 528: 652-667. DOI:10.1016/j.jhydrol.2015.07.011

Qi, W., Zhang, C., Fu, G., Sweetapple, C., Zhou, H., (2016a). Evaluation of global fine-resolution precipitation products and their uncertainty quantification in ensemble discharge simulations. *Hydrology and Earth System Sciences*, 20(2): 903-920. DOI:10.5194/hess-20-903-2016

Qi, W., Zhang, C., Fu, G., Zhou, H., (2016b). Quantifying dynamic sensitivity of optimization algorithm parameters to improve hydrological model calibration. *Journal of Hydrology*, 533: 213-223. DOI:10.1016/j.jhydrol.2015.11.052

Rabus, B., Eineder, M., Roth, A., Bamler, R., (2003). The shuttle radar topography mission—a new class of digital elevation models acquired by spaceborne radar. *ISPRS Journal of Photogrammetry and Remote Sensing*, 57(4): 241-262. DOI:10.1016/s0924-2716(02)00124-7

Rodell, M., Houser, P.R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J.K., Walker, J.P., Lohmann, D., Toll, D., (2004). The Global Land Data Assimilation System. *Bulletin of the American Meteorological Society*, 85(3): 381-394. DOI:10.1175/bams-85-3-381

Sellers, P.J., Mintz, Y., Sud, Y.C., Dalcher, A., (1986). A Simple Biosphere Model (SIB) for Use within General Circulation Models. *Journal of the Atmospheric Sciences*, 43(6): 505-531. DOI:10.1175/1520-0469(1986)043<0505:ASBMFU>2.0.CO;2

Sellers, P.J., Randall, D.A., Collatz, G.J., Berry, J.A., Field, C.B., Dazlich, D.A., Zhang, C., Collelo, G.D., Bounoua, L., (1996a). A Revised Land Surface Parameterization (SiB2) for Atmospheric GCMS. Part I: Model Formulation. *Journal of Climate*, 9(4): 676-705. DOI:10.1175/1520-0442(1996)009<0676:ARLSPF>2.0.CO;2

-
- Sellers, P.J., Tucker, C.J., Collatz, G.J., Los, S.O., Justice, C.O., Dazlich, D.A., Randall, D.A., (1996b). A Revised Land Surface Parameterization (SiB2) for Atmospheric GCMS. Part II: The Generation of Global Fields of Terrestrial Biophysical Parameters from Satellite Data. *Journal of Climate*, 9(4): 706-737. DOI:10.1175/1520-0442(1996)009<0706:ARLSPF>2.0.CO;2
- Sheffield, J., Goteti, G., Wood, E.F., (2006). Development of a 50-Year High-Resolution Global Dataset of Meteorological Forcings for Land Surface Modeling. *Journal of Climate*, 19(13): 3088-3111. DOI:10.1175/JCLI3790.1
- Shen, Y., Xiong, A., (2016). Validation and comparison of a new gauge - based precipitation analysis over mainland China. *International Journal of Climatology*, 36(1): 252-265. DOI:10.1002/joc.4341
- Simmons, A.J., Jones, P.D., da Bechtold, C.V., Beljaars, A.C.M., Kållberg, P.W., Saarinen, S., Uppala, S.M., Viterbo, P., Wedi, N., (2004). Comparison of trends and low - frequency variability in CRU, ERA - 40, and NCEP/NCAR analyses of surface air temperature. *Journal of Geophysical Research: Atmospheres (1984 - 2012)*, 109(D24). DOI:10.1029/2004JD005306
- Stackhouse, J., W., P., Gupta, S.K., Cox, S.J., Mikovitz, J.C., Zhang, T., Hinkelman, L.M., (2011). The NASA/GEWEX Surface Radiation Budget Release 3.0: 24.5-Year Dataset. *GEWEX News*, 21(1): 10-12.
- Su, F., Zhang, L., Ou, T., Chen, D., Yao, T., Tong, K., Qi, Y., (2016). Hydrological response to future climate changes for the major upstream river basins in the Tibetan Plateau. *Global and Planetary Change*, 136: 82-95. DOI:10.1016/j.gloplacha.2015.10.012
- Su, J., Lü, H., Wang, J., Sadeghi, A.M., Zhu, Y., (2017). Evaluating the Applicability of Four Latest Satellite–Gauge Combined Precipitation Estimates for Extreme Precipitation

and Streamflow Predictions over the Upper Yellow River Basins in China. *Remote Sensing*, 9(11): 1176. DOI:10.3390/rs9111176

Tolson, B.A., Shoemaker, C.A., (2008). Efficient prediction uncertainty approximation in the calibration of environmental simulation models. *Water Resources Research*, 44(4): W04411. DOI:10.1029/2007wr005869

Tolson, B.A., Asadzadeh, M., Maier, H.R., Zecchin, A., (2009). Hybrid discrete dynamically dimensioned search (HD-DDS) algorithm for water distribution system design optimization. *Water Resources Research*, 45(12): W12416. DOI:10.1029/2008WR007673

Tong, K., Su, F., Yang, D., Hao, Z., (2014). Evaluation of satellite precipitation retrievals and their potential utilities in hydrologic modeling over the Tibetan Plateau. *Journal of Hydrology*, 519: 423-437. DOI:10.1016/j.jhydrol.2014.07.044

Wang, F., Wang, L., Koike, T., Zhou, H., Yang, K., Wang, A., Li, W., (2011). Evaluation and application of a fine-resolution global data set in a semiarid mesoscale river basin with a distributed biosphere hydrological model. *Journal of Geophysical Research: Atmospheres*, 116(D21). DOI:10.1029/2011jd015990

Wang, F., Wang, L., Zhou, H., Saavedra Valeriano, O.C., Koike, T., Li, W., (2012). Ensemble hydrological prediction-based real-time optimization of a multiobjective reservoir during flood season in a semiarid basin with global numerical weather predictions. *Water Resources Research*, 48(7). DOI:10.1029/2011wr011366

Wang, L., Koike, T., Yang, D.W., Yang, K., (2009a). Improving the hydrology of the Simple Biosphere Model 2 and its evaluation within the framework of a distributed hydrological model. *Hydrological Sciences Journal*, 54(6): 989-1006. DOI:10.1623/hysj.54.6.989

-
- Wang, L., Koike, T., Yang, K., Jackson, T.J., Bindlish, R., Yang, D., (2009b). Development of a distributed biosphere hydrological model and its evaluation with the Southern Great Plains Experiments (SGP97 and SGP99). *Journal of Geophysical Research*, 114(D8). DOI:10.1029/2008jd010800
- Wang, L., Koike, T., Yang, K., Yeh, P.J.-F., (2009c). Assessment of a distributed biosphere hydrological model against streamflow and MODIS land surface temperature in the upper Tone River Basin. *Journal of Hydrology*, 377(1-2): 21-34. DOI:10.1016/j.jhydrol.2009.08.005
- Wang, L., Koike, T., Yang, K., Jin, R., Li, H., (2010a). Frozen soil parameterization in a distributed biosphere hydrological model. *Hydrology and Earth System Sciences*, 14(3): 557-571. DOI:10.5194/hess-14-557-2010
- Wang, L., Wang, Z., Koike, T., Yin, H., Yang, D., He, S., (2010b). The assessment of surface water resources for the semi-arid Yongding River Basin from 1956 to 2000 and the impact of land use change. *Hydrological Processes*, 24(9): 1123-1132. DOI:10.1002/hyp.7566
- Wang, L., Zhou, J., Qi, J., Sun, L., Yang, K., Tian, L., Lin, Y., Liu, W., Shrestha, M., Xue, Y., Koike, T., Ma, Y., Li, X., Chen, Y., Chen, D., Piao, S., Lu, H., (2017a). Development of a land surface model with coupled snow and frozen soil physics. *Water Resources Research*, 53: 5085–5103. DOI:10.1002/2017WR020451
- Wang, S., Liu, S., Mo, X., Peng, B., Qiu, J., Li, M., Liu, C., Wang, Z., Bauer-Gottwein, P., (2015). Evaluation of remotely sensed precipitation and its performance for streamflow simulations in basins of the southeast Tibetan Plateau. *Journal of Hydrometeorology*, 16(6): 2577-2594.
- Wang, W., Cui, W., Wang, X., Chen, X., (2016). Evaluation of GLDAS-1 and GLDAS-2

-
- Forcing Data and Noah Model Simulations over China at the Monthly Scale. *Journal of Hydrometeorology*, 17(11): 2815-2833. DOI:10.1175/JHM-D-15-0191.1
- Wang, X., Pang, G., Yang, M., (2017b). Precipitation over the Tibetan Plateau during recent decades: a review based on observations and simulations. *International Journal of Climatology*. DOI:10.1002/joc.5246
- Weedon, G.P., Balsamo, G., Bellouin, N., Gomes, S., Best, M.J., Viterbo, P., (2014). The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA - Interim reanalysis data. *Water Resources Research*, 50(9): 7505-7514. DOI:10.1002/2014WR015638
- Xu, W., Ma, L., Ma, M., Zhang, H., Yuan, W., (2016). Spatial-temporal variability of snow cover and depth in Qinghai-Tibetan Plateau. *Journal of Climate*. DOI:10.1175/JCLI-D-15-0732.1
- Xue, B.L., Wang, L., Yang, K., Tian, L., Qin, J., Chen, Y., Zhao, L., Ma, Y., Koike, T., Hu, Z., Li, X., (2013). Modeling the land surface water and energy cycles of a mesoscale watershed in the central Tibetan Plateau during summer with a distributed hydrological model. *Journal of Geophysical Research: Atmospheres*, 118(16): 8857-8868. DOI:10.1002/jgrd.50696
- Yang, D., 1998. Distributed hydrological model using hillslope discretization based on catchment area function: development and applications, University of Tokyo, Tokyo.
- Yang, F., Lu, H., Yang, K., He, J., Wang, W., Wright, J.S., Li, C., Han, M., Li, Y., (2017). Evaluation of multiple forcing data sets for precipitation and shortwave radiation over major land areas of China. *Hydrology and Earth System Sciences*, 21(11): 5805-5821. DOI:10.5194/hess-21-5805-2017
- Yang, K., Wu, H., Qin, J., Lin, C., Tang, W., Chen, Y., (2014). Recent climate changes over

the Tibetan Plateau and their impacts on energy and water cycle: A review. *Global and Planetary Change*, 112: 79-91. DOI:10.1016/j.gloplacha.2013.12.001

Zaitchik, B.F., Rodell, M., Olivera, F., (2010). Evaluation of the Global Land Data Assimilation System using global river discharge data and a source-to-sink routing scheme. *Water Resources Research*, 46(6): W06507. DOI:10.1029/2009wr007811

Zhao, Y., Zhu, J., (2015). Assessing Quality of Grid Daily Precipitation Datasets in China in Recent 50 Years (in Chinese). *PLATEAU METEOROLOGY*, 34(1): 50-58.

Zhou, J., Wang, L., Zhang, Y., Guo, Y., Li, X., Liu, W., (2015). Exploring the water storage changes in the largest lake (Selin Co) over the Tibetan Plateau during 2003 - 2012 from a basin - wide hydrological modeling. *Water Resources Research*, 51(10): 8060-8086. DOI:10.1002/2014WR015846

Zhou, X.Y., Zhang, Y.Q., Yang, Y.H., Yang, Y.M., Han, S.M., (2013). Evaluation of anomalies in GLDAS-1996 dataset. *Water Science and Technology*, 67(8): 1718-1727. DOI:10.2166/wst.2013.043

Table 1 Details of the studied river basins in the Tibetan Plateau

Basins	Area (km ²)	Mean elevation (m)	Mean annual precipitation (mm/year)	Mean annual temperature (°C)
Yarlung Tsangpo River basin	256,864	4627	559	0.1
Upper Yangtze River basin	137,371	4762	471	1.1
Upper Yellow River basin	123,580	4125	550	-1.7
Upper Lancang River basin	53,656	4557	641	-0.8

Accepted Article

Table 2 Details of the global precipitation products

Product	Spatial resolution	Temporal resolution	Areal coverage	Temporal coverage	Data sources
TRMM3B42 V7	0.25°	3h	Global 50°N-S	1998-present	G, S
GSMaP-Gauge V5	0.1°	1h	Global 60°N-S	2000-present	G, S
CMORPH-BLD 1.0	0.25°	Daily	Global 90°N-S	1998-present	G, S
CHIRPS V2.0	0.25°	Daily	Global 50°N-S	1981-present	G, S, R
MSWEP V2.01	0.25°	3h	Global	1979-present	G, S, R
WFDEI-CRU	0.5°	3h	Global	1979-2016	G, R
PERSIANN-CDR	0.25°	Daily	Global 60°N-S	1983-present	G, S
GLDAS2.0	0.25°	3h	Global 90°N-60°S	1948-2010	G, S, R
GLDAS2.1	0.25°	3h	Global 90°N-60°S	2000-present	G, S, R

G: gauge; S: satellite; R: reanalysis.

Table 3 Basin average values of calibrated model parameters

	Yarlung Tsangpo River	Upper Lancang River	Upper Yangtze River	Upper Yellow River
$KS(m/s)$	1.33E-05	4.86E-06	5.99E-06	6.44E-05
$anik$	4.92	5.18	3.37	2.09
$S_{max}(m)$	1.82E-05	5.01E-03	2.54E-02	7.32E-03
a	4.05E-02	6.55E-02	3.88E-02	1.68E-02
n	2.47	2.62	2.06	2.04
θ_s	0.21	0.41	0.60	0.69
f	2.16	4.24	2.09	2.68

KS : saturated hydraulic conductivity for soil surface; $anik$: hydraulic conductivity anisotropy ratio; S_{max} : maximum surface water storage; a : van Genuchten parameter; n : van Genuchten parameter; θ_s : saturated water content; f : hydraulic conductivity decay factor.

Table 4 NSE and RB values in the calibration and validation periods in the four river basins

		Calibration period	Validation period
Yarlung Tsangpo River	NSE	0.92	0.87
	RB	0%	-10%
Upper Lancang River	NSE	0.92	0.85
	RB	-6%	2%
Upper Yangtze River	NSE	0.79	0.80
	RB	-11%	10%
Upper Yellow River	NSE	0.93	0.76
	RB	-4%	-7%

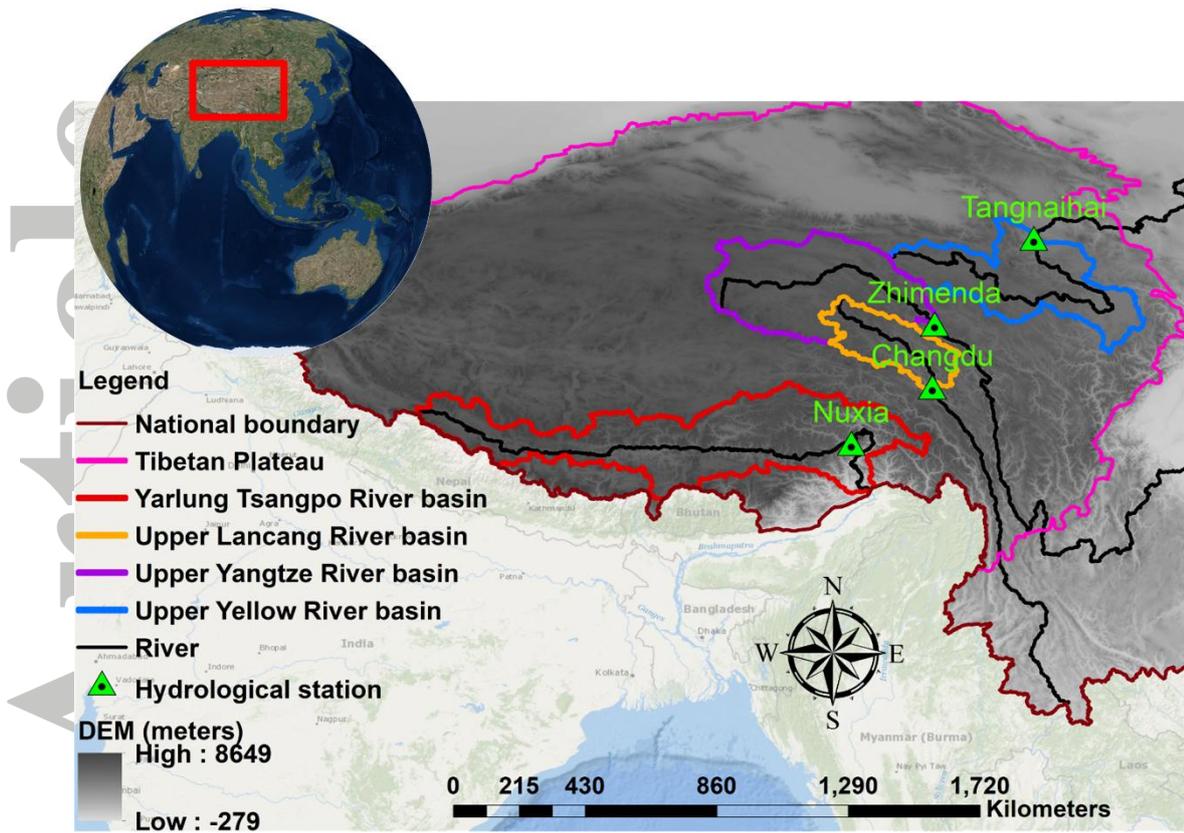


Figure 1 The four river basins studied in the Tibetan Plateau.

Accepted

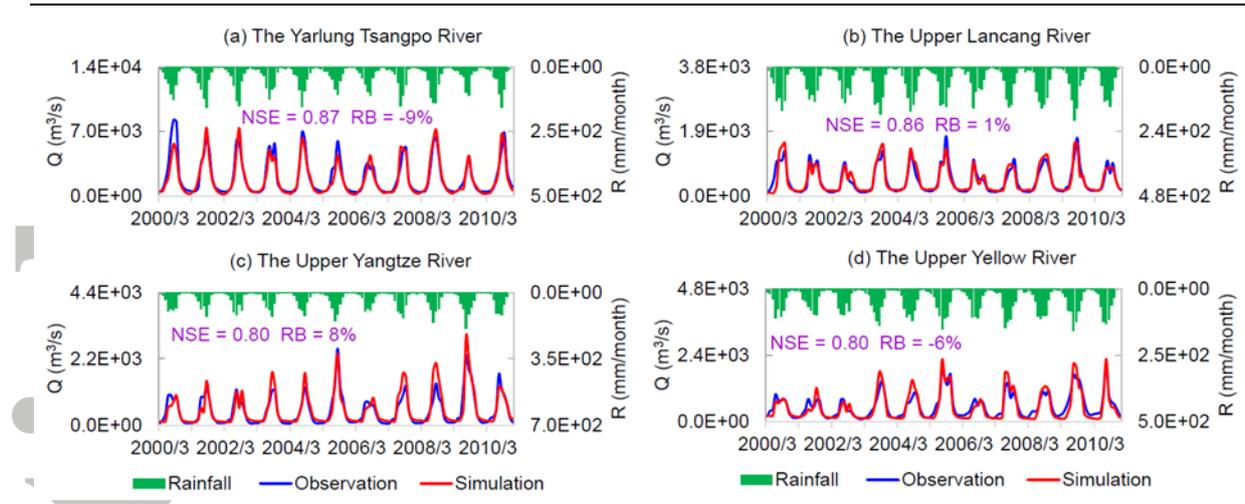


Figure 2 WEB-DHM evaluations based on monthly runoff. ‘Q’ represents runoff; ‘R’ represents rainfall intensity.

Accepted Article

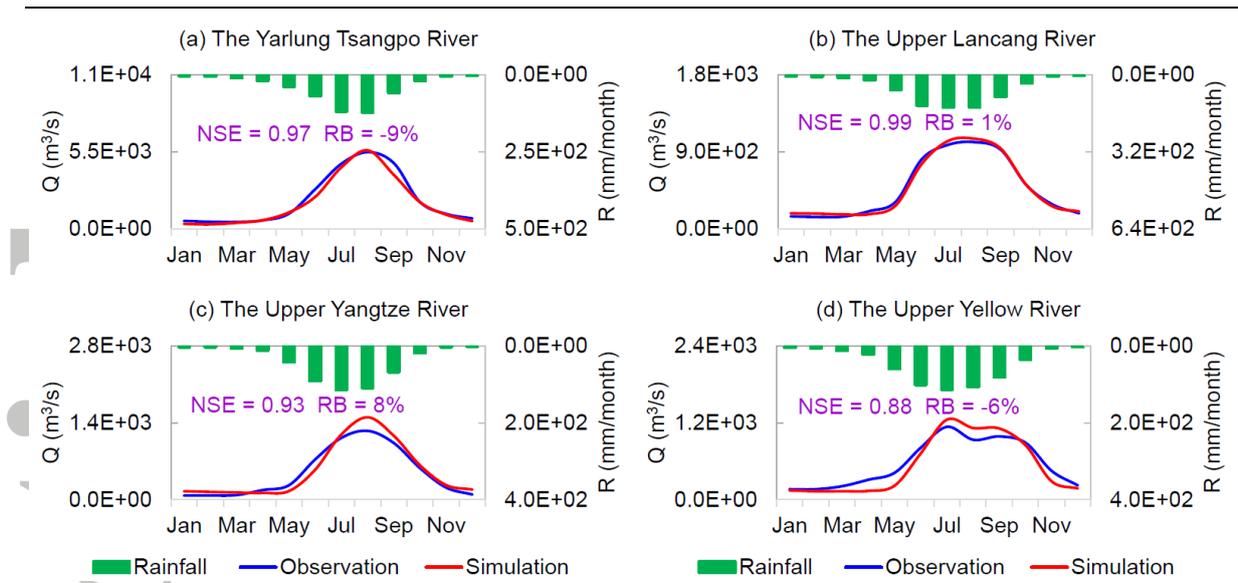


Figure 3 WEB-DHM evaluations based on multi-year average monthly runoff. ‘Q’ represents runoff; ‘R’ represents rainfall intensity.

Accepted Article

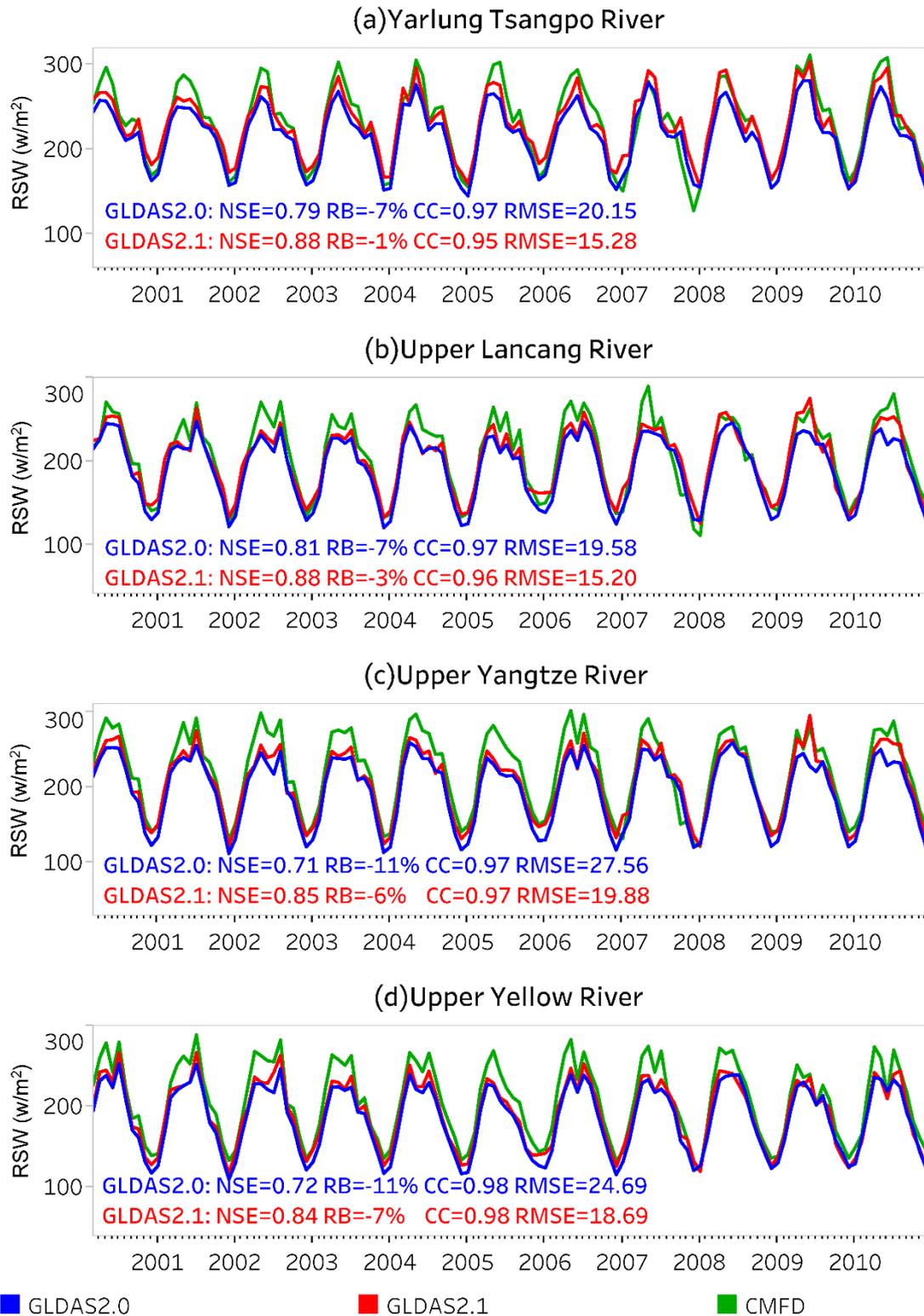


Figure 4 Downward solar radiation (RSW) comparison among the three datasets in the four river basins studied in the Tibetan Plateau.

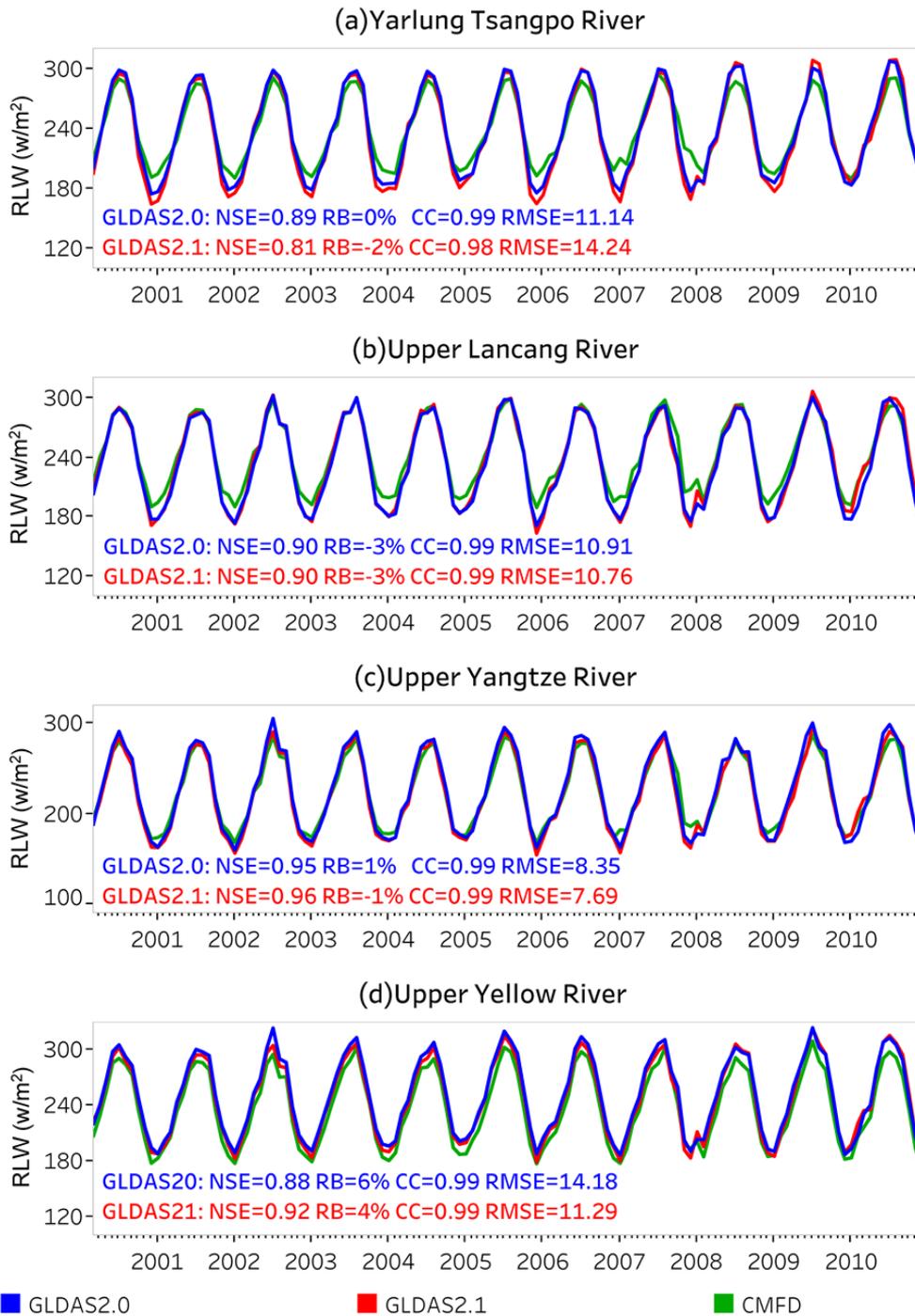


Figure 5 Downward longwave radiation (RLW) evaluations in the four river basins studied in the Tibetan Plateau.

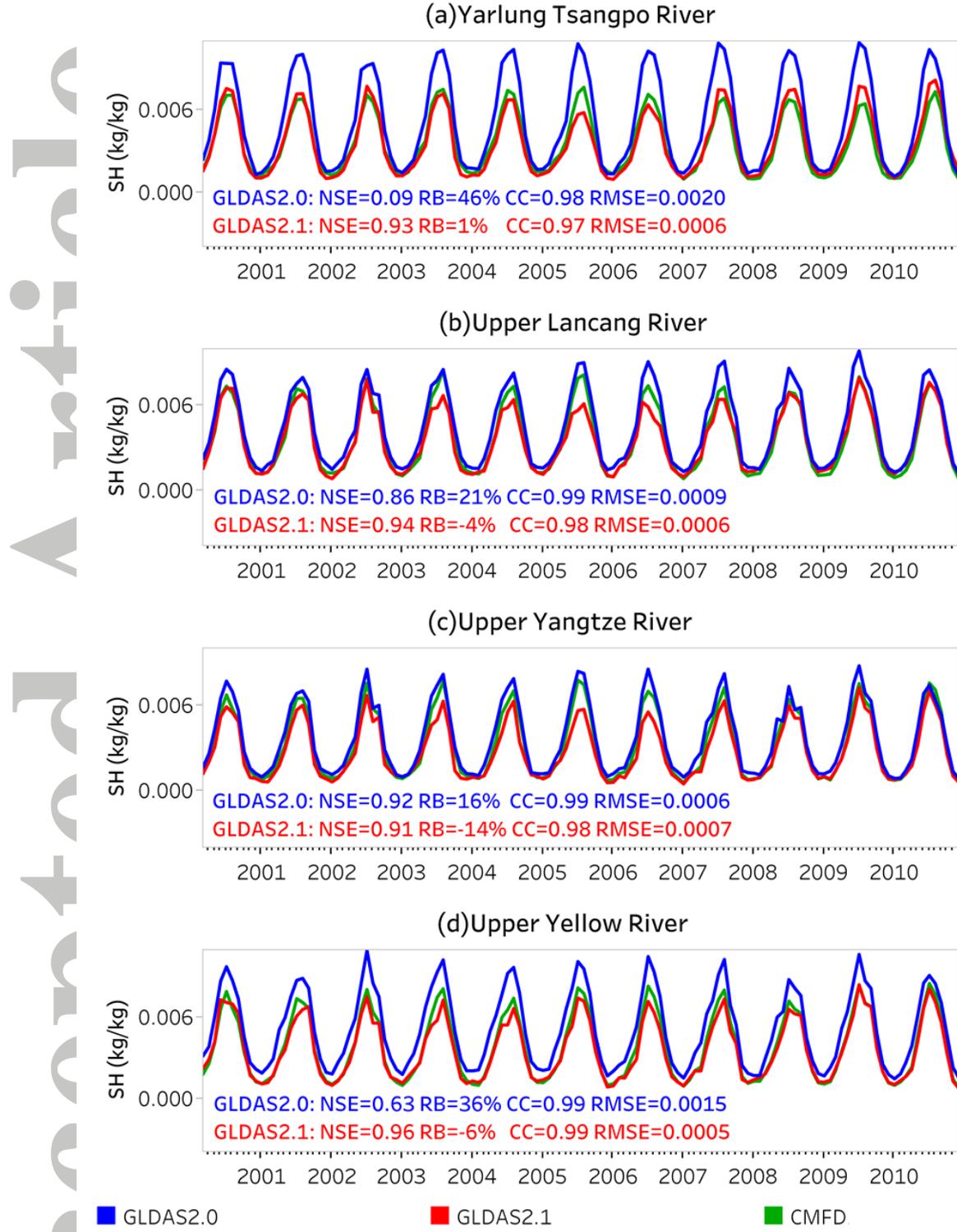


Figure 6 Specific Humidity (SH) evaluations in the four river basins studied in the Tibetan Plateau.

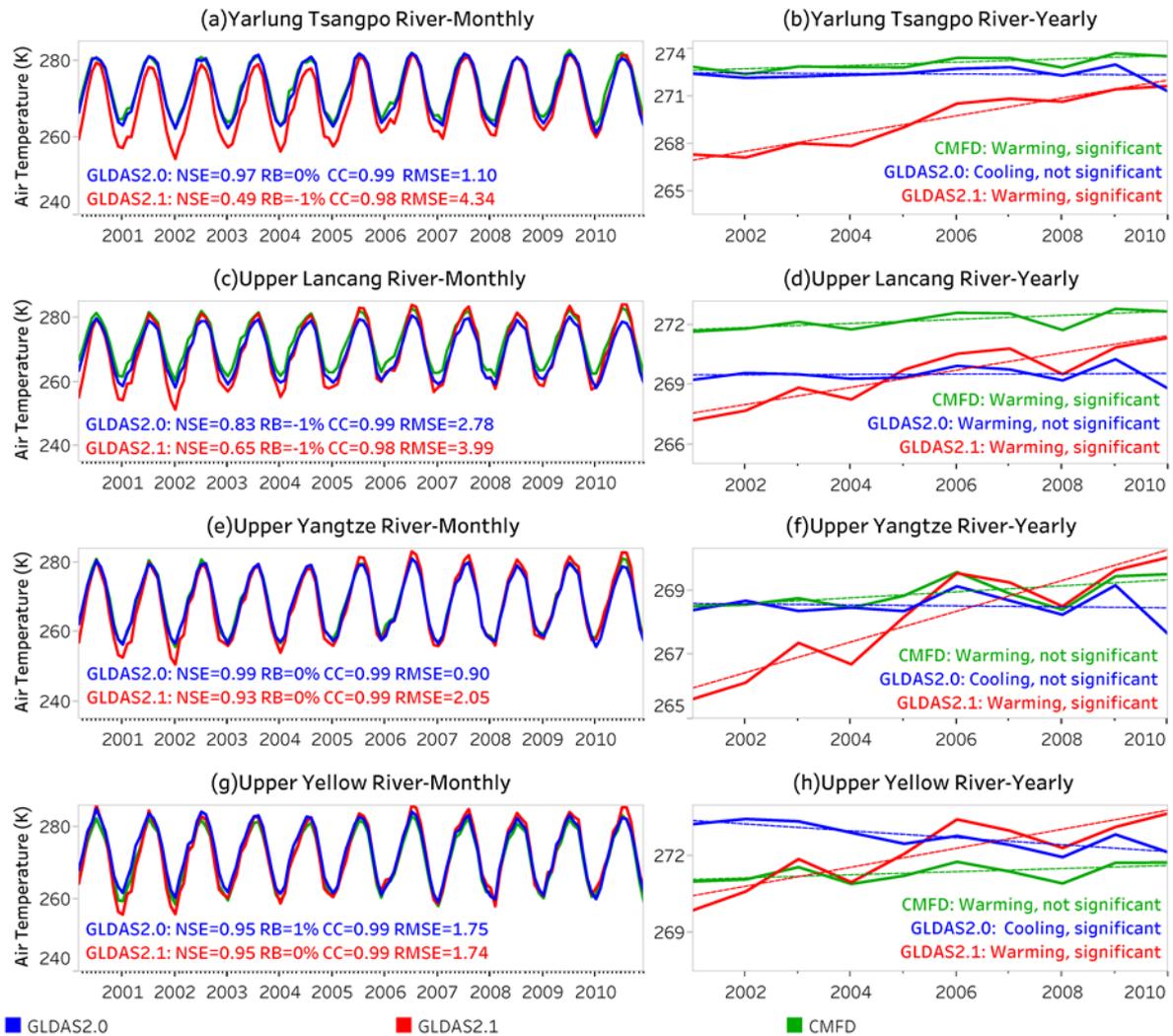


Figure 7 Comparison of air temperature (K) among CMFD, GLDAS2.0 and GLDAS2.1. The dashed lines represent the trend lines.

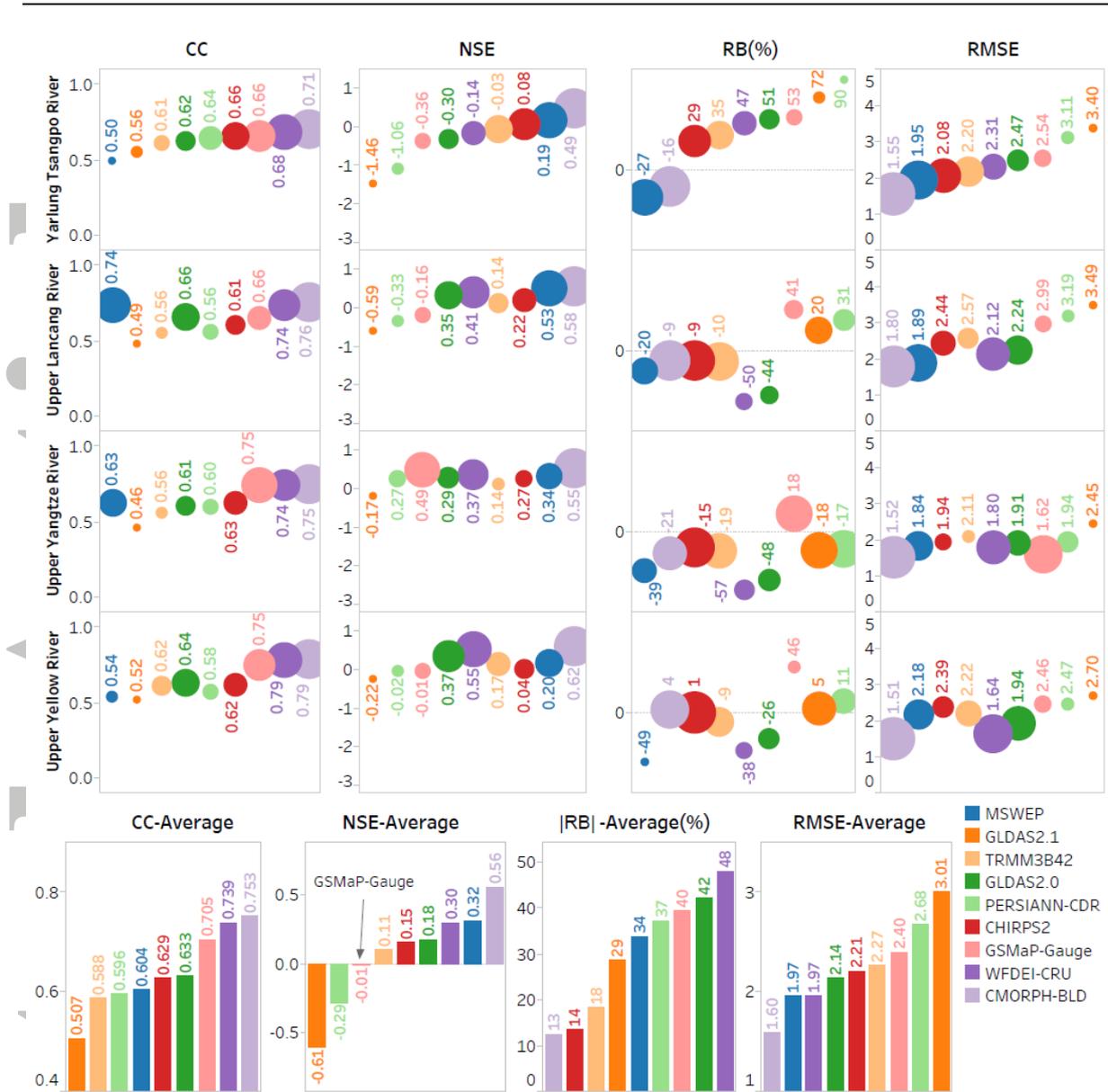


Figure 8 Evaluation results of GLDAS2.0 and 2.1 precipitation products and comparison with global precipitation products based on four criteria on a daily scale. |RB| represents the absolute values of Relative Bias (RB). The larger the circles, the better the precipitation products.

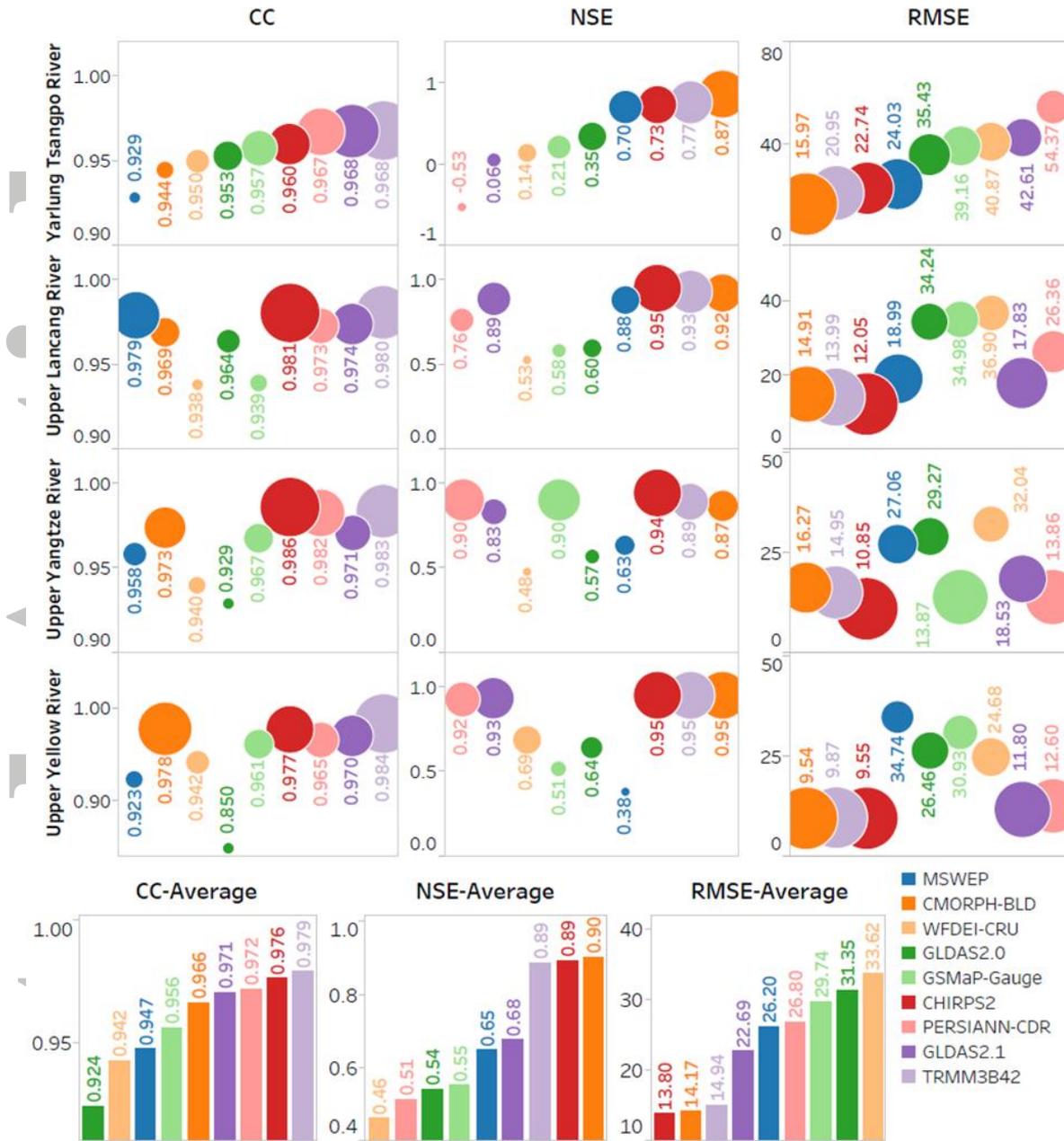


Figure 9 Evaluation results of GLDAS2.0 and 2.1 precipitation data and comparison with global precipitation products on a monthly scale.

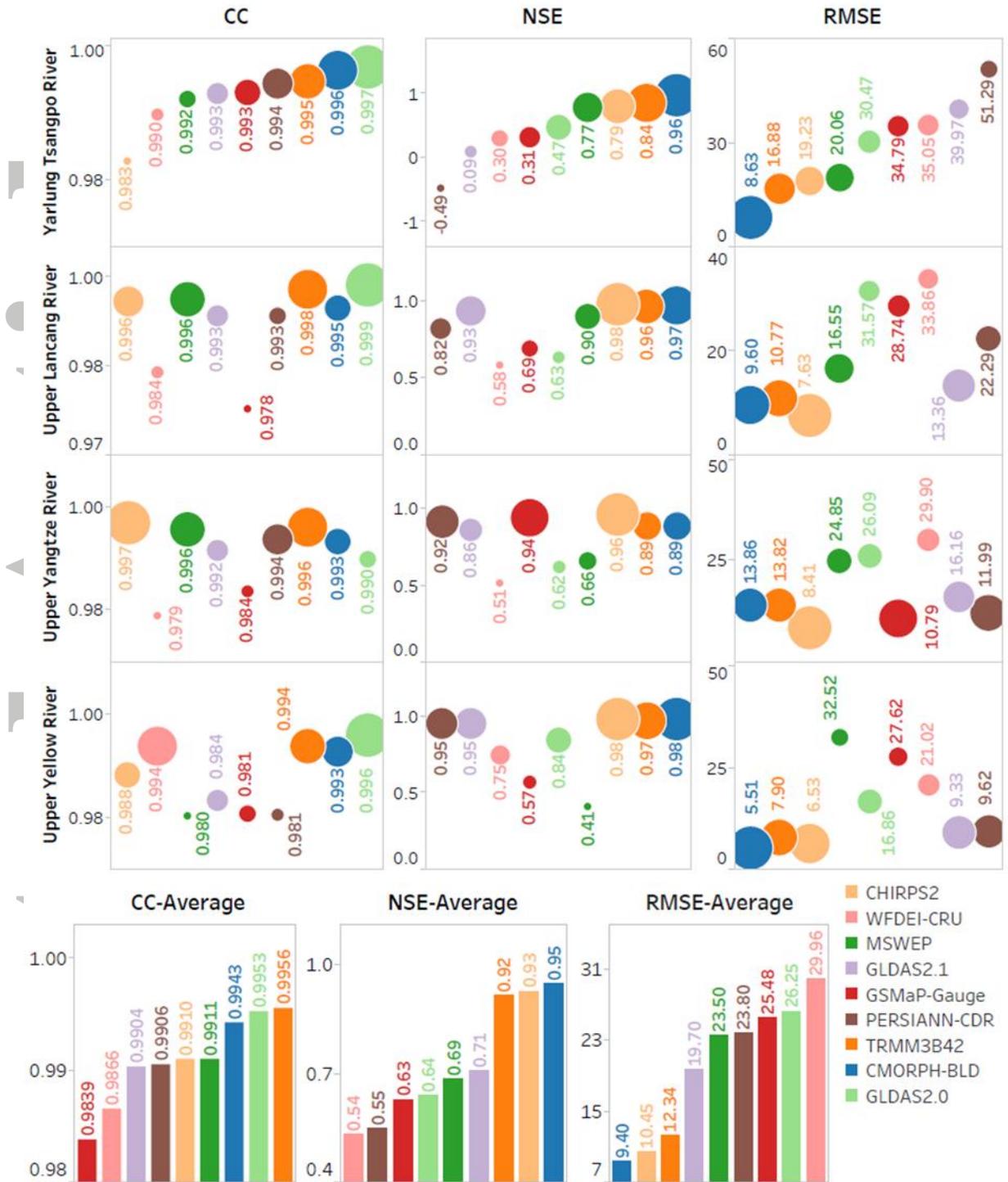


Figure 10 Evaluation results of GLDAS2.0 and 2.1 precipitation data and comparison with global precipitation products on a multi-year average monthly scale.

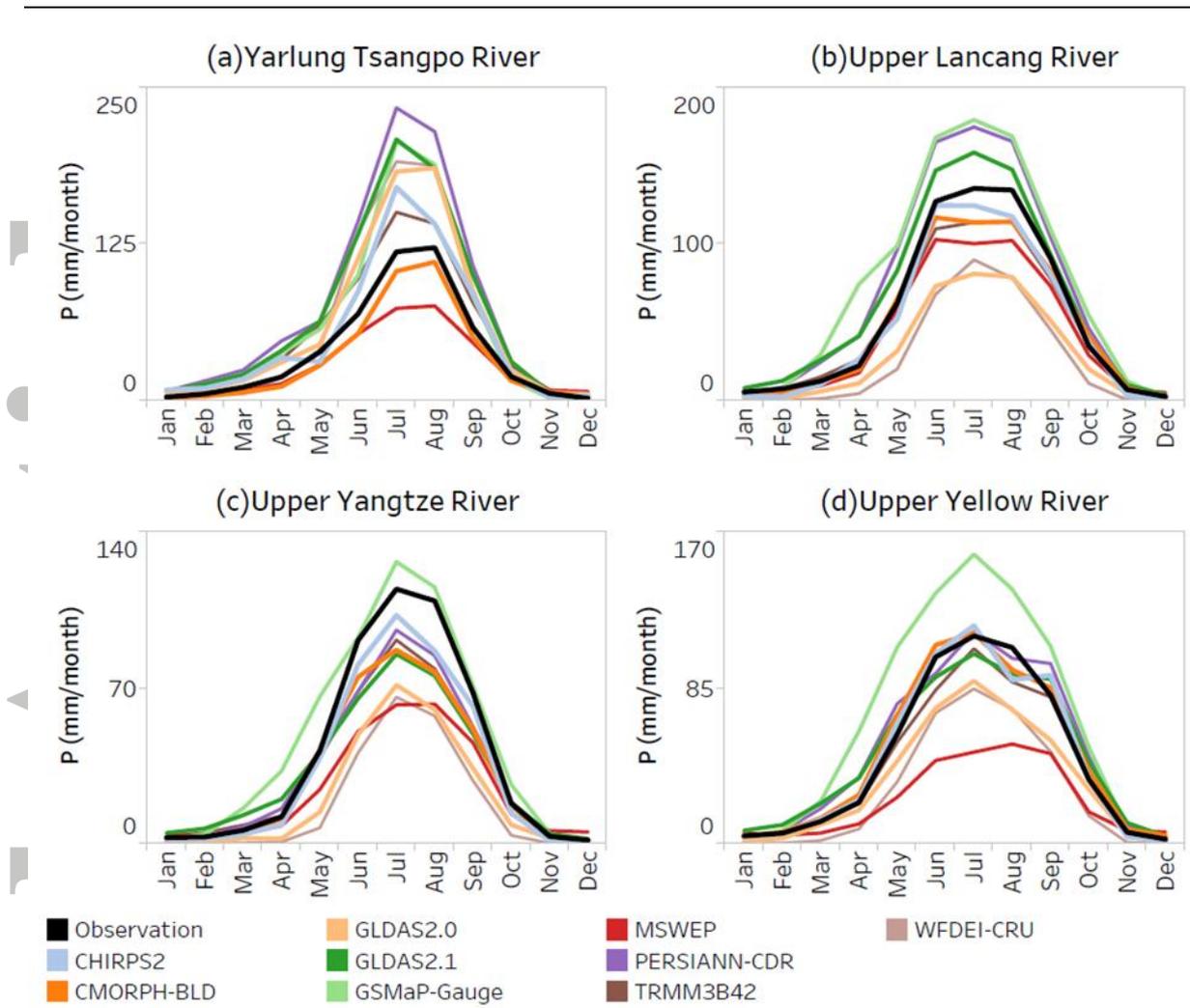


Figure 11 Precipitation data evaluations on a multi-year average monthly scale in time series plots.

Accepted

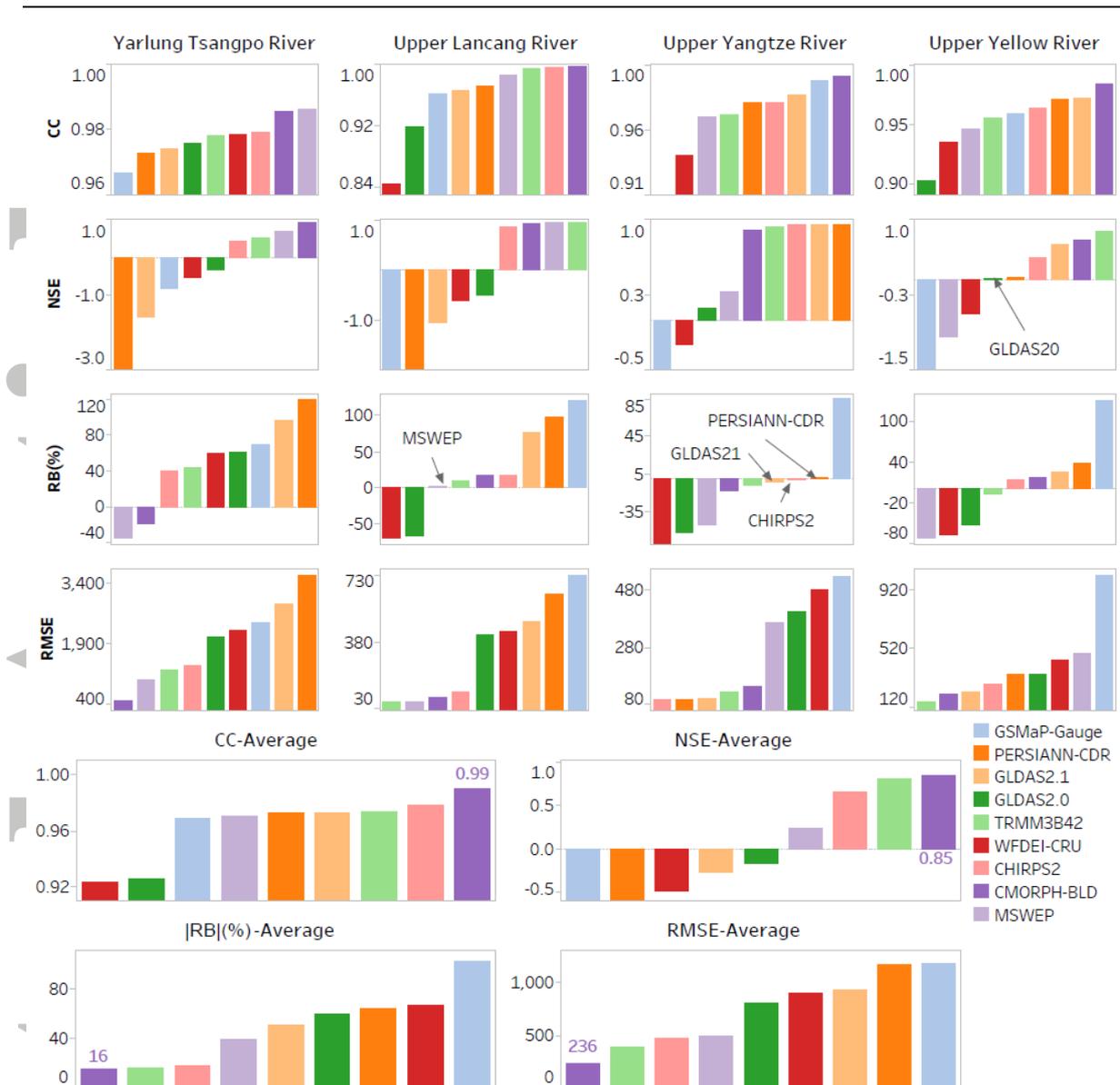


Figure 12 Evaluations of simulated runoff using GLDAS2.0 forcing data and precipitation product data on a multi-year average monthly scale. The ‘Average’ represents the average values of the evaluation criteria over the four river basins studied. The average values of |RB| represent the average of absolute RB values.

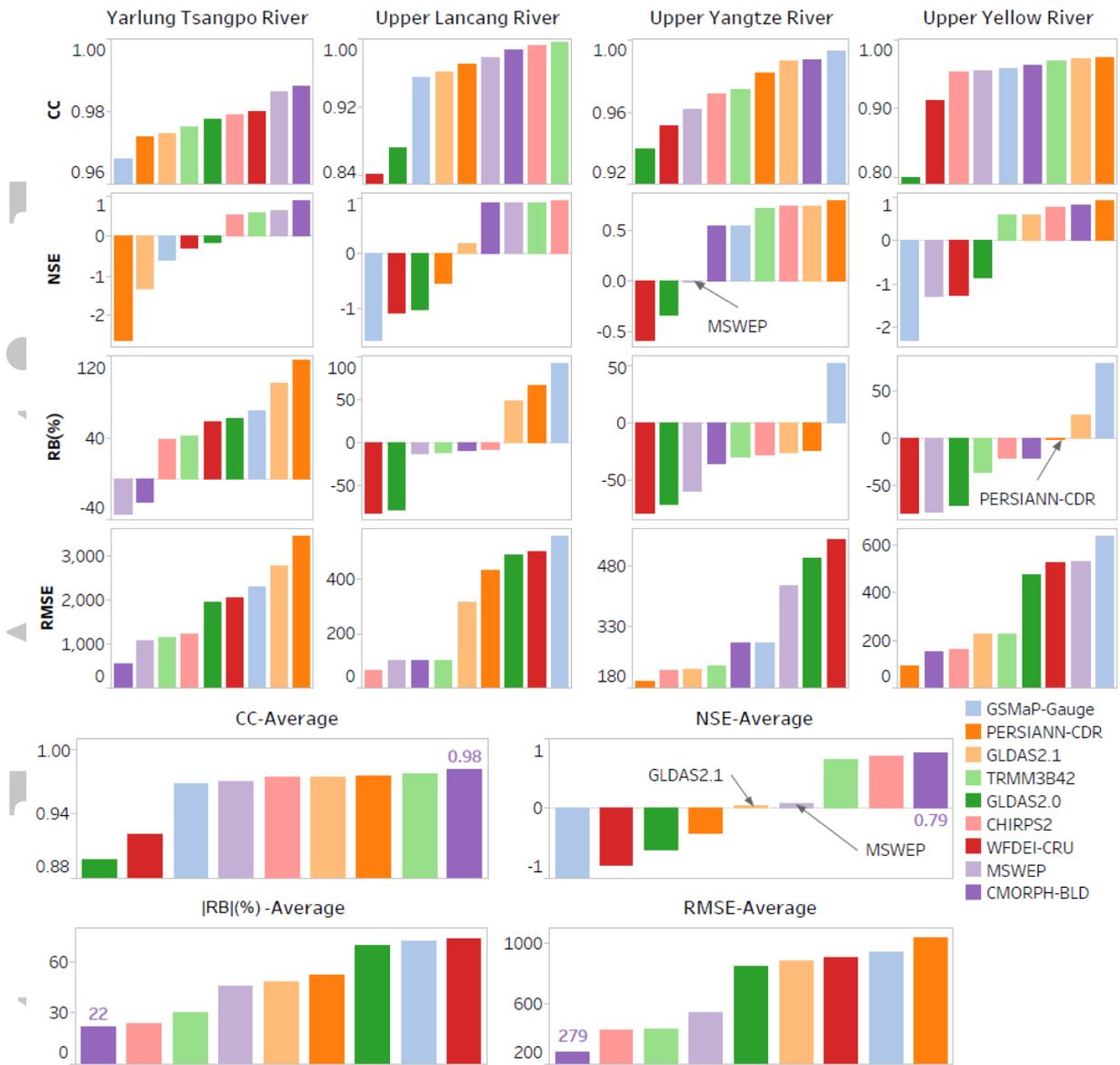


Figure 13 Evaluations of simulated runoff using GLDAS2.1 forcing data and precipitation product data on a multi-year average monthly scale.