Earth and Space Science

RESEARCH ARTICLE
10.1029/2020EA001576

Key Points:

- Specific humidity and wind speed have large uncertainty, and precipitation has large uncertainty in summer
- Trends of precipitation, air temperature, downward longwave radiation, and wind speed are consistent with observation
- Equations are developed to reduce forcing data uncertainty, and validations show that the equations are effective

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

J. Liu,
junguo.liu@gmail.com;
liujg@sustech.edu.cn

Citation:

Qi, W., Liu, J., Yang, H., Chen, D., & Feng, L. (2022). Assessments and corrections of GLDAS2.0 forcing data in four large transboundary rivers in the Tibetan Plateau and Northeast China. *Earth and Space Science*, *9*, e2020EA001576. https://doi.org/10.1029/2020EA001576

Received 20 NOV 2020
Accepted 19 NOV 2021

Author Contributions:

Conceptualization: Wei Qi
Data curation: Wei Qi
Formal analysis: Wei Qi
Investigation: Wei Qi
Methodology: Wei Qi
Resources: Wei Qi
Supervision: Junguo Liu
Validation: Wei Qi
Writing – original draft: Wei Qi

Assessments and Corrections of GLDAS2.0 Forcing Data in Four Large Transboundary Rivers in the Tibetan Plateau and Northeast China

Wei Qi[1], Junguo Liu[1], Hong Yang[2,3], Deliang Chen[4], and Lian Feng[1]

[1]School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen, China, [2]Eawag, Swiss Federal Institute of Aquatic Science and Technology, Duebendorf, Switzerland, [3]Department of Environmental Science, MUG, University of Basel, Basel, Switzerland, [4]Department of Earth Sciences, Regional Climate Group, University of Gothenburg, Gothenburg, Sweden

Abstract GLDAS2.0 provides long-term fine resolution gridded hydrometeorological data sets, which are necessary for water-related studies, particularly in some transboundary rivers that are partially without observation. Yet, GLDAS2.0 has only been validated at limited locations, and few studies have been conducted to develop approaches to correct the GLDAS2.0 data for transboundary rivers. This work assessed the GLDAS2.0 data and developed approaches to correct their uncertainties for studies in large transboundary rivers in the Tibetan Plateau and Northeast China (NC). To achieve these goals, observational data from 1982 to 2010 and a water and energy budget-based distributed hydrological model including biosphere after calibration and validation were employed. We find that the GLDAS2.0 data (except for wind speed) can reasonably replicate observed seasonal variations. However, its specific humidity and wind speed have large uncertainty, and precipitation has large uncertainty in summer. In NC, the trends of its precipitation, air temperature, downward longwave radiation, and wind speed are consistent with the observations. In the Yarlung Tsangpo, Lancang, and Nu Rivers, the trends of all GLDAS2.0 data reproduce the observation very well, that is, wetting, warming, and dimming trends. Validations show that the corrections are effective and the corrected forcing data can be successfully used in hydrological simulation with improved performance than the raw GLDAS2.0 data, which demonstrates the usefulness of the methodology and corrected forcing data to hydrometeorological studies in transboundary rivers in China as well as in other nearby regions/countries.

Plain Language Summary GLDAS2.0 provides long-term hydrometeorological data sets, which are necessary for water-related studies, particularly in some transboundary rivers that are partially without observation. This work assessed GLDAS2.0 data and developed approaches to correct their uncertainties for studies in large transboundary rivers in the Tibetan Plateau (TP) and Northeast China (NC). In NC, the trends of GLDAS2.0 precipitation, air temperature, downward longwave radiation, and wind speed are consistent with observations. In TP, the trends of all the GLDAS2.0 data reproduce observation very well. Validations show that the corrections are effective and the corrected data have improved performance than raw GLDAS2.0 data in hydrological simulation.

1. Introduction

Many large river basins in the world have vast geographical coverages, spreading across regions and countries. The hydrological data observed and accessible are usually unevenly distributed for different parts/catchments of a given river basin: Some have relatively abundant data and others hardly any. Lack of observed data in parts of a river basin often impedes the application of hydrological models to adequately simulate river flows and to provide information to support a holistic water resources management on the river basin scale. Filling this data gap is therefore important. This is particularly so for transboundary rivers where the data in some parts are not available or difficult to obtain.

The Global Land Data Assimilation System (GLDAS) provides long-term fine resolution gridded meteorological data on a global scale (Rodell et al., 2004), which are beneficial to water-related studies in some transboundary rivers lacking in situ observation. However, GLDAS used satellite and reanalysis-based meteorological data sets, which have large uncertainties and therefore, applicability evaluations of GLDAS data are a key issue to its successful application (Qi et al., 2018; Scanlon et al., 2018; Zaitchik et al., 2010). Many studies have

Writing – review & editing: Wei Qi, Junguo Liu, Hong Yang, Deliang Chen, Lian Feng

evaluated uncertainty in GLDAS1.0 (Chen et al., 2013; Huang et al., 2013; Kato et al., 2007; Qi et al., 2015; Wang et al., 2011; Zaitchik et al., 2010; Zhou et al., 2013). For example, Qi et al. (2015) and Wang et al. (2011) investigated input and output of GLDAS1.0 and developed correction equations to reduce the uncertainty of GLDAS1.0 forcing data in two rivers (less than 18,000 km^2 in total) in Northeast China; Bai et al. (2016) investigated runoff uncertainty in GLDAS1.0 in the Tibetan Plateau and found overestimation of runoff by GLDAS1.0. Regarding GLDAS2.0, a few studies have investigated its applicability. For example, Wang, Li, et al. (2016) investigated the uncertainty of land surface temperature estimation at a point scale and found underestimation generally; Wang, Cui, et al. (2016) investigated the applicability of GLDAS2.0 precipitation, evapotranspiration, air temperature, water storage, and runoff in China and also found underestimation; Liu et al. (2020) evaluated the precipitation data of GLDAS2.0; and Qi, Liu, Yang, et al. (2020) studied the runoff data of GLDAS2.0. Despite these progresses, few studies have been carried out for evaluation and correction of GLDAS2.0 data, such as radiation, specific humidity, and wind speed in a large area for transboundary rivers specifically in the Tibetan Plateau and Northeast China.

China has several large transboundary rivers, for example, the Amur River/Heilong Jiang, Yarlung Tsangpo—Brahmaputra River, Lancang—Mekong River, and Nu-Salween River. The Amur River forms the border between Northeast China and the Russian Far East region and sustains the developments of agriculture and industry of the region (Simonov & Egidarev, 2017). The Brahmaputra River, Mekong River, and Salween River originate from the Tibetan Plateau of China and flow to South and Southeast Asian countries, such as Thailand, Laos, Myanmar, Cambodia, Vietnam, and Bangladesh. Because of the rich water resources, high hydropower potential, upstream and downstream flood and drought management, and interests from multiple countries on them, these large transboundary rivers have drawn much attention globally.

The overall objective of this study is to evaluate GLDAS2.0 forcing data and to explore correction approaches for GLDAS2.0 data in four large transboundary rivers in China. They are the Amur River, Brahmaputra River, Mekong River, and Salween River. The data used are the GLDAS_NOAH025_3H_2.0 product (Matthew & Hiroko Kato, 2015). A distributed hydrological model including biosphere after calibration and validation was implemented. The data from 1982 to 2010 were studied considering the availability of observation data. By comparing the differences of the uncertainty in Northeast China and Tibetan Plateau, the necessity of developing equations for the uncertainty corrections specifically in the Tibetan Plateau was revealed. This paper is unique in that, for the first time, it evaluates GLDAS2.0 forcing data specifically for the four large transboundary rivers in the Tibetan Plateau and Northeast China and also provides the corrected data with reduced uncertainty.

2. Study Region and Data Sets

2.1. Study Region

The Amur River is the tenth longest river in the world and has a catchment area of around 2 million km^2. The Amur River basin is located in Northeast China and an adjacent area outside China. Figure 1a shows the locations of Northeast China, which covers a total area of 1.22 million km^2. The Brahmaputra River basin has a catchment area of about 0.65 million km^2. The Yarlung Tsangpo River basin is located in the upstream of the Brahmaputra River and accounts for about 51% of the Brahmaputra River basin. The Salween River basin has a catchment area of about 0.26 million km^2, and its upper stream is located within China (i.e., the Nu River) accounting for about 47% of the total area of the Salween River basin. The Mekong River basin is about 0.81 million km^2. In China, the Mekong River is named the Lancang River, which has a catchment area of about 0.16 million km^2. Figure 1b shows the locations of the Brahmaputra River basin, the Salween River basin, and the Mekong River basin. The Yarlung Tsangpo River, Lancang River, and Nu River basins will be termed as Southwest China river basins hereafter.

2.2. Data Sets

In the Southwest China river basins, the China Gauge-based Daily Precipitation Analysis (CGDPA) data were used. CGDPA is based on about 2,400 precipitation gauges (Shen & Xiong, 2016; Zhao & Zhu, 2015). Other data were from the China Meteorological Forcing Dataset (CMFD) (He & Yang, 2011; Xue et al., 2013; Yang et al., 2017; Zhou et al., 2015). For Northeast China, the input data used were the same as our previous study, and more details can be found in the study by Qi, Feng, et al. (2020).

10.1029/2020EA001576

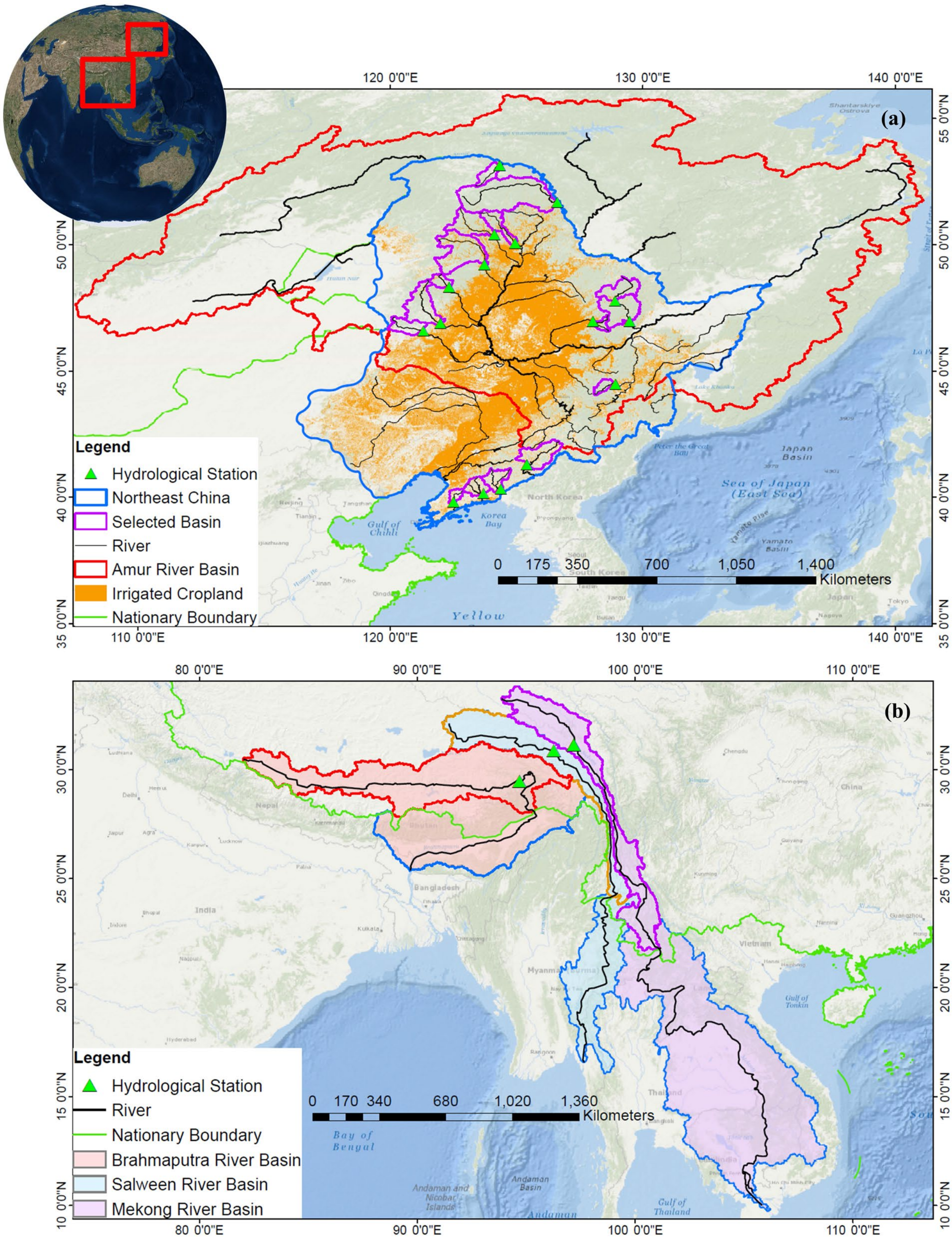

Figure 1. Study regions. (a) Northeast China and basins off the main irrigated cropland for model calibration. (b) Southeast China river basins and hydrological gauges in the Tibetan Plateau, which are close to natural conditions. The irrigated cropland distribution in Northeast China is from http://waterdata.iwmi.org/applications/irri_area/.

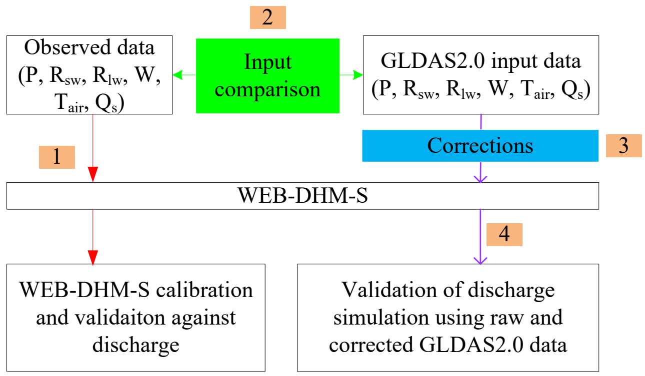

Figure 2. Flowchart of this study. P = Precipitation. Rsw = Downward shortwave radiation. Rlw = Downward longwave radiation. W = Wind speed. Tair = Air temperature. Qs = Specific humidity.

The GLDAS2.0 forcing data originate from the Princeton meteorological data set (Sheffield et al., 2006). Specifically, the precipitation originates from disaggregated Global Precipitation Climatology Project (GPCP) monthly data (Huffman et al., 2001) using Tropical Rainfall Measuring Mission (TRMM) 3B42 real-time data, National Center for Atmospheric Research precipitation data, and Climatic Research Unit (CRU) TS2.0 data (Kalnay & Cai, 2003; Sheffield et al., 2006). Downward shortwave and longwave radiation data of GLDAS2.0 are produced based on the National Centers for Environmental Prediction (NCEP) (Kalnay et al., 1996) and NASA Langley Surface Radiation Budget (SRB) data (Stackhouse et al., 2011). GLDAS2.0 air temperature is generated based on the CRU TS2.0 and NCEP data. GLDAS2.0 uses wind speed and specific humidity from NCEP without correcting bias (Sheffield et al., 2006). Our study included the data before 2000, which is important for long-term studies. GLDAS2.1 does not have data before 2000 and therefore, it was not included in this study.

3. Methodology

3.1. Overview

The flowchart of the method is shown in Figure 2. The methodology included four main steps. First, the water and energy budget-based distributed biosphere hydrological model (WEB-DHM-S) (Wang, Koike, Yang, & Yang, 2009; Wang, Koike, Yang, Jackson, et al., 2009; Wang, Koike, Yang, & Yeh, 2009) was calibrated and validated using observed discharge. Second, GLDAS2.0 data were compared with in situ observed data sets. Third, correction approaches were developed to correct the uncertainty in the GLDAS2.0 data. Fourth, simulated discharge using corrected GLDAS2.0 forcing data was evaluated by comparing it against observation.

3.2. Assessment Criteria

The uncertainty and temporal trends of annual averages of a number of variables from 1982 to 2010 were studied. The criteria used for uncertainty evaluation were correlation coefficient (R), Mean Bias Error (MBE), Root Mean Square Error (RMSE), and unbiased RMSE (ubRMSE).

$$\mathrm{MBE} = \frac{\left(\sum_{i=1}^{n} X_{si} - \sum_{i=1}^{n} X_{oi}\right)}{n} \tag{1}$$

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

$$\mathrm{ubRMSE} = \sqrt{\frac{\sum_{i=1}^{n} \left[(X_{si} - \overline{X_s}) - (X_{oi} - \overline{X_o})\right]^2}{n}} \tag{3}$$

where X_{si} and X_{oi} represent GLDAS data and observation at time i, respectively; n represents number of data points; $\overline{X_s}$ and $\overline{X_o}$ represent the average of GLDAS data and observation, respectively.

Relative Bias (RB) and Nash-Sutcliffe Efficiency (NSE) were used for evaluating discharge simulation, which has been commonly implemented (Qi et al., 2015; Qi, Zhang, Fu, Sweetapple, & Zhou, 2016; Qi, Zhang, Fu, & Zhou, 2016; Qi, Zhang, Fu, Zhou, & Liu, 2016):

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{n} (Q_{pi} - Q_{ti})^2}{\sum_{i=1}^{n} (Q_{ti} - \bar{Q}_t)^2} \tag{4}$$

$$\mathrm{RB} = \frac{\sum_{i=1}^{n} Q_{pi} - \sum_{i=1}^{n} Q_{ti}}{\sum_{i=1}^{n} Q_{ti}} \times 100\% \tag{5}$$

where Q_{pi} and Q_{ti} represent simulated and observed runoff at time i, respectively; $\overline{Q_t}$ represents average of observed runoff. The criteria were used in the cross-validation by evaluating the model performance in the calibration and validation periods, respectively. In the calibration period, NSE and RB were used to evaluate the performance of the calibrated model parameters. In the validation period, NSE and RB were used to validate the model performance in discharge predictions using the calibrated model parameters.

3.3. The Hydrological Model, Its Calibration, and Validation

The WEB-DHM-S combines a geomorphology-based hydrological model (Yang, 1998) with a Simple Biosphere scheme (SiB2) and improved snow physics (Shrestha et al., 2010; Qi et al., 2015, 2018, 2019; Qi, Liu, et al., 2020; Qi, Zhang, Fu, Sweetapple, & Zhou, 2016; Wang, Koike, Yang, & Yang, 2009; Wang, Koike, Yang, & Yeh, 2009). More detailed descriptions of the model can be found in Wang, Koike, Yang, Jackson, et al. (2009).

WEB-DHM-S was calibrated and validated in Northeast China in our previous study (i.e., Qi, Feng, et al., 2020) and showed good performances in discharge simulation. More details about the model calibration, validation, calibrated model parameters, and model performance evaluations can be found in Qi, Feng, et al. (2020). The Yarlung Tsangpo, Lancang, and Nu Rivers are close to the natural environment, and therefore, the hydrological gauge data in their mainstreams were used for calibration and validation of the model parameters. Data from 1982 to 1996 were used for calibration for the Yarlung Tsangpo River and Lancang River; data from 1997 to 2010 were used for validation. Data from 1982 to 1992 were used for calibration for the Nu River (7-year data and no data for 1986, 1988, 1989, and 1990), and data from 1993 to 2000 (8-year data) for validation. The Dynamically Dimensioned Search algorithm was used in the calibration on a monthly scale because daily scale observed runoff data were not available. In the three rivers originated from the Tibetan Plateau, the sum of NSE and |RB| was used as the objective function, that is, minimize (|NSE−1| + |RB|).

4. Results and Discussion

4.1. Model Evaluations

Average parameter values after calibration are listed in Table S1 in Supporting Information S1 for the river basins in the Tibetan Plateau. Table S2 in Supporting Information S1 shows the NSE and RB values for the river basins in the Tibetan Plateau. The overall NSE and RB are up to 0.95% and 1.0%, respectively, in the Southwest China river basins. The overall NSE and |RB| are 0.80% and 8.0%, respectively, in Northeast China (Qi, Feng, et al., 2020). Overall, these results show that WEB-DHM-S can simulate the hydrological processes well in the studied regions. In addition, these results suggest that the 1,000 objective function evaluations and the data used in the calibration are acceptable for obtaining reasonable hydrological simulations. Therefore, WEB-DHM-S will be utilized to validate the correction methods for GLDAS2.0 forcing data later with the help of discharge simulation.

4.2. GLDAS2.0 Data Evaluations

Figure 3 shows the assessments of regional averages of GLDAS2.0 data from January 1982 to December 2010 on a monthly scale in Northeast China.

GLDAS2.0 has higher precipitation estimation than observation especially in summer with RMSE, ubRMSE, MBE, RB, and R being 0.36, 0.36, 0.05, 3.95%, and 0.98, respectively. Different from precipitation, GLDAS2.0

10.1029/2020EA001576

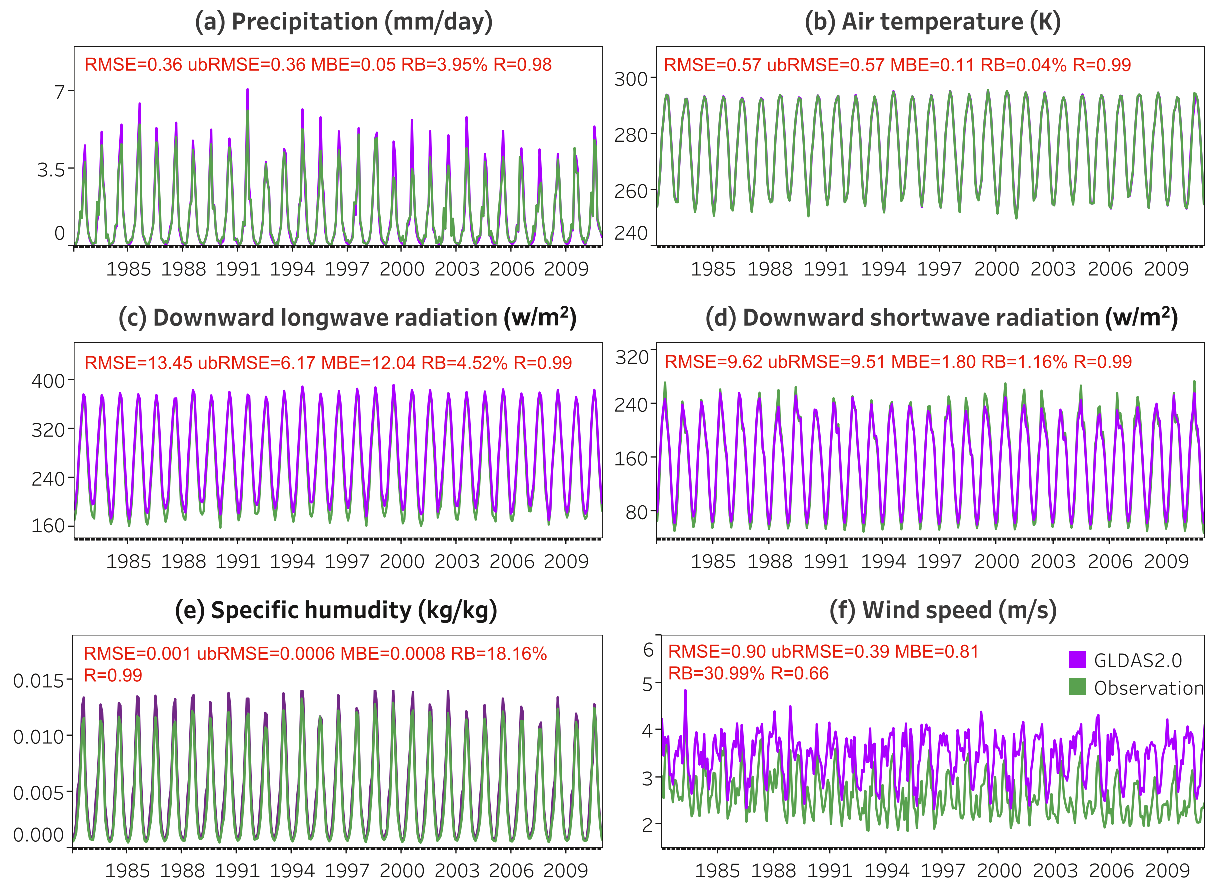

Figure 3. Assessments of regional averages of GLDAS2.0 forcing data in Northeast China from January 1982 to December 2010 on a monthly scale.

downward longwave radiation has lower error in summer periods than other months. Different from both precipitation and downward longwave radiation, GLDAS2.0 downward shortwave radiation has higher estimation in summer and lower estimation in winter than the observation. For specific humidity and wind speed, GLDAS2.0 largely overestimates with RB being 18.16% and 30.99%, respectively. Comparatively, GLDAS2.0 and the observation have a good agreement for air temperature estimation with RMSE, ubRMSE, MBE, RB, and R being 0.57, 0.57, 0.11, 0.04%, and 0.99, respectively. Similar to the results in Northeast China, GLDAS2.0 largely overestimates precipitation (especially in summer), specific humidity, and wind speed in the Southwest China river basins (as shown in Figure 4). GLDAS2.0 also has higher downward longwave radiation estimation than the observation (Figure 4c). Downward shortwave radiation of GLDAS2.0 has higher estimation than the observation in summer (Figure 4d). GLDAS2.0 and observation have a good agreement in air temperature estimation, though GLDAS2.0 slightly overestimates air temperature in summer (Figure 4b).

Figure 5 shows assessment results of GLDAS2.0 forcing data in Northeast China from January 1982 to December 2010 on a multiyear mean monthly scale. All the forcing data can reasonably replicate seasonal variations of observation with R being over 0.99 with the exception of wind speed, which has an R value of only 0.59. Figure 6 shows assessments of regional averages of GLDAS2.0 forcing data in the Southwest China river basins. Similar to the results in Northeast China, the seasonal variations of GLDAS2.0 and observation are similar except for wind speed data.

Figure 7 shows spatial and temporal variation comparisons in Northeast China. GLDAS2.0 reproduces well the seasonal and spatial variations of precipitation, air temperature, downward longwave radiation, downward shortwave radiation, and specific humidity. However, the spatial and temporal variations of wind speed data are not well replicated, which is similar to the results shown in Figure 5. Figure 8 shows the comparisons of spatial and temporal variations in the Southwest China river basins. Similar to the results in Northeast China, GLDAS2.0 replicates seasonal and spatial variations well with the exception of wind speed data.

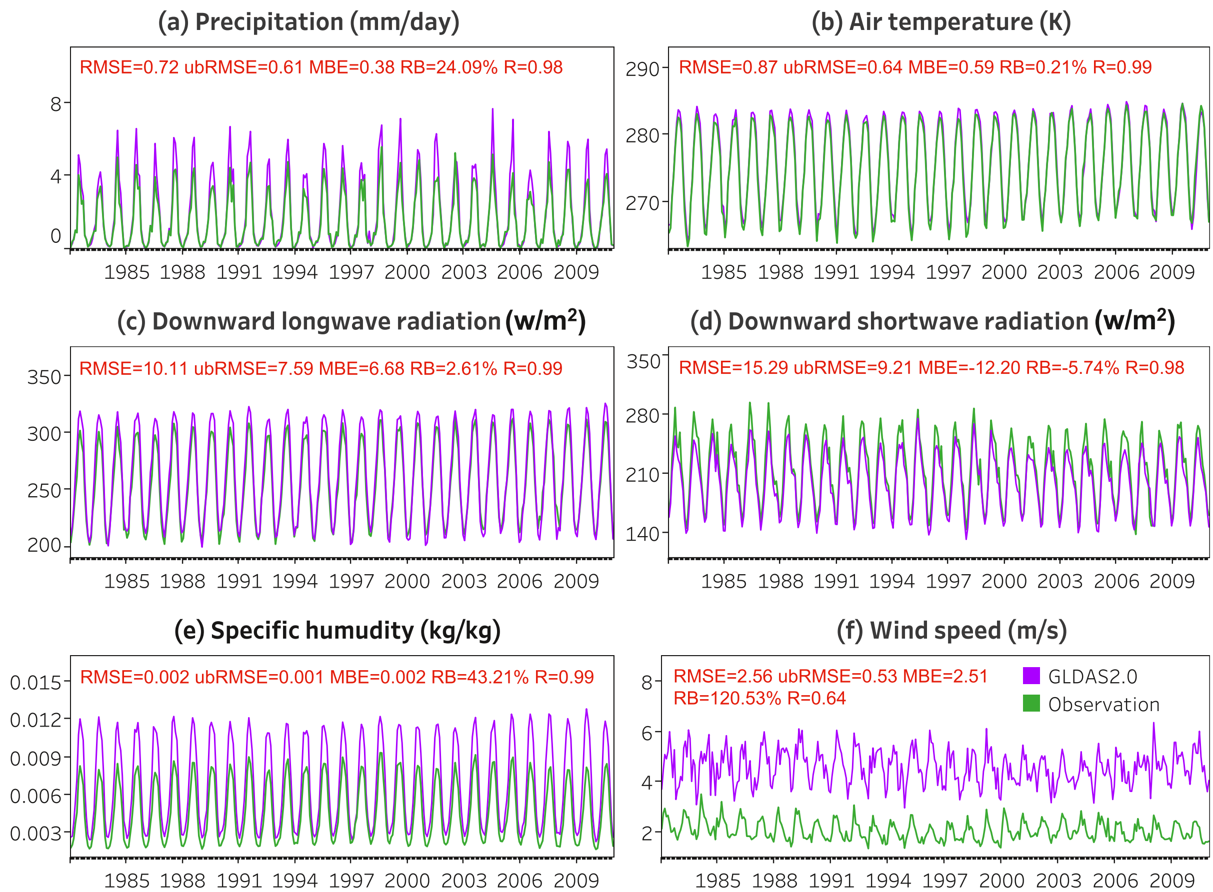

Figure 4. Assessments of regional averages of GLDAS2.0 forcing data in the Southwest China river basins from January 1982 to December 2010 on a monthly scale.

Figure 9 compares changing trends of annual averages of observation and GLDAS2.0 data from 1982 to 2010 in Northeast China at a significance level of 0.05. The nonparametric Mann-Kendall test was used to test the significance. Both GLDAS2.0 and observation show decreasing trends in precipitation and wind speed, though the decreasing trend of wind speed is significant only for the observation. Different from precipitation and wind speed, observation shows that air temperature has a significant increasing trend, but temperature of GLDAS2.0 shows an insignificant increasing trend, which is similar to the results of downward longwave radiation (Figure 9d). For downward shortwave radiation and specific humidity, the observation shows insignificant increasing trends, whereas GLDAS2.0 shows decreasing trends. Figure 10 compares the trends of annual averages of observation and GLDAS2.0 data from 1982 to 2010 in the Southwest China river basins. Both GLDAS2.0 and observation show that there are increasing trends for precipitation, air temperature, downward shortwave radiation, and specific humidity, and that there are decreasing trends for downward longwave radiation and wind speed. Therefore, the changing trends of all GLDAS2.0 forcing data reproduce the observations very well in the Southwest China river basins, and both observational data and GLDAS2.0 show that the Southwest China river basins have a wetting, warming, and dimming trend.

The high bias in heavy rainfall intensity estimation in GLDAS2.0 may result from uncertainty in TRMM data. In other months, the reason that GLDAS2.0 estimation is close to the observation may be that GLDAS2.0 utilized GPCP and CRU data to correct bias. GLDAS2.0 utilized the SRB data sets (Stackhouse et al., 2011) to correct its bias in downward shortwave and longwave radiation estimation, which may result in its good performance of representing the seasonal variations of observation. The large uncertainties in the specific humidity and wind speed are most likely due to the fact that NCEP reanalysis data are directly used to generate them without bias corrections using observations (Sheffield et al., 2006).

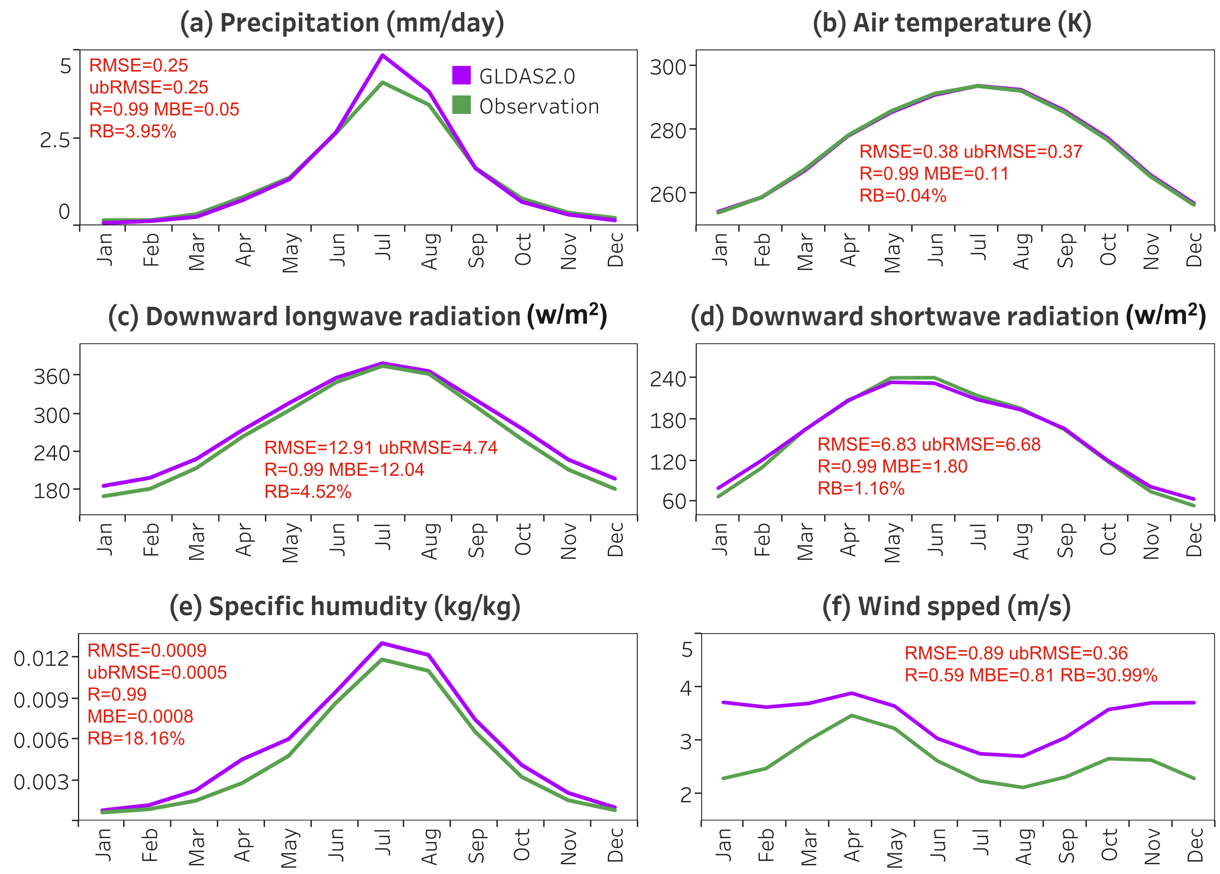

Figure 5. Assessments of regional averages of GLDAS2.0 forcing data in Northeast China from January 1982 to December 2010 on a multiyear mean monthly scale.

4.3. Correction of the GLDAS2.0 Data

Due to the uncertainties in GLDAS2.0 forcing data, it is necessary to reduce the uncertainties before the application for hydrological simulation. In this study, based on visual inspection on the uncertainty patterns in different months, linear and nonlinear regression equations derived from their scatterplots were utilized as correction functions. In Northeast China, the developed correction functions are as follows:

$$R_{sw} = 1.1015 \times R_{G,sw} - 17.688 \tag{6}$$

$$R_{lw} = 1.0636 \times R_{G,lw} - 29.749 \tag{7}$$

$$Q_s = 0.9329 \times Q_{G,s} - 0.0005 \tag{8}$$

$$\begin{cases} \text{Jan, Feb, Dec}: Wind = 0.640124 \times Wind_G \\ \text{Sep, Oct, Nov}: Wind = 0.738065 \times Wind_G \\ \text{Mar to Aug}: Wind = 0.842757 \times Wind_G \end{cases} \tag{9}$$

In the Southwest China river basins, the developed correction functions are as follows:

$$R_{sw} = 1.1529 \cdot R_{G,sw} - 18.216 \tag{10}$$

$$R_{lw} = 0.8603 \cdot R_{G,lw} + 30.079 \tag{11}$$

$$Q_s = 0.6975 \cdot Q_{G,s} + 0.000005 \tag{12}$$

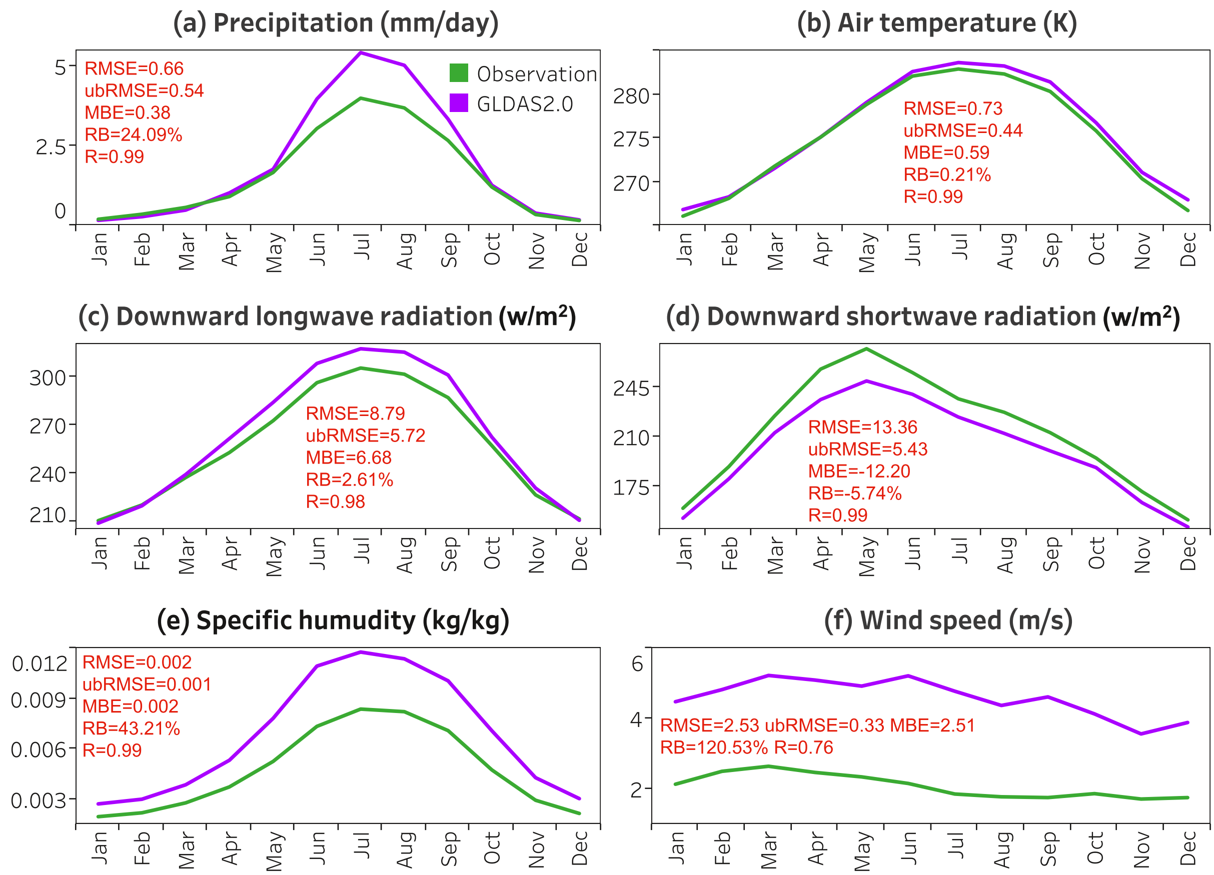

Figure 6. Assessments of regional averages of GLDAS2.0 forcing data in the Southwest China river basins from January 1982 to December 2010 on a multiyear mean monthly scale.

$$\begin{cases} \text{Jan to May}: Wind = 0.2752 \cdot {Wind_G}^2 - 1.9585 \cdot Wind_G + 5.1974 \\ \text{Jun to Sep}: Wind = 0.2752 \cdot {Wind_G}^2 - 1.9585 \cdot Wind_G + 4.7717 \\ \text{Oct to Dec}: Wind = 0.2752 \cdot {Wind_G}^2 - 1.9585 \cdot Wind_G + 5.0881 \end{cases} \tag{13}$$

where $R_{G,sw}$, $R_{G,lw}$, $Q_{G,s}$, and Wind_G represent downward shortwave radiation, downward longwave radiation, specific humidity, and wind speed data of GLDAS2.0, respectively. Figures 11 and 12 show comparisons between observations and corrected GLDAS2.0 forcing data on a multiyear mean monthly scale. It can be seen that the uncertainty is reduced significantly after the corrections. For example, RMSE, ubRMSE, MBE, and RB of wind speed reduce to 0.12, 0.12, −0.005, and −0.19%, respectively, and R improves from 0.59 to 0.96 in Northeast China. Therefore, the developed functions perform well in reducing the uncertainties.

Table 1 shows the NSE and RB/|RB| values of simulated discharge using GLDAS2.0 forcing with and without corrections on a multiyear mean monthly scale. The simulated discharge with corrected GLDAS2.0 forcing performs much better than simulation using the data without corrections. In Northeast China, average NSE increases from 0.69 to 0.77, and |RB| decreases from 17.4% to 13.1%. Similar to the results in Northeast China, in Southwest China river basins, simulated discharge with corrected GLDAS2.0 forcing performs much better than the results without correction on average: The average NSE and |RB| improve from 0.89 to 0.92 and from 9.3% to 2.8%.

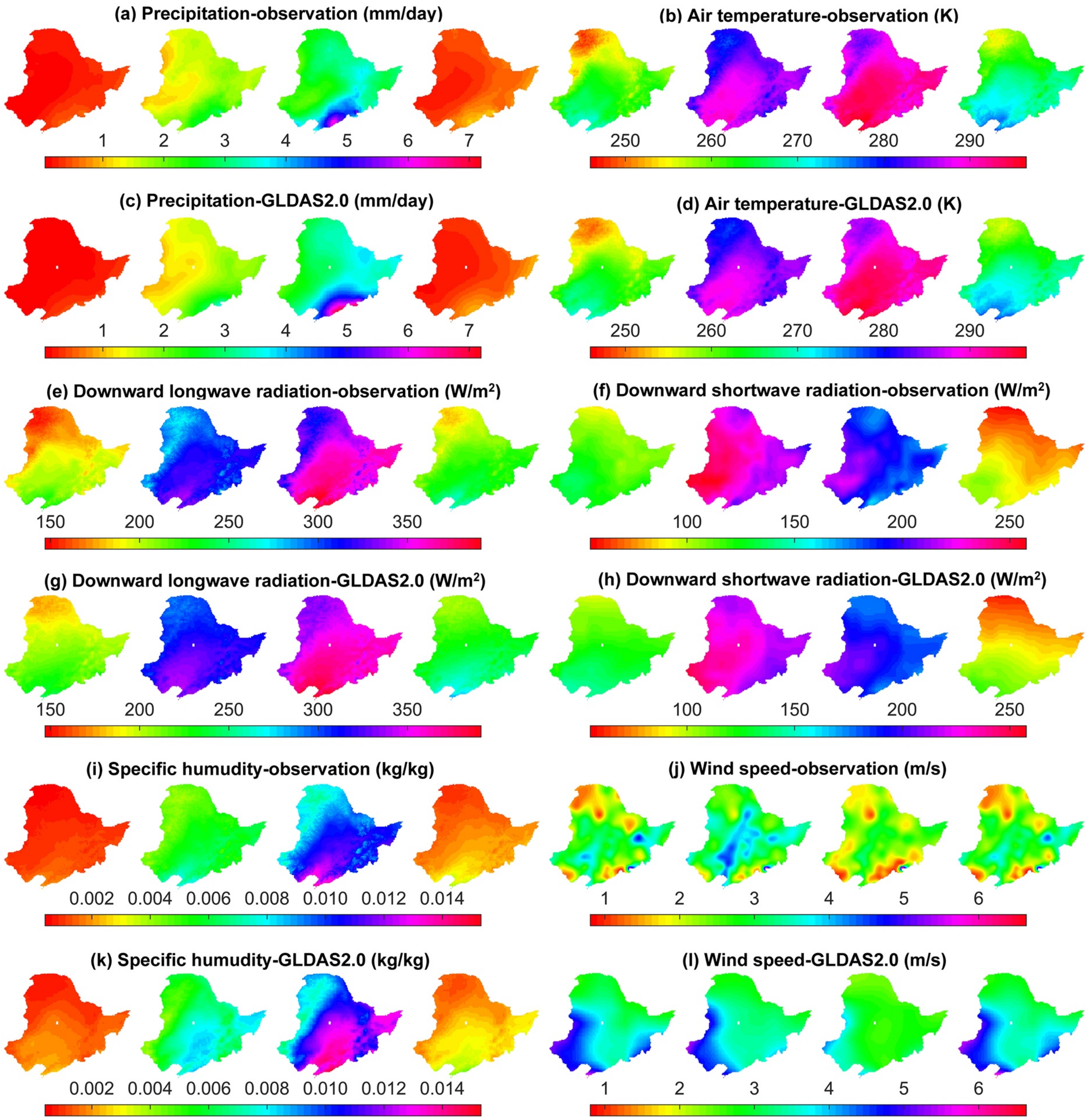

Figure 7. Spatial and temporal comparison between GLDAS2.0 forcing data and CMFD in Northeast China. Averages from January to March, April to June, July to September, and October to December are listed from left to right.

4.4. Discussion

For radiation, specific humidity, and wind speed data, there are no gauge data readily available in the studied river basins outside of China. Therefore, detecting their uncertainties in the GLDAS2.0 data set and conducting corrections are necessary for hydrological simulation. No similar studies have been carried out to develop GLDAS2.0 data correction equations for radiation, specific humidity, and wind speed data for the large transboundary rivers in the Tibetan Plateau and Northeast China. In this study, we developed methodologies for correcting the data in GLDAS2.0 for potential use in the transboundary rivers. In the west and north parts of the Tibetan Plateau, Hindu Kush, Pamir, Karakorum, and southern Himalayan ranges, in situ meteorological gauge data are sparse, which have disrupted comprehensive hydrological studies in the regions. Meanwhile, these regions are important water sources for Nepal, India, Pakistan, Afghanistan, Tajikistan, etc. The developed correction equations could also be beneficial to hydrological studies in these regions. Northeast China is adjacent to the Russian Far East

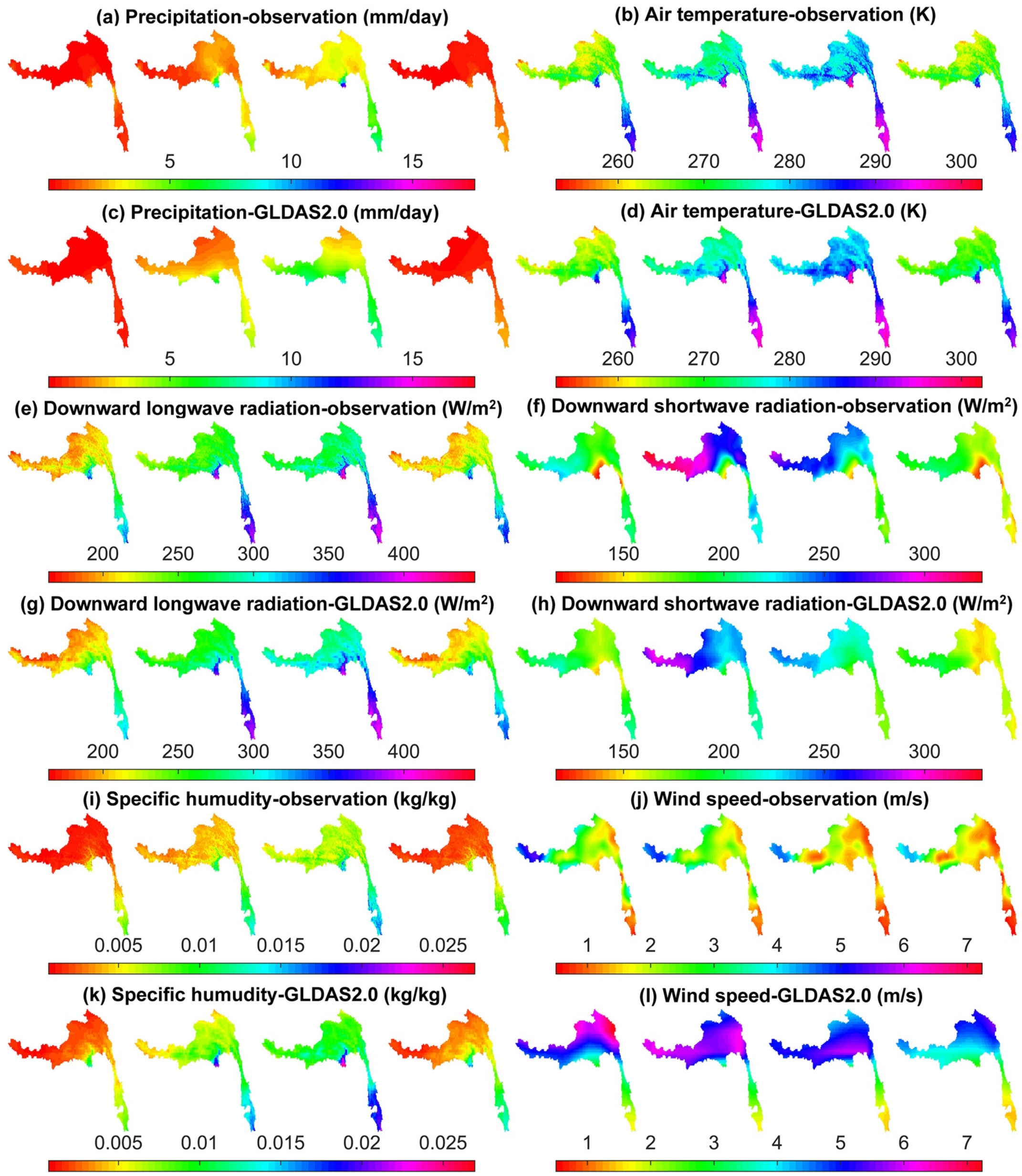

Figure 8. Spatial and temporal comparison between GLDAS2.0 forcing data and CMFD in the Southwest China river basins. Averages from January to March, April to June, July to September, and October to December are listed from left to right.

region of the Amur River basin. Therefore, studies in Northeast China on the uncertainty of GLDAS2.0 data and uncertainty correction approaches for GLDAS2.0 forcing data are beneficial to the potential applications of GLDAS2.0 data in the Amur River basin. When there are no other reliable sources of radiation, specific humidity, and wind speed data, the corrected GLDAS2.0 data based on the developed correction equations could be considered to use.

Comparisons with other studies including GLDAS data corrections are shown in Table 2. Qi et al. (2015) and Wang et al. (2011) assessed the uncertainty of GLDAS1.0/Noah in two river basins in Northeast China with the largest river basin area being only 14,700 km^2. Compared to the studies by Qi et al. (2015) and Wang et al. (2011), our study assessed the new version of GLDAS (i.e., GLDAS2.0) in entire Northeast China, which covers an area of 1.22 million km^2. Our correction methodology is based on 29 years of data. The study by Wang et al. (2011) used 7 years of data, and the study by Qi et al. (2015) used 8 years of data. Therefore, our methodology can be

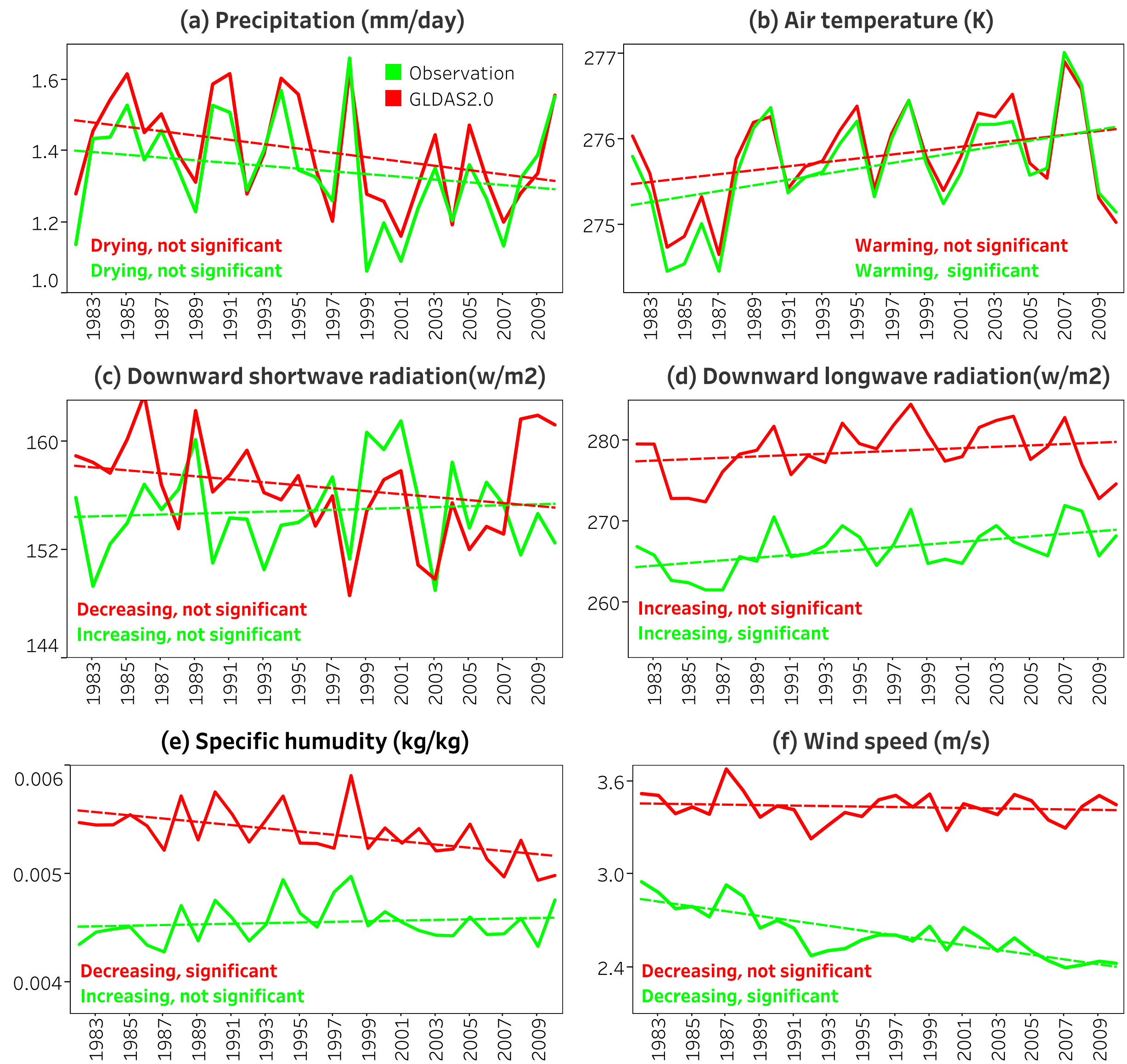

Figure 9. Comparison of changing trends of annual averages from 1982 to 2010 at a significance level of 0.05 in Northeast China.

considered more robust than the previous studies. We used a hydrological with a snow physics module. The Tibetan Plate and Northeast China have a lot of snow, and the previous studies used hydrological models that cannot consider snow melt, which may influence the evaluation of corrected GLDAS data based on hydrological simulation.

Because the changing trends of GLDAS2.0 forcing data are consistent with observation in the Southwest China river basins and the corrections do not change the trends, the corrected GLDAS2.0 forcing data are also suitable for long-term change studies around the Southwest China river basins. The corrected GLDAS2.0 forcing data are suitable for the multiyear mean monthly scale studies in regions around Northeast China, because the GLDAS2.0 data's changing trends are different from observation in Northeast China.

The Asian Precipitation—Highly-Resolved Observational Data Integration Towards Evaluation of Water Resources (APHRODITE) (Xie et al., 2007; Yatagai et al., 2012) precipitation product is freely available for the entire Asian continent and is considered the closest precipitation data to gauge observation (Malsy et al., 2015; Pritchard, 2019; Qi, Zhang, Fu, Sweetapple, & Zhou, 2016). This precipitation data have been used in many studies as a surrogate for real precipitation data (e.g., Chen et al., 2018; Pritchard, 2019). APHRODITE can be

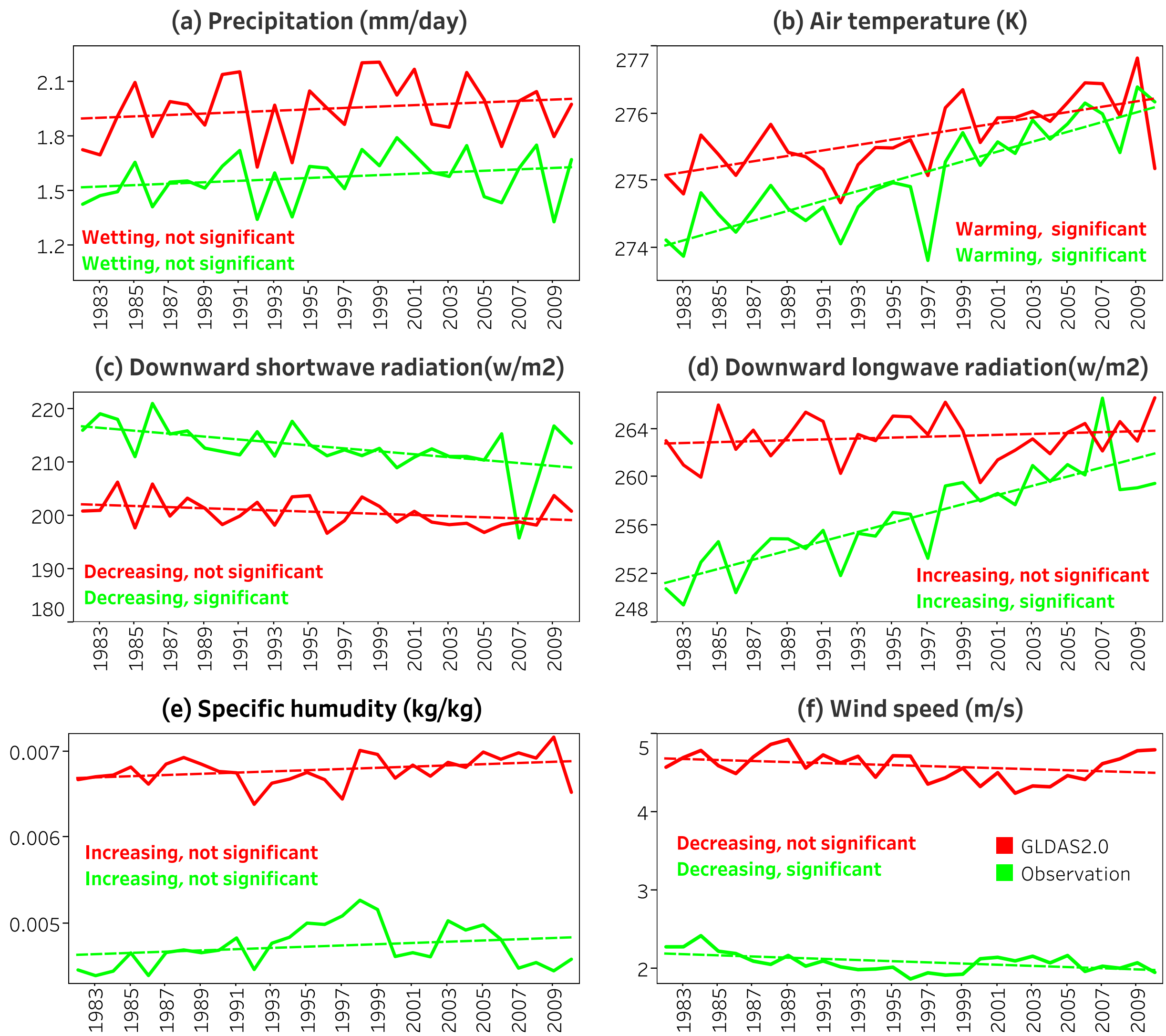

Figure 10. Comparison of changing trends of annual averages from 1982 to 2010 at a significance level of 0.05 in the Southwest China river basins.

directly used in the studied regions. Therefore, there is no need to correct the precipitation data in GLDAS2.0. In addition, APHRODITE also provides air temperature data set that is purely based on meteorological gauge data, covering the whole of Asia. Therefore, we did not correct the air temperature data of GLDAS2.0.

Our analysis based on comparisons between observation data and the corrected data has shown that the corrected data were comparable to observations. In addition, the discharge simulation using the corrected data also showed that the simulated discharge was comparable to observed discharge. Therefore, the correct equations in this study are reliable. The uncertainty in wind speed is more complex than other data. The quadratic equation for wind speed correction was selected based on the trial-and-error approach. Our trials showed the quadratic equations can effectively remove the uncertainty in wind speed and therefore, the quadratic equation was selected. The correction methodologies in this study are generic and therefore can be used for the corrections of the GLDAS2.0 data in other regions of the world. We used simple equations with only a few parameters to fit the data, which can avoid the overparameterization problems. The simple fitting methods have been validated in previous studies, and the results show that the fitting methods are acceptable, for example, Qi et al. (2015) and Wang et al. (2011). Therefore, our results are not influenced by overfitting problems.

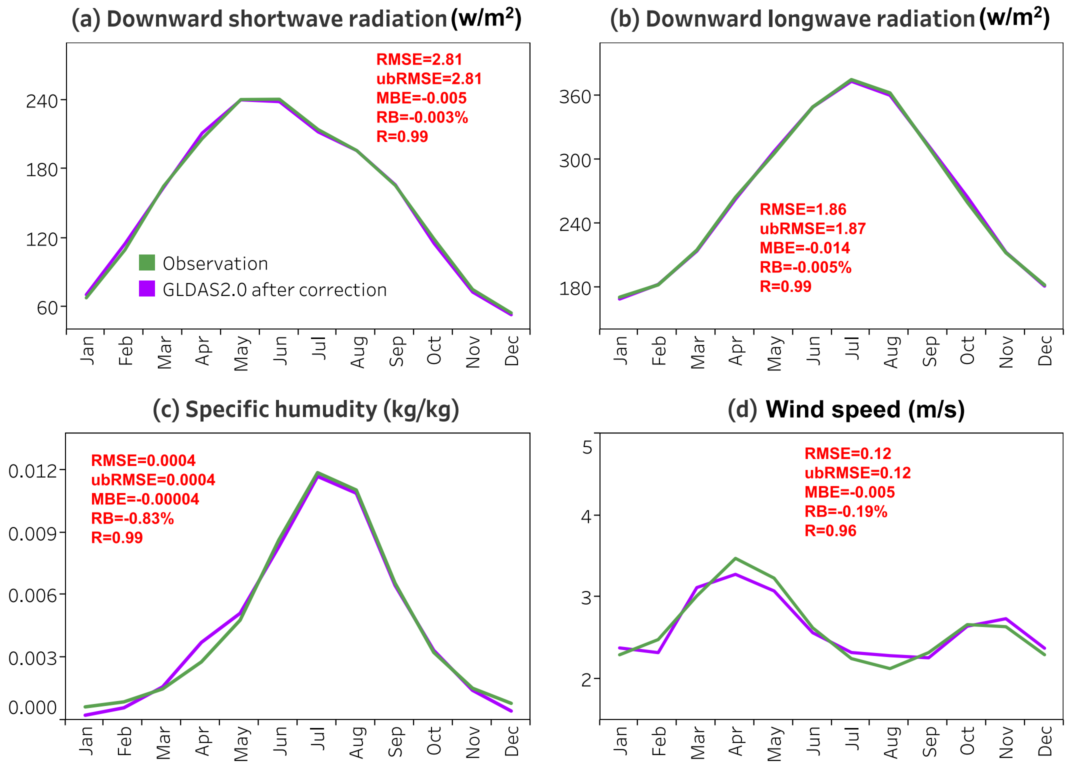

Figure 11. Comparison between observations and corrected GLDAS2.0 forcing data on a multiyear mean monthly scale in Northeast China.

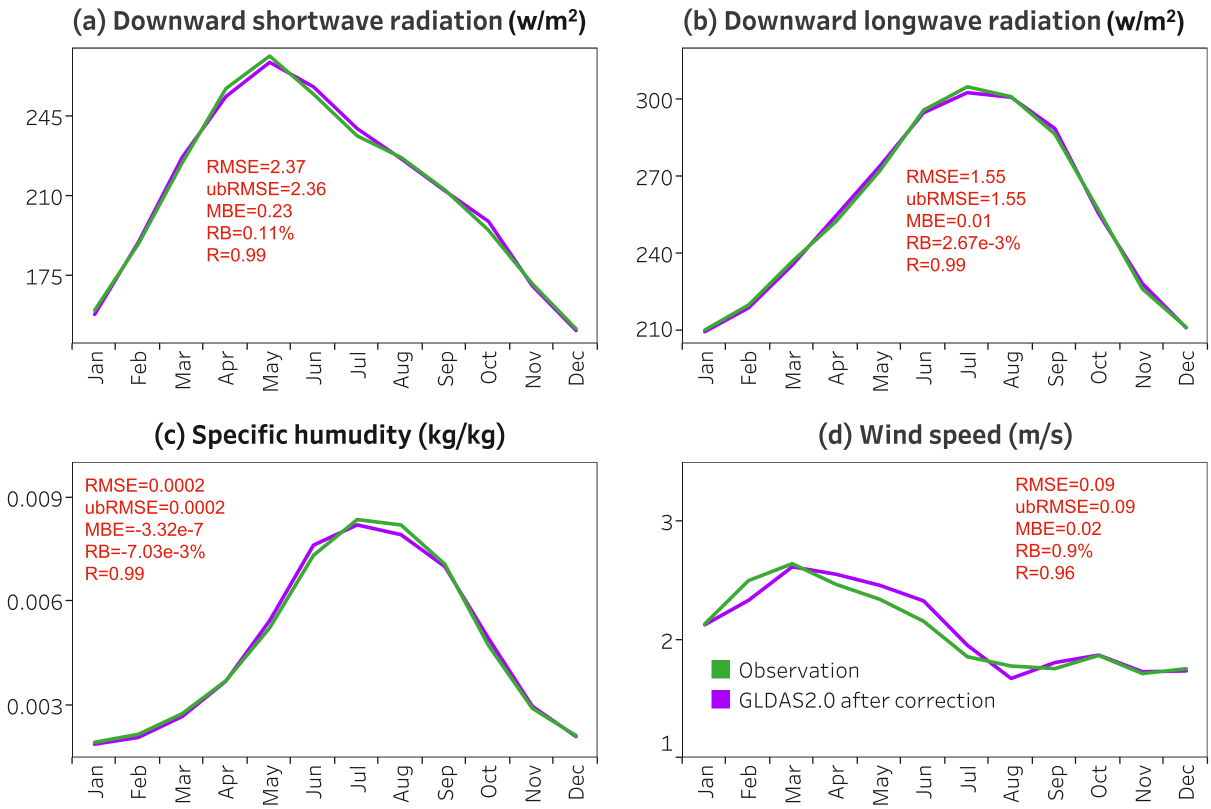

Figure 12. Comparison between observations and corrected GLDAS2.0 forcing data on a multiyear mean monthly scale in the Southwest China river basins.

10.1029/2020EA001576

Table 1
NSE and RB (|RB|) Values of Simulated Discharge Using GLDAS2.0 Forcing Data With and Without Corrections on a Multiyear Mean Monthly Scale

	GLDAS2.0: No correction		GLDAS2.0: Corrected forcing	
Region	NSE	RB/|RB|	NSE	RB/|RB|
Yarlung Tsangpo River	0.86	10.8%	0.93	−1.6%
Nu River	0.92	15.0%	0.95	5.0%
Lancang River	0.90	2.2%	0.89	−1.8%
Average	0.89	**9.3%**	0.92	**2.8%**
Northeast China	0.69	**17.4%**	0.77	**13.1%**

Note. Bold numbers represent the averages of absolute values of RB. "Average" represents the average values in the Yarlung Tsangpo, Nu, and Lancang Rivers. In Northeast China, the average values represent the average in the 16 river basins shown in Figure 1a.

Table 2
Comparisons in Terms of Data, Area, and Model Used

Study	Data	Area (km^2)	Model
Qi et al. (2015)	GLDAS1.0, 2000–2007 (8 years)	2,814	Do not consider snowmelt
Wang et al. (2011)	GLDAS1.0, 2000–2006 (7 years)	14,700	Do not consider snowmelt
Our study	GLDAS2.0, 1982–2010 (29 years)	1.8 million in total	Consider snowmelt

5. Conclusions

GLDAS2.0 provides long-term fine resolution gridded meteorological data on the global land. Yet, few studies have been conducted to evaluate GLDAS2.0 forcing data and develop correction approaches for radiation, specific humidity, and wind speed for hydrological simulations in transboundary rivers. This study assessed the GLDAS2.0 forcing data and developed correction equations for the data for four large transboundary rivers in the Tibetan Plateau and Northeast China for the first time. Observation and a water and energy budget-based distributed hydrological model after calibration were employed. The results can instill the confidence that can be placed in the GLDAS2.0 data's applicability and provide ways to improve the confidence using them. The following conclusions are presented on the basis of this study.

First, specific humidity and wind speed of GLDAS2.0 have large uncertainty in Northeast China and the Southwest China river basins; GLDAS2.0 precipitation has large uncertainty in summer in both regions. However, all the forcing data of GLDAS2.0 can reasonably replicate seasonal variations of observations except for wind speed.

Second, the temporal trends of annual average precipitation, air temperature, downward longwave radiation, and wind speed in GLDAS2.0 data are consistent with observations in Northeast China, that is, drying and warming trends. Meanwhile, the temporal trends of all the GLDAS2.0 forcing data in the Southwest China river basins reproduce the observations very well.

Third, equations are developed to reduce the uncertainty in GLDAS2.0 forcing data. Validations show that the correction equations are effective, for example, RMSE, ubRMSE, MBE, and RB of wind speed reduce to 0.12, 0.12, −0.005, and −0.19%, respectively, and correlation coefficient improves from 0.59 to 0.96 in Northeast China. Discharge simulation shows that corrected GLDAS2.0 forcing data can be used in hydrological simulation with improved performance relative to the use of the raw GLDAS2.0 data with Nash-Sutcliffe Efficiency up to 0.95 and relative bias up to −1.8%.

Data Availability Statement

GLDAS2.0 data were downloaded from https://hydro1.gesdisc.eosdis.nasa.gov/data/GLDAS/. The forcing data set 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, and Chinese Academy of Sciences (https://data.

tpdc.ac.cn/en/data/8028b944-daaa-4511-8769-965612652c49/). CGDPA was developed by the National Meteorological Information Center of China. The discharge data in Northeast China were from hydrology bureaus and can be found by contacting Songliao Water Resources Commission (http://www.slwr.gov.cn/slwjj/slwgk/). The data in the Yarlung Tsangpo, Lancang, and Nu Rivers were from hydrology bureaus and can be found by contacting the Climate Change and Water Resources in the Great River Regions in the Southeast and South Asia (http://www.lancang-mekong.net/). The generated data in this study can be found by contacting the Climate Change and Water Resources in the Great River Regions in the Southeast and South Asia (http://www.lancang-mekong.net/).

Acknowledgments

This study was supported by the Young Scientists Fund of the National Natural Science Foundation of China (51809136), the Strategic Priority Research Program of Chinese Academy of Sciences (XDA20060402), the National Natural Science Foundation of China (41971304), the Shenzhen Science and Technology Innovation Committee (JCYJ20190809155205559), the Stable Support Plan Program of Shenzhen Natural Science Fund (20200925155151006), the Shenzhen Science and Technology Program (KCXFZ20201221173007020), and the High-level Special Funding of the Southern University of Science and Technology (G02296302 and G02296402). 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.

References

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), 12180–12197. https://doi.org/10.1002/2016JD025501

Chen, A., Chen, D., & Azorin-Molina, C. (2018). Assessing reliability of precipitation data over the Mekong River Basin: A comparison of ground-based, satellite, and reanalysis datasets. *International Journal of Climatology*, *38*(11), 4314–4334. https://doi.org/10.1002/joc.5670

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. https://doi.org/10.1002/jgrd.50301

He, J., & Yang, K. (2011). *China meteorological forcing dataset*. Cold and Arid Regions Science Data Center at Lanzhou. https://doi.org/10.3972/westdc.002.2014.db

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. https://doi.org/10.5194/hess-17-1985-2013

Huffman, G. J., Adler, R. F., Morrissey, M. M., Bolvin, D. T., Curtis, S., Joyce, R., et al. (2001). Global precipitation at one-degree daily resolution from multisatellite observations. *Journal of Hydrometeorology*, *2*(1), 36–50. https://doi.org/10.1175/1525-7541(2001)002<0036:GPAODD>2.0.CO;2

Kalnay, E., & Cai, M. (2003). Impact of urbanization and land-use change on climate. *Nature*, *423*(6939), 528–531. https://doi.org/10.1038/nature01675

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., et al. (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, *77*(3), 437–471. https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2

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. https://doi.org/10.2151/jmsj.85A.187

Liu, Z., Yao, Z., Wang, R., & Yu, G. (2020). Estimation of the Qinghai-Tibetan Plateau runoff and its contribution to large Asian rivers. *Science of the Total Environment*, *749*, 141570. https://doi.org/10.1016/j.scitotenv.2020.141570

Malsy, M., aus der Beek, T., Flörke, M., Malsy, M., aus der Beek, T., & Flörke, M. (2015). Evaluation of large-scale precipitation data sets for water resources modelling in Central Asia. *Environmental Earth Sciences*, *73*, 1–13. https://doi.org/10.1007/s12665-014-3107-y

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). https://doi.org/10.5067/342OHQM9AK6Q

Pritchard, H. D. (2019). Asia's shrinking glaciers protect large populations from drought stress. *Nature*, *569*(7758), 649–654. https://doi.org/10.1038/s41586-019-1240-1

Qi, W., Feng, L., Liu, J., & Yang, H. (2020). Snow as an important natural reservoir for runoff and soil moisture in Northeast China. *Journal of Geophysical Research: Atmospheres*, *125*(22), e2020JD033086. https://doi.org/10.1029/2020JD033086

Qi, W., Liu, J., & Chen, D. (2018). Evaluations and improvements of GLDAS2.0 and GLDAS2.1 forcing data's applicability for basin scale hydrological simulations in the Tibetan Plateau. *Journal of Geophysical Research: Atmospheres*, *123*(23), 13128–13148. https://doi.org/10.1029/2018JD029116

Qi, W., Liu, J., & Leung, F. (2019). A framework to quantify impacts of elevated CO_2 concentration, global warming and leaf area changes on seasonal variations of water resources on a river basin scale. *Journal of Hydrology*, *570*, 508–522. https://doi.org/10.1016/j.jhydrol.2019.01.015

Qi, W., Liu, J., Xia, J., & Chen, D. (2020). Divergent sensitivity of surface water and energy variables to precipitation product uncertainty in the Tibetan Plateau. *Journal of Hydrology*, *581*, 124338. https://doi.org/10.1016/j.jhydrol.2019.124338

Qi, W., Liu, J., Yang, H., Zhu, X., Tian, Y., Jiang, X., et al. (2020). Large uncertainties in runoff estimations of GLDAS versions 2.0 and 2.1 in China. *Earth and Space Science*, *7*(1), e2019EA000829. https://doi.org/10.1029/2019EA000829

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

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. https://doi.org/10.1016/j.jhydrol.2015.07.011

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

Qi, W., Zhang, C., Fu, G., Zhou, H., & Liu, J. (2016). Quantifying uncertainties in extreme flood predictions under climate change for a medium-sized basin in northeast China. *Journal of Hydrometeorology*, *17*(12), 3099–3112. https://doi.org/10.1175/jhm-d-15-0212.1

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C. J., et al. (2004). The Global Land Data Assimilation System. *Bulletin of the American Meteorological Society*, *85*(3), 381–394. https://doi.org/10.1175/bams-85-3-381

Scanlon, B. R., Zhang, Z., Save, H., Sun, A. Y., Schmied, H., van Beek, L. P. H., et al. (2018). Global models underestimate large decadal declining and rising water storage trends relative to GRACE satellite data. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(6), E1080–E1089. https://doi.org/10.1073/pnas.1704665115

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. https://doi.org/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. https://doi.org/10.1002/joc.4341

Shrestha, M., Wang, L., Koike, T., Xue, Y., & Hirabayashi, Y. (2010). Improving the snow physics of WEB-DHM and its point evaluation at the SnowMIP sites. *Hydrology and Earth System Sciences*, *14*(12), 2577–2594. https://doi.org/10.5194/hess-14-2577-2010

Simonov, E., & Egidarev, E. (2017). Intergovernmental cooperation on the Amur River basin management in the twenty-first century. *International Journal of Water Resources Development*, *34*, 771–791. https://doi.org/10.1080/07900627.2017.1344122

Stackhouse, W. P., Jr., 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.

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*, *116*(D21), D21108. https://doi.org/10.1029/2011jd015990

Wang, L., Koike, T., Yang, D. W., & Yang, K. (2009). 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. https://doi.org/10.1623/hysj.54.6.989

Wang, L., Koike, T., Yang, K., Jackson, T. J., Bindlish, R., & Yang, D. (2009). 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), D08107. https://doi.org/10.1029/2008jd010800

Wang, L., Koike, T., Yang, K., & Yeh, P.J.-F. (2009). 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. https://doi.org/10.1016/j.jhydrol.2009.08.005

Wang, L., Li, X., Chen, Y., Yang, K., Chen, D., Zhou, J., et al. (2016). Validation of the global land data assimilation system based on measurements of soil temperature profiles. *Agricultural and Forest Meteorology*, *218*, 288–297. https://doi.org/10.1016/j.agrformet.2016.01.003

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. https://doi.org/10.1175/JHM-D-15-0191.1

Xie, P., Chen, M., Yang, S., Yatagai, A., Hayasaka, T., Fukushima, Y., & Liu, C. (2007). A gauge-based analysis of daily precipitation over East Asia. *Journal of Hydrometeorology*, *8*(3), 607–626. https://doi.org/10.1175/jhm583.1

Xue, B. L., Wang, L., Yang, K., Tian, L., Qin, J., Chen, Y., et al. (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. https://doi.org/10.1002/jgrd.50696

Yang, D. (1998). *Distributed hydrological model using hillslope discretization based on catchment area function: Development and applications*. PHD. University of Tokyo.

Yang, F., Lu, H., Yang, K., He, J., Wang, W., Wright, J. S., et al. (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. https://doi.org/10.5194/hess-21-5805-2017

Yatagai, A., Kamiguchi, K., Arakawa, O., Hamada, A., Yasutomi, N., & Kitoh, A. (2012). APHRODITE: Constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges. *Bulletin of the American Meteorological Society*, *93*(9), 1401–1415. https://doi.org/10.1175/bams-d-11-00122.1

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. https://doi.org/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. https://doi.org/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. https://doi.org/10.2166/wst.2013.043