Hydrology Research

Hydrological evaluation of satellite and reanalysis-based rainfall estimates over the Upper Tekeze Basin, Ethiopia

Kidane Welde Reda [a,b,c], Xingcai Liu[a,*], Gebremedhin Gebremeskel Haile[c], Siao Sun[d] and Qiuhong Tang[a,b]

[a] Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China
[b] University of Chinese Academy of Sciences, Beijing, China
[c] Tigray Agricultural Research Institute, Mekelle, Ethiopia
[d] Key Laboratory of Regional Sustainable Development Modeling, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China
*Corresponding author. E-mail: xingcailiu@igsnrr.ac.cn

KWR, 0000-0002-8522-3857

ABSTRACT

Spatial rainfall data is an essential input to parametrically distributed hydrological models and is a main contributor to hydrological model uncertainty. Two important issues should be addressed before the use of satellite and reanalysis rainfall product at the basin level: (1) How useful are these rainfall estimates as forcing data for regional hydrological modelling? and (2) Which product should be selected at high-flow and low-flow conditions? This paper presents the hydrological performance of satellite and reanalysis rainfall products (CHIRPSv8 and EWEMBI) at three hydrological stations in the Upper Tekeze River Basin (UTB), northern Ethiopia. Results showed that the daily rainfall data from both CHIRPSv8 and EWEMBI are close to the rain gauge data, with relative errors 2.12 and 3.85%, respectively. The monthly streamflow simulated by the Soil and Water Assessment Tool (SWAT) model driven by CHIRPSv8 and EWEMBI had a Kling-Gupta Efficiency of 0.6–0.79 and 0.58–0.64, respectively. The hydrological performance during high-flow seasons is superior to low-flow seasons for both CHIRPSv8 and EWEMBI. In particular, CHIRPSv8 showed a relatively better hydrological performance than EWEMBI. This study provides insights on the usefulness of gridded rainfall products for hydrological modelling and conditions under which they can be used over the UTB and other similar basins.

Key words: hydrological modelling, rainfall products, SWAT model, satellite data, Upper Tekeze River Basin

HIGHLIGHTS

- Hydrological performances of using satellite and reanalysis rainfall products are evaluated over three gauge stations.
- CHIRPSv8 and EWEMBI rainfall products are close to the rain gauge data.
- Hydrological simulation using CHIRPSv8 and EWEMBI shows better performances in high-flow season.
- The CHIRPSv8 and EWEMBI rainfall estimates can be used for ungauged basins such as the UTB, Ethiopia.

GRAPHICAL ABSTRACT
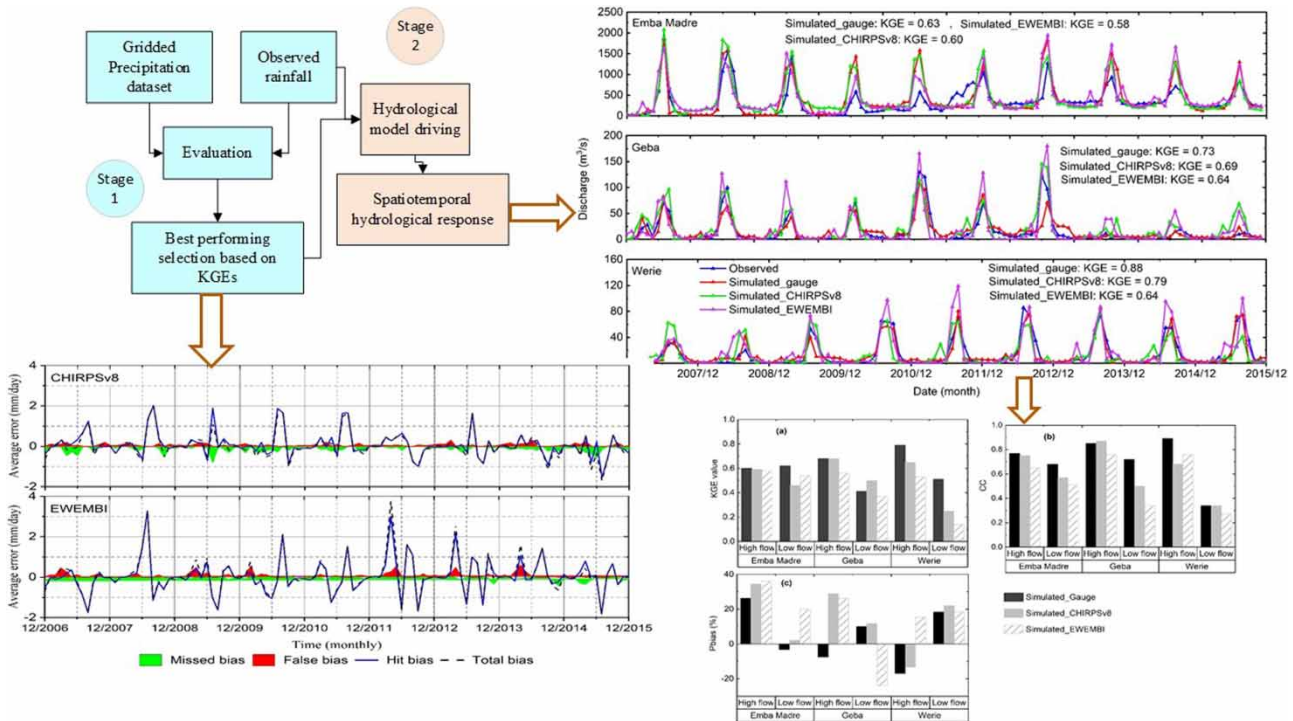
INTRODUCTION

The hydrological process affects the water supply and demand for irrigation, drinking water, hydropower production, industrial need, and ecosystems (Li *et al.* 2018). To a certain degree, this also affects the economic, agricultural, and ecological development of a particular region. Hydrological models provide simplified and obstructed representations of the hydrological processes (Singh *et al.* 2014; Ouermi *et al.* 2019). The development of a given hydrological model could be well defined in physical theory basis and empirical justification. However, a hydrological model is likely to be subject to errors and fail to produce an accurate hydrograph, particularly when input data is not sufficiently precise. Rainfall is a main input and a substantial contributor to hydrological model uncertainty (Meng *et al.* 2014; Sperna Weiland *et al.* 2015). Because the variabilities of rainfall in space and time make its measurement difficult, an accurate estimate of spatial rainfall over a region of interest is critical for hydrological modelling.

The primary means of measuring point rainfall is gauge stations. The areal precipitation can be estimated by means of different interpolation methods (e.g., Thiessen polygons, geostatistical Kriging, and inverse distance weighting) and be used as input to hydrological models (Zhao *et al.* 2015; Li *et al.* 2018). However, the reliability of the areal rainfall estimate is highly dependent on gauge density. In many regions, rain gauge observations are sparse and unevenly distributed due to topographic limitations and economic problems, especially in developing countries (Zeng *et al.* 2018). Such problems are more pronounced in Ethiopia (Bayissa *et al.* 2017; Gebremicael *et al.* 2019) where it has a sparse and low-quality gauge network.

To address the above issue, a number of satellite and reanalysis-based rainfall products with high spatial and temporal resolutions have been developed in recent years at continental and global levels. The rainfall products differ from each other in terms of data source (satellite, gauge, radar, analysis, and reanalysis could be a combination of some of these), spatial resolution (0.05°–2.5°), temporal resolution (30 min to monthly), design objective (instantaneous accuracy or temporal homogeneity), and spatial and temporal coverages (continental to fully global and nearly 1 year to 115 years) (Beck *et al.* 2017a, 2017b). These products have been used as meteorological forcing in hydrological process modelling, climatic research, and water resource management (Kneis *et al.* 2014). Easy access and the ability to mimic point-scale measurements and their spatial grid and constant time step have favoured the widespread use of these products (Raimonet *et al.* 2017). Hence, these

satellite and reanalysis rainfall estimates provide alternative sources of input data for hydrological modelling in regions where rain gauge observations are sparse. Xue et al. (2013) validate the application of Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) products (3B42V6 and 3B42V7 algorisms) for an ungauged catchment in Bhutan. The recent improvement in spatial and temporal resolutions has largely increased their applicability to large-scale distributed hydrological models at both global and regional conditions. For example, Beck et al. (2017a, 2017b) evaluated global runoff using WATCH Forcing Data ERA-Interim (WFDEI) meteorological dataset-driven hydrological models. Others used for global draught analysis (Herrera-Estrada et al. 2017; Wang et al. 2021). At the regional level, Schuol et al. (2008) estimated fresh water availability in West African sub-continent, Haile et al. (2020) identified spatio-temporal patterns of drought in east Africa, and Faramarzi et al. (2013) modelled fresh water availability in response to climate change in Africa using satellite and reanalysis rainfall products.

However, the performance of different precipitation products varies across regions. For instance, the magnitude of 30 different annual precipitation estimates over global land deviated by as much as 300 mm/yr among the products (Sun et al. 2018). Bayissa et al. (2017) also showed the performance of five satellite rainfall products varied across the upper Blue Nile. The adequacy of using these rainfall products to drive hydrological models at a specific basin level is as yet less well understood, which needs a thorough evaluation (Hu et al. 2017; Li et al. 2018; Awange et al. 2019).

Many recent studies have evaluated the adequacy of using satellite and reanalysis rainfall products in driving hydrological models for water balance analysis over different regions of the globe (Tan et al. 2017; Roy et al. 2018; Awange et al. 2019; Azimi et al. 2020). The results showed that hydrological performances driven by satellite and reanalysis rainfall products vary with region, season, and elevation. There are no regions on which all precipitation datasets show a consistent and statistically significant trend in precipitation seasonality (Tan et al. 2020). These inconsistent changes in precipitation seasonality within various precipitation datasets imply the importance of choosing the dataset when studying changes in regional precipitation seasonality and its implications in hydrological conditions. In general, one rainfall product that performs reasonably well in driving the hydrological process in one region may not result in adequate simulation over other regions. In addition, the source of uncertainty in hydrological modelling is not only limited to the forcing rainfall data but also the models, i.e., model structure and model parameters.

Because of these limitations in rainfall products and hydrological models, their use in hydrological regime simulation should be evaluated before any application in a particular region. Most previous studies focused on validating input rainfall data and comparison of models. In the Upper Tekeze River Basin (UTB), Northern Ethiopia, the applicability of rainfall products to drive a hydrological model is not sufficiently tested due to lack of data. Validating rainfall products are either through ground observations or model-based applications. Some studies have been conducted using the first approach over the Nile river basin catchments, Ethiopia (Bayissa et al. 2017; Lakew et al. 2017). The second approach is based on hydrological modelling performance within target applications. For instance, satellite and reanalysis rainfall products can be evaluated based on their ability in reproducing observed hydrograph referred to as hydrological evaluation. Only a few studies using the first approach are available (e.g., Gebremicael et al. 2019), but rainfall products have not yet been assessed using the second approach in the UTB.

In view of the above considerations, this study aimed at assessing the suitability and adequacy of one satellite-based rainfall product, namely, Climate Hazards Group InfraRed Precipitation with Stations (CHIRPSv8), and one reanalysis product called the EartH2Observe, WFDEI and ERA-Interim reanalysis data Merged and Bias-corrected for the Inter-Sectoral Impact Model Intercomparison Project (EWEMBI) rainfall product (Lange 2016) for simulating regional water balance over the UTB. CHIRPSv8 and EWEMBI products were selected based on our previous study (Reda et al. 2021). These two products showed good performances out of nine gridded global rainfall products in representing daily and monthly precipitation under the UTB. The CHIPRSv8 is the latest high-resolution rainfall dataset that combines satellite and gauge observations, which could be promising data for regional hydrological applications. The EWEMBI is a new bias-corrected global dataset at the half-degree resolution and has been used for both global and regional hydrological evaluations (Frieler et al. 2017; Chawanda et al. 2020). In this study, a physically based hydrological model, the Soil and Water Assessment Tool (SWAT) model, was used to evaluate the hydrological performance of the two global datasets applied in the UTB. The SWAT is a popular model for many studies at sub-basins under the Nile river where it is overall poorly gauged (Griensven et al. 2012; Akoko et al. 2021). The evaluation would also provide useful information for hydrological simulations using the global datasets in other regions where limited meteorological observations are available.

STUDY AREA AND DATA

Study area

With an area of 45,695 km^2, the UTB is the northern part of the Nile river basin in Ethiopia (Figure 1). It is located at 37.5°–39.8°E longitude and 11.5°–14.3°N latitude. The Tekeze River is a major tributary of the Nile River and is exploited for hydropower production and irrigation in Ethiopia (Zenebe 2009). The mean sea level elevation varies from 830 to 4,529 m, with a complex topography consisting of mountains, highlands, and lowlands with terrains of gentle slopes. The mean annual flow of the UTB at the Emba Madre gauge station (located at 12.6°N latitude and 39.19°E longitude) is around 6.9 billion m^3/year.

Generally, the basin covers semi-arid (North part) and semi-humid (South part) regions. The mean annual temperature is 17 °C, and the annual rainfall is ~800 mm. Its climate is characterized by a short rainy season (June–September) concentrating 60–70% of annual rainfall and a long dry season (October–May). The spatial variation of rainfall over the UTB is high (Belete 2007; Gebremicael et al. 2017). Farmland is the dominant land use in the UTB. It covers more than 70% of the land within the basin. The farmland is dominated by rainfed agriculture, with main crops such as 'Teff', wheat, barley, maize, sorghum, and pulses. The rest of the basin is mainly grassland and residential areas. The main soils of the Tekeze basin as per the FAO classification are Eutric Cambisols and Calcric Cambisols. Eutric Vertisols are mainly found in the highland and lowland areas of the watershed, and Eutric Leptosols are in the central part of the basin. The dominant soil texture groups are sandy and clay loam, which in total account for 70% area coverage of the basin (Gebremeskel et al. 2018).

Gauge rainfall and streamflow data

Daily rainfall, maximum and minimum temperature, humidity, wind speed, and solar radiation data at the 21 rain gauge stations (Figure 1) were collected for the period of 2006–2015 from the National Meteorological Agency of Ethiopia (NMAE). Homogenization and outlier processing were implemented in a quality control procedure. The homogeneity of monthly rainfall was tested using the standard normal homogeneity test, which compares the observations at each rain gauge station with the average of observations at the four nearest gauge stations (see section 1 in Supplementary Information for details). Outliers were tested through comparing with the neighbouring four gauge stations to cross-check if similar extreme values are observed. No homogenization and outlier problems were detected at the rain gauge stations. Hence, the ground rainfall observation datasets were found to be reliable for analyses following screening criteria. Finally, a spatial rainfall distribution from the rain gauge data was estimated using the ordinary Kriging interpolation technique, which is regarded as an accurate interpolation technique, that produces more reliable estimates than other interpolation techniques for daily average rainfall data (Guyot et al. 2015; Yue et al. 2016). All the rain gauge stations in our study area are not part of the Global Precipitation Climatology Centre (GPCC) network used for the calibration of satellite products.

Furthermore, daily measured flow data at the Emba Madre, Geba, and Werie gauging stations during the period 2006–2015 were collected from the hydrology department of the ministry of water, energy and irrigation of Ethiopia. The drainage areas covered by the Emba Madre, Geba, and Werie gauge stations are 45,694, 4,342, and 1,770 km^2, respectively.

CHIRPSv8 and EWEMBI rainfall products

CHIRPS is a quasi-global rainfall dataset developed by the Climate Hazards Group and the US Geological Survey (USGS) at the University of California at Santa Barbara (Funk et al. 2015). CHIRPS has 0.05° spatial and daily temporal resolutions with a 50°S–50°N global coverage. It is produced from the global telecommunications system (GTS) through a combination of gauge observations, satellite observations, and global climatology (Funk et al. 2014; Gebremicael et al. 2019). This product has been used in east Africa in many studies and commonly showed good performance (e.g., Bayissa et al. 2017; Dinku et al. 2018; Lemma et al. 2019).

EWEMBI is a newly compiled reference dataset (Lange 2018). EWEMBI has a daily temporal frequency and 0.5° spatial resolution with a global coverage. In this study, the EWEMBI dataset was downscaled to 0.05° spatial resolution through the statistical downscaling method. The elevation–precipitation relationship, which better predicts the precipitation distribution in the areas with complex terrain, was used (Jia et al. 2011; Chen et al. 2020). The coefficient of determination (R^2) for the linear regression analysis result between the daily EWEMBI precipitation and elevation, from 2006 to 2015, ranges from 0.053 to 0.79, with an average value of 0.672. The EWEMBI dataset is produced from the WATCH forcing data methodology applied to ERA-Interim reanalysis data (WFDEI) and Global Precipitation Climatology Centre (GPCC v5) over the land surface, while EartH2Observe forcing data (E2OBS) over the ocean (Lange 2016). GPCCv5 and v6 monthly precipitation totals were used for bias adjustment. The WFDEI precipitation products included in E2OBS were bias-adjusted with Climatic
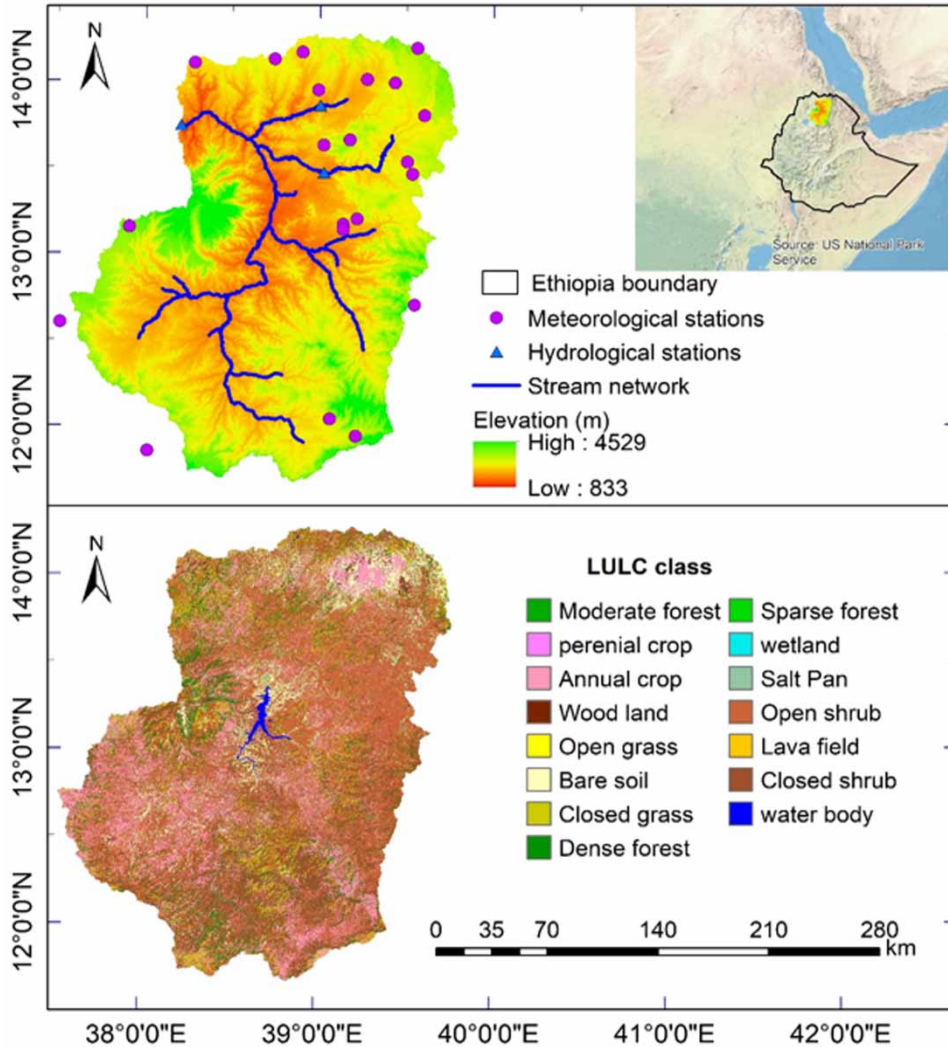
Figure 1 | The geographical location, terrain elevation, land use classes, rain gauge, and hydrological stations of the Upper Tekeze Basin.

Research Unit Time Series (CRU TS3.101/TS3.21) monthly precipitation. The detail about EWEMBI and its data sources can also be found in Dee *et al.* (2011) and Weedon *et al.* (2014). The EWEMBI dataset was compiled to support bias correction of climate input data for the ISIMIP2b.

METHODOLOGY

Evaluation of CHIRPSv8 and EWEMBI rainfall

The total rainfall bias of the CHIRPSv8 and EWEMBI rainfall products was separated into three independent error components following the methodology used by Li *et al.* (2018) and Habib *et al.* (2009). Specifically, (1) the bias due to missed rainfall (missed bias), (2) the bias due to false detection (false bias), and (3) the bias of successful detection (hit bias). The missed bias is that the rainfall products fail to detect the rainfall events indicated by the reference dataset (gauge-based rainfall event) and its value is negative. On the other hand, false bias is that rainfall occurrence is detected by the satellite and reanalysis products but the gauge-based observations indicate no rainfall event. The value of a false bias is positive. In the hit bias case, both the reference and the satellite and reanalysis data detect a rainfall event but with different rainfall amounts. The value of hit bias could be positive, negative, or zero. The monthly average error of the two rainfall products was analysed using the temporal error trend analysis for each bias component.

In addition, daily average (mm/day), standard deviation (std) (mm), maximum daily rainfall (mm/1day), and maximum 5-day rainfall (mm/5day) from CHIRPSv8, EWEMBI, and gauge datasets were compared with an annual basis at both the grid and basin levels. Finally, the daily rainfall distribution and their corresponding contributions to the total rainfall as a function of rainfall intensity bins for each dataset was analysed.

Hydrological model

The applicability of the global rainfall data in driving the hydrological model was tested using the 2012 version of the SWAT model. The SWAT is a temporally continuous and semi-distributed model developed to predict the impact of land management practices in watersheds and river basins on runoff, sediment, and agricultural chemical yields (Neitsch *et al.* 2012). The main SWAT model components include weather, hydrology, soil temperature, plant growth, nutrients, pesticides, land management, bacteria, and pathogens (Francesconi *et al.* 2016). The model is often used to estimate runoff and its pollutant concentrations at various timescales, and it is widely used worldwide. The SWAT has been applied in different watersheds and river basins of Ethiopia and demonstrated satisfactory results (Gebremicael *et al.* 2013; Welde & Gebremariam 2017; Nadew 2018).

The SWAT simulates the hydrological process of a river basin through disaggregating it into sub-basins and hydrological response units (HRUs). Spatial inputs required for delineating sub-basins and HRUs include a land use land cover (LULC) map, soil map, and digital elevation model (DEM). A 30 m × 30 m DEM was used in this study. The LULC (Figure 1) and soil maps were collected from the Ministry of Agriculture and natural resource management of Ethiopia.

Hydrological model and simulation

To evaluate the suitability of different rainfall data for hydrological simulation, experiments based on the SWAT model were conducted using data from rain gauge stations, CHIRPSv8, and EWEMBI datasets in the UTB, while the other forcings such as temperature, relative humidity, wind speed, and solar radiation were from the stations and kept the same for all the modelling experiments. The SWAT model contains parameters that are not physically measurable and need to be obtained using conceptualization and calibration through adapting the model output to the observed data. Before calibration, sensitive parameters were identified and prioritized using the procedure of SWAT sensitivity analysis. Then, these sensitive parameters were calibrated using sequential uncertainty fitting (SUFI-2) algorithm capabilities of SWAT CUP. SUFI-2 is a commonly used tool for calibration and uncertainty analysis of the SWAT model (Wu & Chen 2015). Even though SWAT CUP has different calibration methods like particle swarm optimization (PSO), generalized likelihood uncertainty analysis estimation (GLUE), parameter solution (ParaSol), and Markov chain Monte Carlo (MCMC), previous studies (Wu & Chen 2015; Abbaspour *et al.* 2017; Mengistu *et al.* 2019) have stated that the SUFI-2 method can provide more reasonable and balanced predictions than the others.

Because different rainfall datasets result in different calibration results, the following three scenarios were applied in evaluating the hydrological processes. Scenario I, the daily gauge rainfall data, was first used to calibrate the parameters in the SWAT. Then, the daily CHIRPSv8 and EWEMBI rainfall products were used to run the model using the calibrated parameters.

The SWAT model uses point rainfall data from the gauges and changes them into areal rainfall using ordinary kriging or Thiessen polygon. The kriging method was used to obtain areal rainfall from station data in the SWAT model in this study, which was reported to be more efficient than the Thiessen polygon method in mountainous areas (Mair & Fares 2011). For CHIRPSv8 and EWEMBI, grid-level rainfall estimates were used to better capture the spatial varied information. Finally, the simulated runoff from the three model runs (using the rain gauge, CHIRPSv8, and EWEMBI) was compared with the streamflow observations at the three stations. In scenario II, the daily CHIRPSv8 rainfall product was utilized to calibrate the SWAT model, and then the EWEMBI and rain gauge rainfall products were used subsequently to drive the model. In scenario III, the EWEMBI rainfall product was used first to optimize the model parameters in calibration, and the CHIRPSv8 and rain gauge rainfall data were used to drive the SWAT model. In all the scenarios, the SWAT model is calibrated using the Nash–Sutcliffe coefficient of efficiency (NSCE) index as the objective function. The runoff simulations were evaluated by comparison with measurements at three hydrological stations at daily and monthly time scale, during full time series, and high-flow and low-flow seasons.

Model simulation uncertainties were estimated using the 95% prediction uncertainty (95PPU) capabilities of SUFI-2. In SUFI-2, uncertainty in parameters, expressed as ranges (uniform distributions), considers all uncertainty sources including model structure, parameters, driving variables (e.g., temperature and rainfall), and the measured data. Propagation of the uncertainties in the parameters leads to uncertainties in the model output simulations, which are expressed as the 95%

probability distributions. These are estimated at the 97.5 and 2.5% levels of the cumulative distribution of an output variable using Latin hypercube sampling, referred as the 95PPU (Abbaspour *et al.* 2007). These 95PPUs are the model outputs resulting from a stochastic calibration approach, indicating that there is no single signal representing model output, but an envelope of solutions represented by the 95PPU, produced by certain parameter ranges. Two statistics (i.e., *P*-factor and *R*-factor) were used to quantify the fit between simulation result (expressed as 95PPU) and observation (expressed as a single signal). *P*-factor is the percentage of observed data enveloped by the modelling result, the 95PPU. *R*-factor is the thickness of the 95PPU envelope.

Model performance evaluation

To evaluate the hydrological model performance using the gauge, CHRPSv8, and EWEMBI rainfall products, four commonly used statistical criteria were applied, including the correlation coefficient (CC), NSCE, percent of bias (Pbias), and Modified Kling-Gupta Efficiency (KGE). The mathematical definitions of these metrics are listed in Table 1. The evaluation was for low-flow, peak-flow, and all time-series flows daily and monthly time scales. NSCE is a normalized statistical metric and evaluates the relative magnitude of the residual variance compared with the observed variance and demonstrates how well observed vs. simulated data plots fit the 1:1 line through the entire period. A match of timing and shape of the hydrograph is reflected by the CC. Pbias assesses the systematic error that indicates the average tendency of the simulated data deviates from their observed counterparts. Finally, the model performance on hydrological simulation for the CHIRPSv8, EWEMBI and gauge rainfall was assessed using the KGE. The KGE is a recently developed performance indicator, which takes into account equal weighting of three metrics, such as the coefficient of determination (R^2), the bias ratio (β), and the variability (γ), between observed (obs) and simulated (sim) discharges (Kling *et al.* 2012). Mathematically, β and γ are defined as follows:

$$\beta = \frac{\mu_{sim}}{\mu_{obs}} \tag{1}$$

$$\gamma = \frac{\delta_{sim}/\mu_{sim}}{\delta_{obs}/\mu_{obs}} = \frac{CV_{sim}}{CV_{obs}} \tag{2}$$

where μ and σ are the mean and standard deviation of discharge (m^3/s), respectively. CV is the coefficient of variation. KGE is dimensionless and its optimal value is unity. According to Thiemig *et al.* (2013), hydrological performance can be categorized using the KGE as follows: very poor (KGE $\leq$ 0), poor (0 < KGE < 0.5), intermediate (0.5 $\leq$ KGE < 0.75), and good (KGE $\geq$ 0.75).

Table 1 | Statistical indexes and their mathematical definition

Metric Name	Function
Pearson's CC	$CC = \dfrac{\sum\limits_{i=1}^{n}(Q_{i,obs} - \bar{Q}_{obs})(Q_{i,sim} - \bar{Q}_{sim})}{\sqrt{\sum\limits_{1}^{n}(Q_{i,obs} - \bar{Q}_{obs})^2}\sqrt{\sum\limits_{i}^{n}(Q_{i,sim} - \bar{Q}_{sim})^2}}$
Percentage of bias	$Pbias = \left(\dfrac{\sum\limits_{i=1}^{n}Q_{i,sim} - \sum\limits_{i=1}^{n}Q_{i,obs}}{\sum\limits_{i=1}^{n}Q_{i,obs}}\right)*100$
Nash–Sutcliffe coefficient of efficiency	$NSCE = \dfrac{\sum\limits_{i=1}^{n}(Q_{i,obs} - \bar{Q}_{obs})^2 - \sum\limits_{i=1}^{n}(Q_{i,obs} - Q_{i,sim})^2}{\sum\limits_{i=1}^{n}(Q_{i,obs} - \bar{Q}_{obs})^2}$
Modified Kling-Gupta Efficiency	$KGE = 1 - \sqrt{(R^2 - 1)^2 + (\beta - 1)^2 + (\gamma - 1)^2}$

Q_{obs} and Q_{sim} are the observed and simulated discharges for the *i*th day, respectively, and $\bar{Q}_{obs}$ and $\bar{Q}_{sim}$ are the average values of the observed and simulated discharges for the entire period under consideration.

Water balance analysis

In addition to the flow hydrograph comparisons, the water balance component result can be another important indicator for evaluating the different rainfall products' validity in driving the hydrological model. The SWAT model divides precipitation into evapotranspiration, surface runoff, groundwater recharge, and baseflow. Hence, the comparison was also conducted for these components of these simulation results using the CHIRPSv8, EWEMBI, and gauge rainfall data.

RESULTS

Comparison between gridded rainfall and rain gauge data

Table 2 summarizes statistics of the CHIRPSv8, EWEMBI, and areal-averaged gauge rainfall products. The estimates of areal-averaged rainfall during 2006–2015 were 1.69–2.47 mm day^{-1} (rain gauges), 1.79–2.5 mm day^{-1} (CHIRPSv8), and 1.72–2.93 mm day^{-1} (EWEMBI). The maximum daily rainfall occurred in 2006 for the rain gauge rainfall and the CHIRPSv8 data, whereas it occurred in 2012 according to the EWEMBI rainfall data. The annual average rain gauge rainfall data indicates that 2009 was the driest year, but CHIRPSv8 and EWEMBI overestimate rainfall in 2009 by 6.8 and 6.5%, respectively. The bias of EWEMBI (−4.4 to 31.1%) was larger than that of CHIRPSv8 (−15 to 9.8%) compared with the rain gauge data. The absolute averages of positive bias were higher than those of the negative bias in both CHIRPSv8 (5.44 vs. −2.63) and EWEMBI (11.38 vs. −2.45). The standard deviations (std deviations) of the rainfall are higher than the average rainfall according to the three rainfall datasets.

The observed maximum daily rainfall (mm day^{-1}) and maximum 5-day rainfall (mm 5-day^{-1}) show large variability among different rainfall datasets (Table 2). Particularly, the maximum daily rainfall was 18.72–42.11 mm (gauge data), 19.51–49.66 mm (CHIRPSv8), and 17.72–40.13 mm (EWEMBI). The maximum 5-day rainfall was 47.13–92.11 mm (gauge data), 47.3–105.57 mm (CHIRPSv8), and 50.75–82.06 mm (EWEMBI). The capability of CHIRPSv8 and EWENBI to detect particular heavy rain events is weakened. This shows that the CHIRPSv8 and EWEMBI rainfall products are not able to capture the occurrence of maximum daily and maximum 5-day rainfall accurately. This result is consistent with Rajulapati et al. (2020), which indicates that global precipitation products generally cannot provide a consistent representation of the magnitude and frequency of extreme events. The results indicated that while the range of the data during the time period of interest was similar, the days of the rainy season months (July–September) with the maximum daily and maximum 5-day rainfall occurrences were different. Hence, the differences between the maximum daily and maximum 5-day rainfall from different data sources were large in these months.

Temporal and spatial distributions

Figure 2 shows the trends of average monthly time series of the CHIRPSv8 and EWEMBI data error components (missed, false, hit, and total bias). The CHIRPSv8 product displays a lower total bias range (−1.69, 2.02) as compared with the EWEMBI's product (−1.82, 3.75). In addition, the missed bias was higher in CHIRPSv8 than EWEMBI, but the reverse was for false

Table 2 | Comparison of statistical indexes between areal-averaged CHIRPSv8 and EWEMBI vs. rain gauge rainfall

| Year	Areal average (mm day^{-1}) (bias, %)			Std deviation (mm day^{-1})			Maximum daily rainfall (mm day^{-1})			Maximum 5-day rainfall (mm 5-day^{-1})		
	Gauge	CHIRPSv8	EWEMBI	Gauge	CHIRPSv8	EWEMBI	Gauge	CHIRPSv8	EWEMBI	Gauge	CHIRPSv8	EWEMBI
2006	2.47	2.5 (−3.6)	2.54 (2.9)	4.44	5.3	4.22	42.11	49.66	33.53	79.78	77.85	71.44
2007	2.32	2.43 (6.4)	2.18 (−1.6)	4.41	5.1	3.76	39.9	34.9	28.38	78.2	84.43	64.94
2008	1.95	2.11 (8.7)	2.18 (11.8)	3.07	4.3	3.83	15.46	36.97	39.29	47.13	80.31	82.06
2009	1.69	1.84 (6.8)	1.72 (6.5)	3.45	4.6	2.87	23.84	27.4	36.12	47.16	75.76	50.75
2010	2.22	2.26 (0.9)	2.2 (−1)	3.67	4.5	3.75	22.96	27.26	27.76	60.32	86.51	62.83
2011	1.97	2.21 (9.8)	1.92 (−2.8)	3.67	4.8	3.25	36.47	39.71	17.72	92.11	105.57	56.06
2012	2.23	2.21 (0.02)	2.93 (31.1)	3.98	4.4	5.52	26.01	31.99	40.13	67.01	64.81	81.12
2013	2.00	2.07 (−1.5)	2.24 (11.8)	3.57	4.1	4.51	29.37	30.27	39.77	50.25	67.57	80.82
2014	2.28	2.25 (−3.9)	2.49 (9.2)	3.36	4	3.87	18.72	28	17.73	58.36	66.71	64.98
2015	2.05	1.79 (−1.5)	1.97 (−4.4)	3.08	3.1	3.58	20.15	19.51	21.53	53.83	47.3	63.6
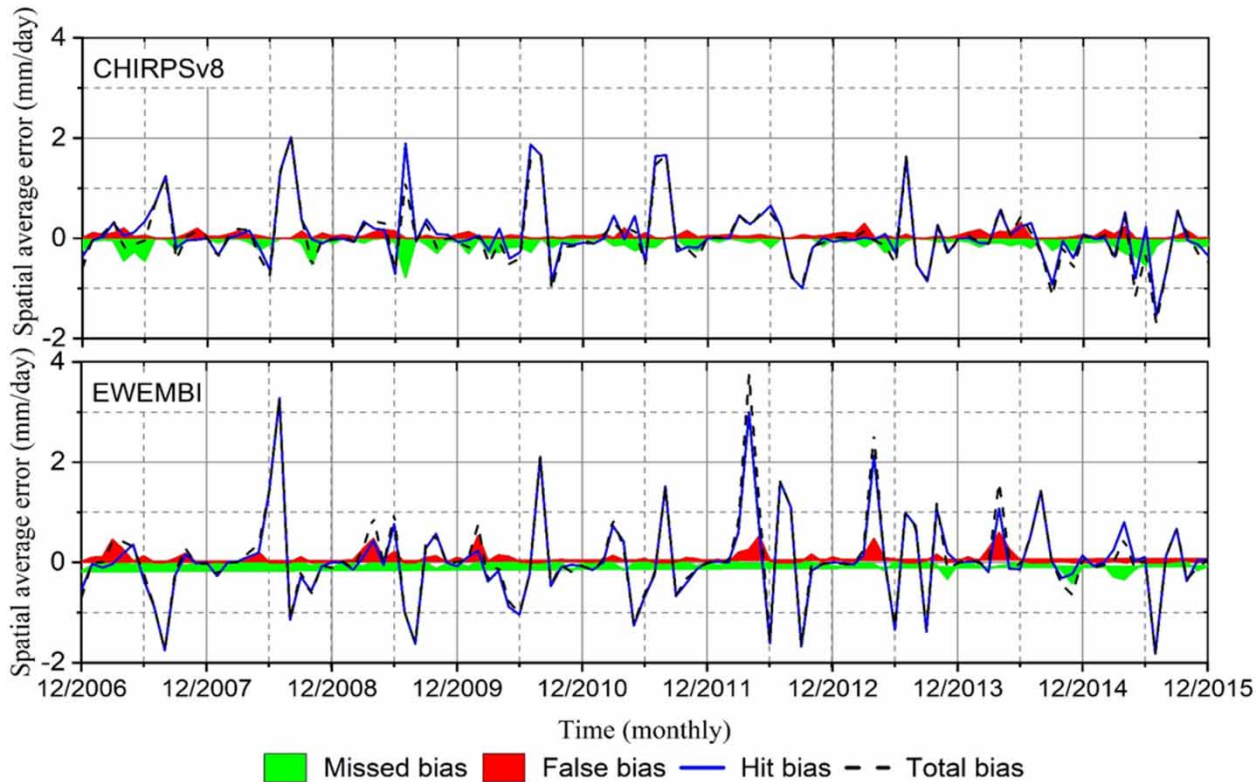
Figure 2 | Time series of error components for CHIRPSv8 and EWEMBI rainfall products over the UTB.

bias. For both the rainfall products, the error components are dominated by hit bias which ranges from −1.54 to 2.02 for CHIRPS and from −1.82 to 3.28 for EWEMBI. CHIRPSv8 and EWEMBI show similar error patterns and both display high false bias in the dry season from January to March and hit bias in the rainy season from June to September.

Figure 3 presents an intercomparison for the CHIRPSv8 and EWEMBI daily rainfall against the gauge observation under the rainy and dry seasons using the CC. The spatial distribution of CC showed that both CHIRPSv8 and EWMBI show relatively better agreement with the gauge rainfall for the rainy season than the dry season period.

The CHIRPSv8 and EWEMBI daily rainfall data and gauge rainfall data were further compared at the grid and basin scales. Figure 4 displays the scatter plots of the CHIRPSv8 and EWEMBI data vs. rain gauge data at the two spatial scales. For the basin-scale comparison, the average value of all the gauge stations was compared with their respective average of CHIRPSv8 and EWEMI data at the grid cells where the gauge stations are located (Figure 4(a) and 4(b)). For the grid-scale comparison, the values of 21 gauge stations were compared with the grid value of CHIRPSv8 and EWEMBI data at each station at a daily time scale (Figure 4(c) and 4(d)). CHIRPSv8 showed better agreement with the gauge data compared with EWEMBI at both the grid and basin scales with the CC values of 0.58 and 0.73, respectively. Similarly, a moderate agreement is found between EWEMBI and gauge rainfall data at grid and basin levels with CC values of 0.53 and 0.62, respectively. The CHIRPSv8 and EWEMBI products slightly overestimate rainfall of the basin with bias of 0.6 and 5.5%, respectively (Figure 4). Since the accuracy of satellite and reanalysis-based rainfall products often improves with increased time-scale aggregation (Lo Conti *et al.* 2014; Li *et al.* 2018), the CHIRPSv8 and EWEMBI performances are expected to be improved at a monthly scale. Hence, the two products could be utilized for seasonal and annual water balance simulation at the UTB.

Rainfall intensity distribution

Figure 5 shows the daily rainfall intensity distributions for different rainfall sources and their relative contributions to the total rainfall in each year during the period 2006–2015. The rainfall intensity class [0, 1 mm) was the most frequent for the rain gauge rainfall (48.5–63.3%), CHIRPSv8 (53.7–64.4%), and EWEMBI (51.8–60.5%). The second most frequent daily rainfall
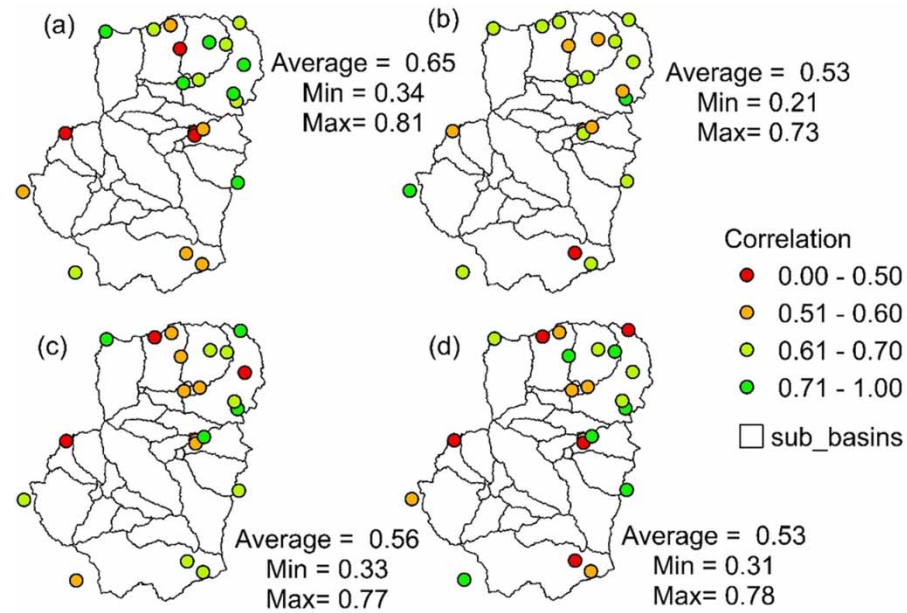
Figure 3 | Spatial distribution of correlation coefficient for the seasonal daily rainfall over the UTB. (a) CHIRPSv8 for the rainy season (June–September); (b) CHIRPSv8 for the dry season (October–May); (c) EWEMBI for the rainy season; (d) EWEMBI for the dry season. The colour of the circles indicates the values of Pearson's CC between gauge data and the gridded data.

class was [1,5 mm) for all the three rainfall datasets, occurring in about 21–34.5% (gauge rainfall), 21.4–29.9% (CHIRPSv8), and 24.7–30.4% (EWEMBI) of the total number of days. It indicates that more days of slight rain or no rain [0, 1 mm) were recorded in all the CHIRPSv8, EWEMBI, and rain gauge than days of heavier rainfall.
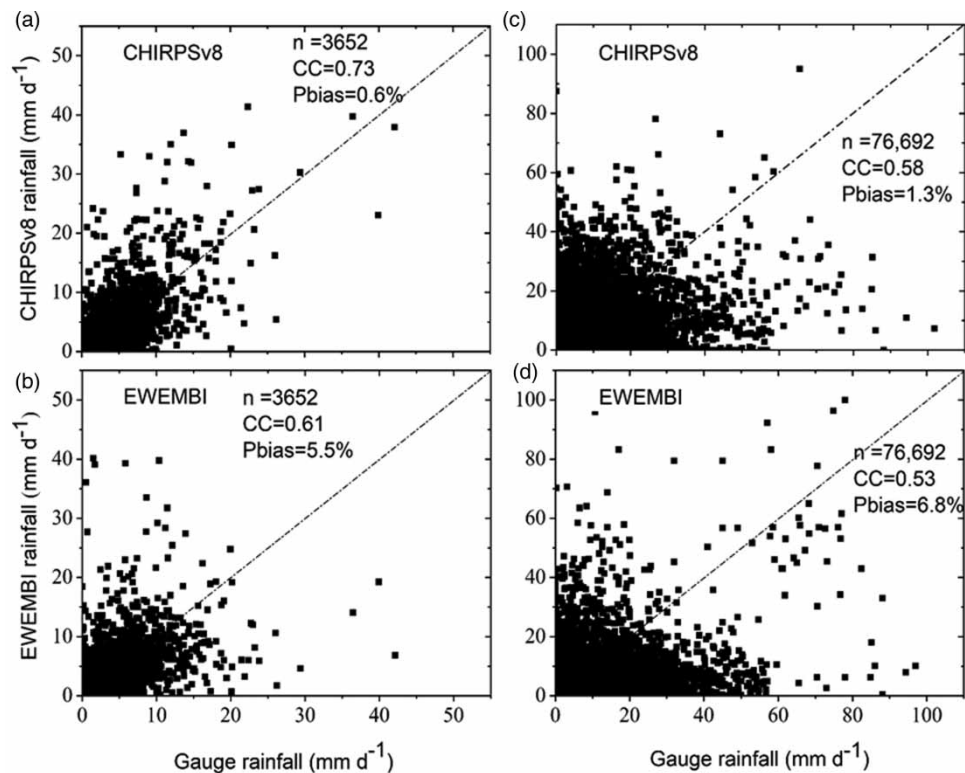

Figure 4 | Scatterplots of daily rainfall from the CHIRPSv8 and EWEMBI vs. the rain gauge data: at the basin level (a and b) and at the grid scales (c and d).
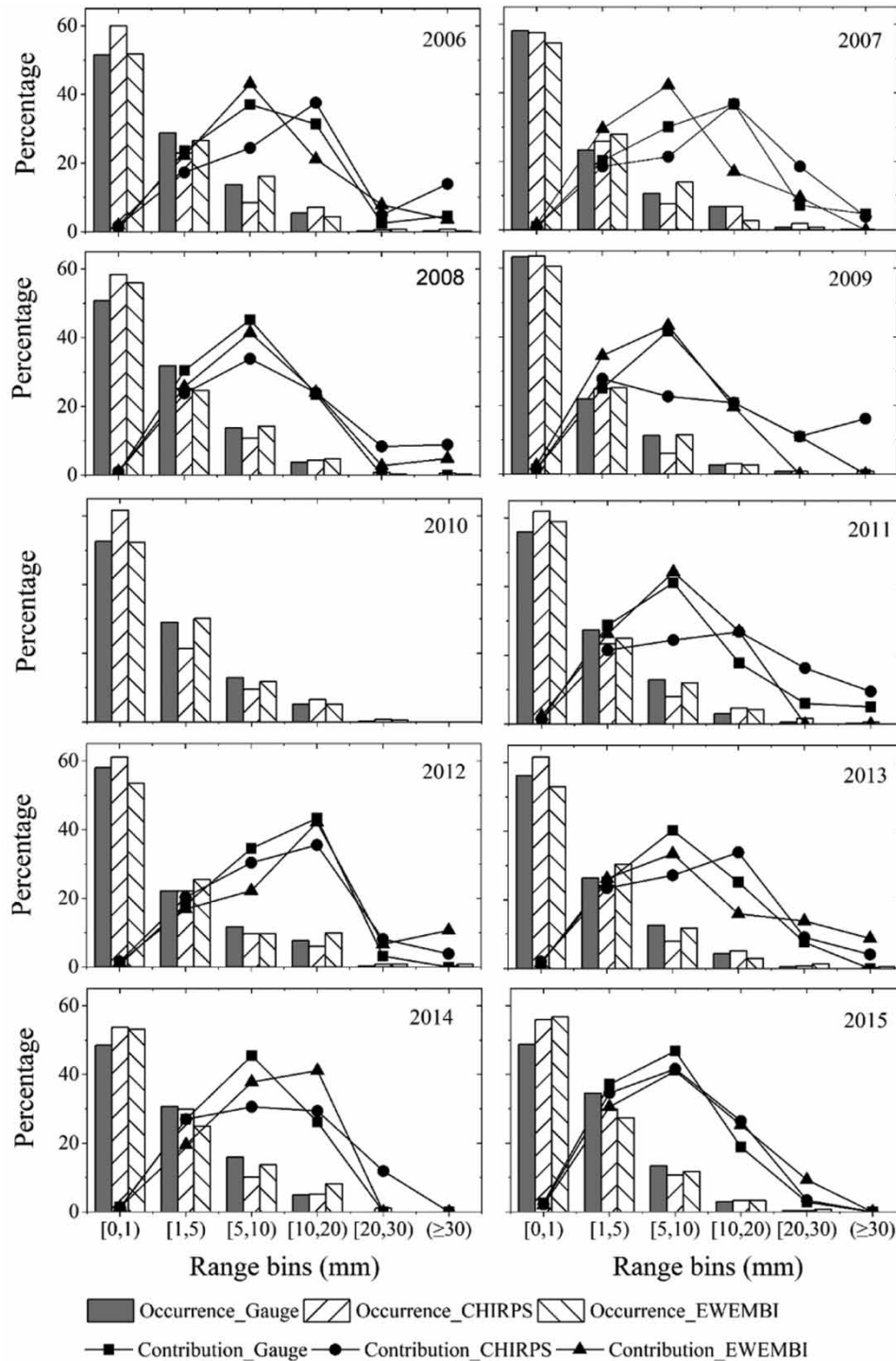
Figure 5 | Distribution and relative contribution to the total rainfall for different daily rainfall bins during different years of the period 2006–2015.

Even though the frequencies of daily rainfall within [0, 1 mm) and [1,5 mm) were high in all the rainfall datasets, their contributions to the total rainfall amount were trivial compared with the other rainfall bins (Figure 5). The [5,10 mm) daily rainfall had the largest contributions to the total rainfall, which were 30.1–45.5%, 21.4–38.4%, and 22.3–44.3% for the gauge, CHIRPSv8, and EWEMBI rainfall, respectively. The highest rainfall class 'rainfall ≥30 mm/day' contributed an

average (range parentheses) of 1.5% (0, 4.7%), 6.1% (0, 16.2%), and 2.8% (0, 10.7%) of the total rainfall in the case of gauge, CHIRPSv8, and EWEMBI rainfall data, respectively. Both the occurrence and contribution of this intensive rainfall class for the CHIRPSv8 and EWEBI datasets were larger than the gauge data. In general, based on the monthly time-scale error components and the temporal and spatial analyses of the products, the CHIRPSv8 and EWEMBI rainfall estimates were able to capture the rainfall structure compared with the rain gauge rainfall data over the UTB.

EVALUATION OF HYDROLOGICAL PERFORMANCE

Calibration of model parameters

Many parameters of the SWAT model are physical in nature and should, in principle, be estimated from field surveyed data, such as land cover, soil, vegetation coverage, and channel characteristics. However, in many large basins, particularly in Africa (Thiemig *et al.* 2013), it is difficult to estimate model parameters from field data. Hence, model performances need to be optimized to determine their parameters through calibration. Prior to model calibration, a sensitivity analysis was implemented to identify the most sensitive parameter using SWAT sensitivity analysis capabilities, and the most sensitive parameters are listed in Table 3. These parameters include the surface runoff, groundwater, evaporation soil, precipitation, and channel components of the basin hydrological process. In this study, SUFI-2, which is commonly used as a tool for the combined calibration and uncertainty analysis (Yang *et al.* 2008; Zhou *et al.* 2014) of the SWAT model, was used. Fifteen flow parameters of the hydrological process were calibrated for the three scenarios, and their numerical parameter optimization results are shown in Table 3. The calibration result indicates that the optimum parameter values vary across the scenarios. This variation directly and indirectly contributes to the performance of the rainfall data in driving the SWAT model for simulating the hydrological process. This implies that the model parameters were sensitive to the input rainfall data (CHIRPSv8, EWEMI, and rain gauge data).

Table 3 | Calibrated sensitive model parameters and their optimal values for the three scenarios (v and r are parameter change qualifiers representing change existing with 'α' and multiply existing with '$1 + \alpha$', respectively)

Parameter	Description	Process	Lower bound	Upper bound	Calibrated value		
					Scenario I	Scenario II	Scenario III
v_PLAPS.sub	Precipitation lapse rate (mm H_2O/km)	Precipitation	0	500	98.2	76.5	58.3
v_TLAPS.sub	Temperature lapse rate (°C/km)	Temperature	−8	−4	−0.05	−0.05	−0.05
r_CN2.mgt	Initial Soil Conservation Service (SCS) runoff curve number for moisture condition II	Runoff	−0.4	0.4	0.25	0.19	0.06
v_Alpha_Bf.gw	Baseflow alpha factor–baseflow recession constant	Runoff	0	1	0.56	0.53	0.54
v_Gwqmn.gw	Threshold depth of water required for return flow to occur	Groundwater	0	2,000	1,050	982.7	800.23
r_Sol_Z.sol	Soil depth (mm)	Soil	−0.3	0.3	−0.23	−0.15	−0.11
v_Esco.bsn	Soil evaporation compensation factor	Evaporation	0.01	1	0.31	0.29	0.4
r_Sol_Awc.sol	Available water capacity of soil layer (mm H_2O/mm soil)	Soil	−0.6	0.6	−0.015	0.03	−0.002
v_Revapmn.gw	Threshold depth of water in shallow aquifer for 'revap' to occur (mm)	Groundwater	0	300	186	159.2	185.3
v_Ch_K2.rte	Effective hydraulic conductivity in main channel alluvium (mm/h)	Channel	0	250	88	116.9	115.1
v_Gw_Revap.gw	Groundwater 'revap' coefficient	Groundwater	0.02	0.2	0.073	0.14	0.09
v_Epco.hru	Plant uptake composition factor	Crop	0.01	1	0.56	0.5	0.49
r_Sol_K	Saturated hydraulic conductivity	Soil	−0.24	0.25	0.13	0.01	0.01
v_Gw_Delay	Groundwater delay time (days)	Groundwater	40	450	49.3	45.14	35.67
v_Ch_N2	Manning's 'n' value for the main channel	Runoff	0.01	0.3	0.06	0.09	0.16

Evaluation of hydrological simulations based on different rainfall data

Table 4 presents model performance results for monthly flow simulations considering the three scenarios at the three stations. In the case of scenario I, the SWAT model using gauge rainfall data showed good simulations of monthly streamflow with NSCE values 0.85, 0.76, and 0.62 for Werie, Geba, and Emba Madre stations, respectively. The relative monthly streamflow errors were 19.3% (Emba Madre), 2.8% (Geba), and −3.7% (Werie). The CC values were high at Werie (0.92) and relatively low at Emba Madre (0.82). The model was considered applicable and provided a sound basis for evaluating the adequacy of using CHIRPSv8 and EWEMBI to drive the model in scenario I. However, the subsequent performance of the SWAT model using CHIRPSv8 and EWEMBI data was not as good as the gauge-driven performance. The NSCE values were decreased to 0.37–0.66 for CHIRPSv8 and 0.32–0.61 for EWEMBI rainfall forced simulated streamflow. The relative simulated monthly flow errors were between −1.9 and 14.3% (CHIRPSv8) and 4.72 and 21.1% (EWEMBI) across the three streamflow stations (Table 4).

For the case of scenario II, the model performance using CHIRPSv8 rainfall data was slightly improved compared with scenario I. The NSCE values at the three streamflow stations ranged from 0.58 to 0.69. The relative monthly streamflow bias was between −2.1 and 23.4%, and the CC ranged from 0.8 to 0.83. For the case of scenario III, Table 4 shows that model performance based on calibration using EWEMBI is almost similar to its performance in scenarios I and II, with NSCE ranging between 0.53 and 0.55, and relative bias ranging between 14 and 23.6%.

Figure 6 illustrates observed and simulated monthly streamflow hydrographs, driving by CHIRPSv8 and EWEMBI rainfall datasets, based on a model calibrated using corresponding rainfall. The simulated streamflow driven by gauge rainfall data showed a better agreement with streamflow observations than model simulations driven by the CHIRPSv8 and EWEMBI rainfall. The average values of KGE for gauge rainfall-driven simulations were 0.63, 0.73, and 0.88 for Emba Madre, Geba, and Werie stations, respectively. Considering the classification of KGE, the SWAT simulations driven by the CHIRPSv8 and EWEMBI rainfall showed intermediate to good hydrological performance, and CHIRPSv8 is better than EWEMBI. The streamflow peaks were often overestimated in the simulations driven by EWEMBI and CHIRPSv8 at all the three stations.

Figure 7 depicts the hydrological performance during low-flow and high-flow seasons when rainfall data-specified calibrations were used. With the exception of using gauge rainfall data at the Emba Madre gauge station, all the rainfall estimates (EWEMBI, CHIRPSv8, and gauge rainfall data) achieve a better hydrological performance during the high-flow season than the low-flow season (Figure 7(a)). The KGE scores in hydrological performance range from 0.6 to 0.79 (gauge rainfall), 0.59–0.68 (CHIRPSv8), and 0.53–0.58 (EWEMBI) for high-flow seasons across the three streamflow stations. The high CCs indicate that the timing and shape of the hydrograph were well reproduced during the high-flow season (Figure 7(b)). This is consistent with the seasonal daily rainfall intercomparison of the rainfall products over the basin. The CC between the CHIRPSv8 (or EWMBI) and gauge rainfall data implies relatively better agreement of the rainfall datasets for the wet season, and this enables a better hydrological performance during the high-flow season (Figure 3). A higher absolute bias (36.1%) in simulated monthly flow was showed in the high-flow season at the Emba Madre gauge station by the EWEMBI rainfall (Figure 7(c)).

The hydrological performance of the model driven by gauge rainfall, CHIRPSv8, and EWEMBI rainfall estimates was relatively poor during low-flow seasons. The low correlation in timing and magnitude of the observed and simulated hydrographs by the three rainfall estimates is the key limiting factor in relatively poor performances in the low-flow seasons. This can be explained by the fact that the calibration process which optimizes the model parameters is often dominated by the high-flow spectrum rather than the low-flow spectrum.

We further inspected model accuracy using the gauge rainfall, CHIRPSv8, and EWEMBI rainfall data at the main outlet of the basin (Emba Madre) for daily streamflow simulation during 2006–2015 (Figure 8). Figure 8 indicates daily streamflow hydrograph and the model performance statistics. A satisfactorily model performance result is obtained using gauge-based rainfall, with the coefficient of determination (R^2) of 0.65, Pbias of 4.2%, and NSCE of 0.57. Similarly, the SWAT model based on CHIRPSv8 and EWEMBI rainfall data showed a satisfactory result, although with performance indicator values a bit lower than those of the gauge-based data-driven model. Even though the R^2 values are close to that of the gauge rainfall-driven model, a relative increment in Pbias is shown for the simulations by CHIRPSv8 (21.9%) and EWEMBI (29.7%). All the model hydrological simulations using gauge rainfall, CHIRPSv8, and EWEMBI data overestimated the daily discharge of the basin. Generally, from the above results, CHIRPSv8 and EWEMBI rainfall can be used for hydrological simulation purposes in the area of interest and then can be further used to estimate hydrological budget for the basin.

Table 4 | Results for the hydrological modelling of the UTB using the SWAT

| Scenario	Station	Gauge observed					CHIRPSv8					EWEMBI				
					95PPU					95PPU					95PPU	
		NSCE	Bias	CC	*P*-factor	*R*-factor	NSCE	Bias	CC	*P*-factor	*R*-factor	NSCE	Bias	CC	*P*-factor	*R*-factor
1	Emba Madre	0.62	19.3	0.82	0.63	0.92	0.37	14.3	0.71	–	–	0.32	21.1	0.64	–	–
	Geba	0.76	2.8	0.87	0.76	1.05	0.66	−1.9	0.74	–	–	0.59	11.74	0.53	–	–
	Werie	0.85	−3.7	0.92	0.82	1.14	0.59	2.7	0.66	–	–	0.61	4.72	0.67	–	–
2	Emba Madre	0.35	23.3	0.77	–	–	0.58	22.1	0.83	0.58	0.67	0.28	25.14	0.59	–	–
	Geba	0.73	3.19	0.86	–	–	0.69	23.4	0.87	0.69	0.94	0.52	23.8	0.63	–	–
	Werie	0.83	−4.22	0.91	–	–	0.63	−2.1	0.80	0.77	1.12	0.49	1.32	0.74	–	–
3	Emba Madre	0.38	23.9	0.72	–	–	0.34	24.7	0.58	–	–	0.53	23.6	0.77	0.67	0.77
	Geba	0.73	2.96	0.84	–	–	0.59	19.7	0.70	–	–	0.51	14.0	0.82	0.71	1.25
	Werie	0.79	−2.6	0.87	–	–	0.61	2.7	0.64	–	–	0.55	20.4	0.85	0.75	0.98
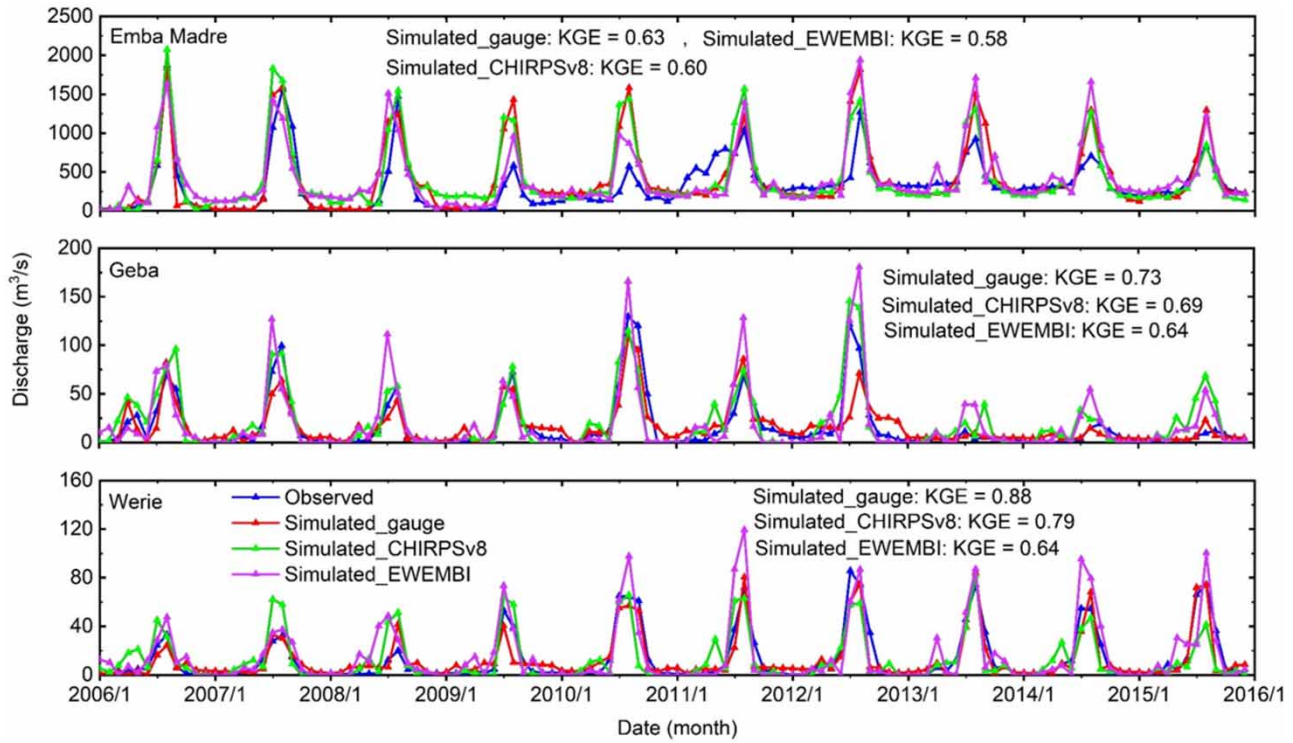
Figure 6 | Comparison of observed vs. simulated monthly hydrographs driven by CHIRPSv8 and EWEMBI rainfall estimates and the rain gauge rainfall data with their respective calibrated model parameter values at the Emba Madre, Geba and the Werie stations.
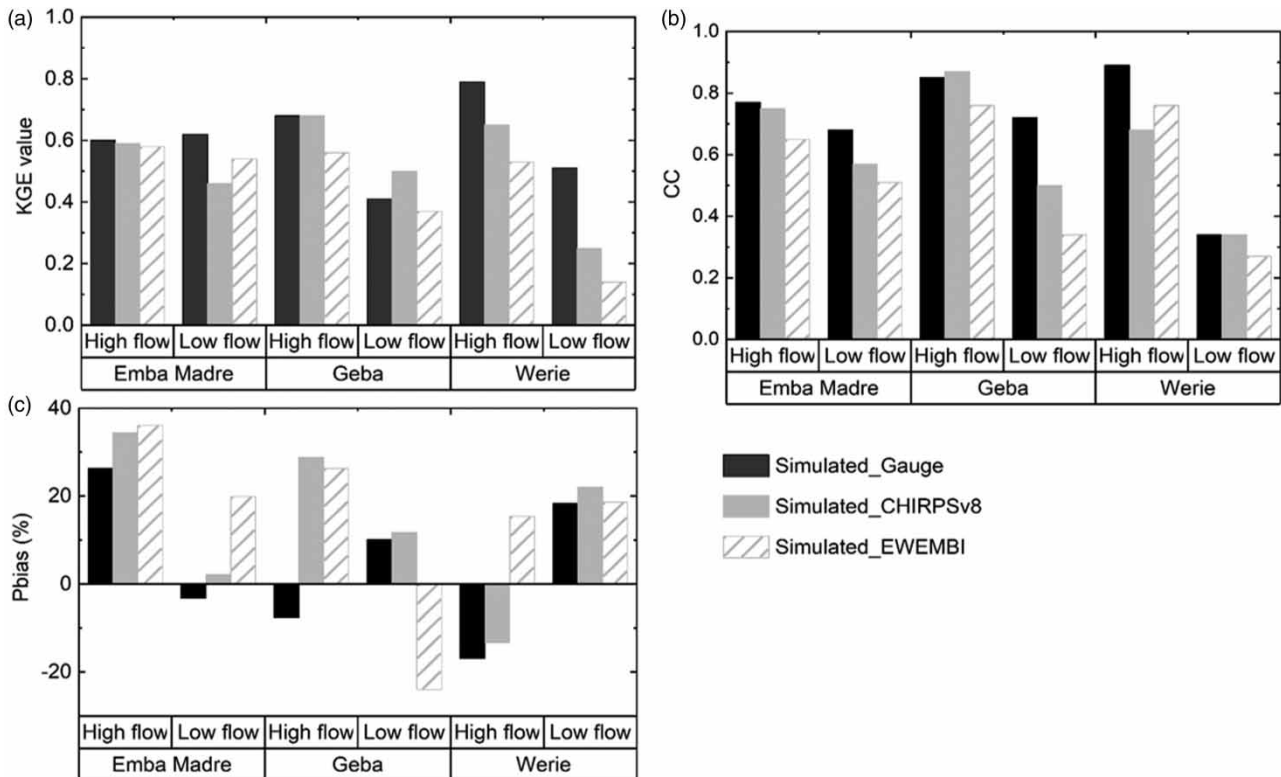

Figure 7 | Hydrological performance during high-flow and low-flow seasons at monthly time scale: (a) KGE, (b) correlation coefficient and (c) percent bias (%).
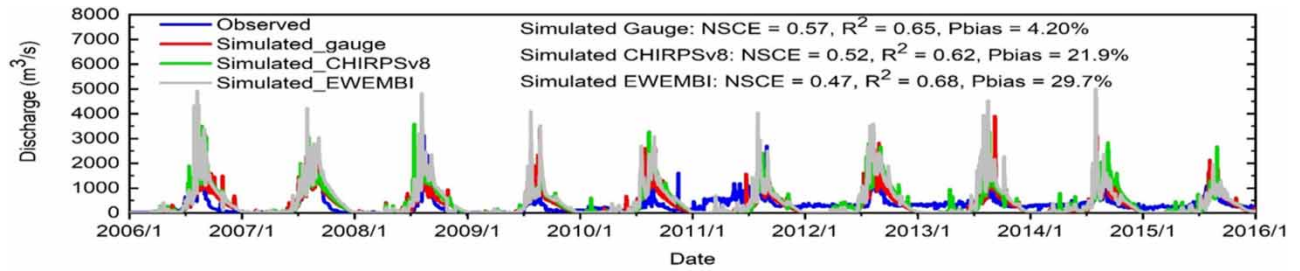
Figure 8 | Comparison of observed vs. simulated daily hydrographs at the Emba Madre station, model driven by the three rainfall estimates during 2006–2016.

Evaluation of water balance

Table 5 presents the yearly averaged depth of water balance components (precipitation, evapotranspiration, surface runoff, and baseflow) in the UTB. The simulated water balance difference is slight. The estimated annual rainfall was 963.3, 949.5, and 1,060.9 mm year^{-1} for gauge rainfall, CHIRPSv8, and EWEMBI, respectively. In regard to the gauge rainfall-driven simulations, 49.9% of the rainfall was exhausted through evapotranspiration, whereas in the case of CHIRPSv8 and EWEMBI rainfall estimates, the corresponding evapotranspiration rate was 50.3 and 41.1%, respectively. The proportion of the rainfall changed to groundwater recharge was about 15.2% (gauge rainfall), 22.1% (CHIRPSv8 rainfall), and 16.4% (EWEMBI rainfall). Because the groundwater recharge is a determinant factor that influences the amount of base flow in the hydrological regime, the base flow component was higher in CHIRPSv8 (38.6%), corresponding to its larger simulated ground water recharge rather than simulation by the other two rainfall products. With regard to the simulated total runoff component, proportionally higher rainfall was transformed into runoff in the CHIRPSv8 rainfall-driven simulation (576.3 mm, 60.7% of its annual rainfall). A total runoff of 596.63 mm (EWEMBI) and 554.07 mm year^{-1} (gauge rainfall) were produced, which comprised 56.2 and 57.5% of annual rainfall, respectively. In general, CHIRPSv8 and EWEMBI are also useful for general water budget calculations and other similar applications.

DISCUSSION AND IMPLICATIONS

In this study, we assessed the hydrological simulation performances using the model calibrated with CHIRPSv8, EWEMBI, and gauge rainfall (i.e., scenarios I, II, and II). In addition, we examined the performance of the CHIRPSv8 and EWEMBI rainfall products during low- and high-flow seasons separately at monthly and daily time scales over the UTB.

The SWAT model performance in simulating monthly streamflow driven by CHIRPSv8, EWEMBI, and gauge rainfall data at three streamflow gauge stations was good. Under the full time-series analysis, results from the hydrological evaluation indicates that the CHIRPSv8 and EWEMBI have the potential to be used as driving forcing for hydrological modelling over the UTB. According to Arnold *et al.* (2012), Thiemig *et al.* (2013), and Li *et al.* (2018), hydrological models are considered

Table 5 | The mean simulated hydrological components using the gauge-, CHIRPSv8-, and EWEMBI-based modelling during 2006–2015

Hydrological regimes	Gauge-based modelling			CHIRPSv8-based modelling			EWEMBI-based modelling		
	Depth (mm/yr)	Precipitation (%)	% from total runoff	Volume (mm/yr)	Precipitation (%)	Total runoff (%)	Volume (mm/yr)	Precipitation (%)	% from total runoff
Precipitation	963.3			949.5			1,060.9		
Evapotranspiration	480.69	49.9		477.69	50.3		436.0	41.1	
Groundwater recharge	146.23	15.2		209.46	22.1		173.6	16.4	
Total runoff	554.07	57.5		576.30	60.7		596.63	56.2	
Surface runoff	214.82	22.3	38.77	190.42	20.1	33.04	209.96	19.8	35.19
Base flow	339.82	35.3	61.33	366.34	38.6	63.57	386.67	36.4	64.81

satisfactory if the NSCE and KGE are greater than 0.5 and the absolute Pbias is less than 25% for monthly streamflow simulations. In addition, this can be supported spatially by the fact that two of the rainfall products achieve sufficiently satisfactory hydrological performances over the three gauge stations located over the lowland (Emba Madre station), midland (Geba station), and highland (Werie station) parts of the basin. And this is sufficiently comparable to the gauge rainfall-driven hydrological simulation result. This outcome is highly desirable, especially for data-sparse and ungauged basins.

Even though the overall hydrological performance through the entire study period is sufficiently satisfactory for hydrological applications, some hydrological and water management applications need high performance for a certain flow condition. For example, for drought and flood analysis, good performance during low-flow and high-flow periods is important. To this end, we analysed the hydrological performance during high-flow (June–September) and low-flow (October–May) seasons separately. The results indicate the potential of the products used as forcing data in flood management, dam reservoir storage modelling, and other hydrological applications that require accurate forecasting of high-flow conditions. The better hydrological performance using satellite rainfall products was also shown by Thiemig *et al.* (2013) in high-flow periods than low-flow periods. The rainfall over the UTB is highly seasonal with greater than 70% of annual rainfall received from June to September (Fentaw *et al.* 2018; Gebremicael *et al.* 2019). CHIRPSv8 and EWEMBI also showed a better hydrological performance during high-flow periods. CHIRPSv8 performs well even during the low-flow seasons. This implies, for our study areas, that the use of CHIRPSv8 could be advisable for hydrological applications that need accurate reproduction of low-flow conditions, for example, in meteorological and hydrological drought monitoring and management.

Beside the daily and monthly streamflow hydrograph comparisons examined above, hydrological water balance component modelling delivers additional important indicators for evaluating rainfall data adequacy in driving hydrological models (Xu *et al.* 2013; Wang *et al.* 2020). Hence, based on rain gauges, CHIRPSv8, and EWEMBI rainfall-driven simulated monthly streamflows, we examined the discrepancy of the simulated water balance components of the basin. The simulated water balance difference is slight among the rainfall datasets. Hence, the two rainfall datasets can be used in the applications that need water balance analysis in the UTB, and then water balances will help to characterize, in a coherent manner, the hydrological functioning of the basin.

Finally, this study could be important for planners, decision-makers, and researchers in the water resources sector in ungauged regions. Our study indicates that satellite and reanalysis-based rainfall products with high spatio-temporal resolutions can be used as an alternative data sources for meteorological forcing in hydrological process modelling in the ungauged basins. For instance, the recommended rain gauge density is one station per 600–900 km^2 for flatlands and 100–250 km^2 for topographically rugged areas (Zeng *et al.* 2018). Africa has only 744 stations and only a quarter of them meet the international standards which is far below the recommended to capture its spatio-temporal climate variabilities (Satgé *et al.* 2020). Hence, this study could be helpful as a reference for other mountainous regions in Africa. Even in gauged catchments, the quality and stability of climatic data from ground meteorological stations may not be sufficient due to limited numbers of stations, uneven spatial distribution, and vulnerability to human and environmental factors. Under such conditions anywhere in the globe, our study is an indicator how the satellite and reanalysis rainfall products are very important to represent ground observations.

CONCLUSIONS

This study compared the daily rainfall from two rainfall products CHIRPSv8 and EWEMBI with rain gauge data and assessed the adequacy of hydrological process modelling using CHIRPSv8 and EWMBI rainfall as forcing data at the UTB, northern Ethiopia. The following conclusions can be drawn.

I. Compared with the rain gauges, the biases of rainfall estimate from the CHIRPSv8 and EWEMBI dataset are within an acceptable range. In both the CHIRPSv8 and EWEMBI rainfall products, the hit bias error component is dominant as compared with the missed and false error components. In addition, both CHIRPSv8 and EWMBI show relatively better agreement with the gauge rainfall for the wet season than the dry season period. The differences of maximum daily and maximum 5-day rainfall between CHIRPSv8 or EWEMBI and rain gauges are large.

II. The SWAT-simulated hydrological performance for the monthly and daily streamflow simulations driven by both CHIRPSv8 and EWEMBI rainfalls is close to streamflow simulation driven by gauge-based rainfall, implying that the satellite-based and reanalysis data, although suffering biases, are applicable for hydrological application over the

UTB. The hydrological simulations driven by CHIRPSv8 rainfall perform relatively better than EWEMBI rainfall in the area of interest.

III. In general, the hydrological simulation shows better performance during the high-flow season than the low-flow season using both CHIRPSv8 and EWEMBI rainfall products.

IV. The estimated CHIRPSv8 and EWEMBI rainfall data can be used as an alternative meteorological forcing for hydrological modelling and applications, especially for data-sparse and ungauged basins such as the UTB, Ethiopia.

ACKNOWLEDGEMENTS

COMPETING INTEREST

The authors declare no conflict of interest

AUTHOR CONTRIBUTION

K.W.R. conceptualized the article, developed the methodology and software, and wrote the original draft. X.L. conceptualized the article and wrote, reviewed and edited the whole article. G.G.H. and S.S. reviewed and edited the article. Q.T. wrote, reviewed and edited the article.

DATA AVAILABILITY STATEMENT

All relevant data are available from EWEMBI: http://dataservices.gfz-potsdam.de/pik/showshort.php?id=escidoc:3928916 and CHIRPSv8: https://www.chc.ucsb.edu/data.

REFERENCES

Abbaspour, K. C., Yang, J., Maximov, I., Siber, R., Bogner, K., Mieleitner, J. & Srinivasan, R. 2007 Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT. *Journal of Hydrology* **333** (2–4), 413–430.

Abbaspour, K., Vaghefi, S. & Srinivasan, R. 2017 A guideline for successful calibration and uncertainty analysis for soil and water assessment: a review of papers from the 2016 international SWAT conference. *Water* **10** (1), 6.

Akoko, G., Le, T. H., Gomi, T. & Kato, T. 2021 A review of SWAT model application in Africa. *Water* **13** (9), 1313.

Arnold, J. G., Kiniry, J. R., Srinivasan, R., Williams, J. R., Haney, E. B. & Neitsch, S. L. 2012 *Soil & Water Assessment Tool (SWAT): Input/Output Documentation Version 2012*. Texas Water Resources Institute, TR-439.

Awange, J. L., Hu, K. X. & Khaki, M. 2019 The newly merged satellite remotely sensed, gauge and reanalysis-based multi-source weighted-ensemble precipitation: evaluation over Australia and Africa (1981–2016). *Science of the Total Environment* **670**, 448–465.

Azimi, S., Dariane, A. B., Modanesi, S., Bauer-Marschallinger, B., Bindlish, R., Wagner, W. & Massari, C. 2020 Assimilation of Sentinel 1 and SMAP-based satellite soil moisture retrievals into SWAT hydrological model: the impact of satellite revisit time and product spatial resolution on flood simulations in small basins. *Journal of Hydrology* **581**, 124367.

Bayissa, Y., Tadesse, T., Demisse, G. & Shiferaw, A. 2017 Evaluation of satellite-based rainfall estimates and application to monitor meteorological drought for the Upper Blue Nile Basin, Ethiopia. *Remote Sensing* **9** (7), 669.

Beck, H. E., van Dijk, A. I. J. M., de Roo, A., Dutra, E., Fink, G., Orth, R. & Schellekens, J. 2017a Global evaluation of runoff from 10 state-of-the-art hydrological models. *Hydrology and Earth System Sciences* **21** (6), 2881–2903.

Beck, H. E., Vergopolan, N., Pan, M., Levizzani, V., van Dijk, A. I. J. M., Weedon, G. P. & Wood, E. F. 2017b Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. *Hydrology and Earth System Sciences* **21** (12), 6201–6217.

Belete, K. 2007 *Sedimentation and Sediment Handling at Dams in Tekeze River Basin, Ethiopia*. PhD Thesis. Norwegian University of Science and Technology, Trondheim, Norway.

Chawanda, C. J., Arnold, J., Thiery, W. & van Griensven, A. 2020 Mass balance calibration and reservoir representations for large-scale hydrological impact studies using SWAT+. *Climatic Change* **163** (3), 1307–1327.

Chen, C., Chen, Q., Qin, B., Zhao, S. & Duan, Z. 2020 Comparison of different methods for spatial downscaling of GPM IMERG v06B satellite precipitation product over a typical arid to semi-arid area. *Frontiers in Earth Science* **8**, 536337.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S. & Vitart, F. 2011 The ERA-interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society* **137** (656), 553–597.

Dinku, T., Funk, C., Peterson, P., Maidment, R., Tadesse, T., Gadain, H. & Ceccato, P. 2018 Validation of the CHIRPS satellite rainfall estimates over eastern Africa. *Quarterly Journal of the Royal Meteorological Society* **144** (S1), 292–312.

Faramarzi, M., Abbaspour, K. C., Ashraf Vaghefi, S., Farzaneh, M. R., Zehnder, A. J. B., Srinivasan, R. & Yang, H. 2013 Modeling impacts of climate change on freshwater availability in Africa. *Journal of Hydrology* **480**, 85–101.

Fentaw, F., Hailu, D., Nigussie, A. & Melesse, A. M. 2018 Climate change impact on the hydrology of Tekeze Basin, Ethiopia: projection of rainfall-runoff for future water resources planning. *Water Conservation Science and Engineering* **3** (4), 267–278.

Francesconi, W., Srinivasan, R., Pérez-Miñana, E., Willcock, S. P. & Quintero, M. 2016 Using the Soil and Water Assessment Tool (SWAT) to model ecosystem services: a systematic review. *Journal of Hydrology* **535**, 625–636.

Frieler, K., Lange, S., Piontek, F., Reyer, C. P. O., Schewe, J., Warszawski, L. & Yamagata, Y. 2017 Assessing the impacts of 1.5 °C global warming – simulation protocol of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP2b). *Geoscientific Model Development* **10** (12), 4321–4345.

Funk, C., Peterson, P. J., Landsfeld, M. F., Pedreros, D. H., Verdin, J. P., Rowland, J. D. & Verdin, A. P. 2014 A quasi-global precipitation time series for drought monitoring. *U.S. Geological Survey Data Series* **832** (4). https://dx.doi.org/10.3133/ds832.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S. & Michaelsen, J. 2015 The climate hazards infrared precipitation with stations – a new environmental record for monitoring extremes. *Scientific Data* **2**, 150066.

Gebremicael, T. G., Mohamed, Y. A., Betrie, G. D., van der Zaag, P. & Teferi, E. 2013 Trend analysis of runoff and sediment fluxes in the Upper Blue Nile basin: a combined analysis of statistical tests, physically-based models and landuse maps. *Journal of Hydrology* **482**, 57–68.

Gebremicael, T. G., Mohamed, Y. A., Zaag, P. v. & Hagos, E. Y. 2017 Temporal and spatial changes of rainfall and streamflow in the Upper Tekezē–Atbara river basin, Ethiopia. *Hydrology and Earth System Sciences* **21** (4), 2127–2142.

Gebremeskel, G., Gebremicael, T. G. & Girmay, A. 2018 Economic and environmental rehabilitation through soil and water conservation, the case of Tigray in northern Ethiopia. *Journal of Arid Environments* **151**, 113–124.

Gebremicael, T. G., Mohamed, Y. A., Zaag, P. v. d., Gebremedhin, A., Gebremeskel, G., Yazew, E. & Kifle, M. 2019 Evaluation of multiple satellite rainfall products over the rugged topography of the Tekeze-Atbara basin in Ethiopia. *International Journal of Remote Sensing* **40** (11), 4326–4345.

Griensven, A., Ndomba, P., Yalew, S. & Kilonzo, F. 2012 Critical review of SWAT applications in the upper Nile basin countries. *Hydrology and Earth System Sciences* **16** (9), 3371–3381.

Guyot, J.-L., Lavado-Casimiro, W., Buytaert, W., Onof, C., Wang, L.-P., Zulkafli, Z. & Nerini, D. 2015 A comparative analysis of TRMM–Rain Gauge Data Merging Techniques at the daily time scale for distributed rainfall–runoff modeling applications. *Journal of Hydrometeorology* **16** (5), 2153–2168.

Habib, E., Larson, B. F. & Graschel, J. 2009 Validation of NEXRAD multisensor precipitation estimates using an experimental dense rain gauge network in south Louisiana. *Journal of Hydrology* **373** (3–4), 463–478.

Haile, G. G., Tang, Q., Leng, G., Jia, G., Wang, J., Cai, D. & Zhang, Q. 2020 Long-term spatiotemporal variation of drought patterns over the Greater Horn of Africa. *Science of the Total Environment* **704**, 135299.

Herrera-Estrada, J. E., Satoh, Y. & Sheffield, J. 2017 Spatiotemporal dynamics of global drought. *Geophysical Research Letters* **44** (5), 2254–2263.

Hu, S., Qiu, H., Yang, D., Cao, M., Song, J., Wu, J. & Gao, Y. 2017 Evaluation of the applicability of climate forecast system reanalysis weather data for hydrologic simulation: a case study in the Bahe River Basin of the Qinling Mountains, China. *Journal of Geographical Sciences* **27** (5), 546–564.

Jia, S., Zhu, W., Lű, A. & Yan, T. 2011 A statistical spatial downscaling algorithm of TRMM precipitation based on NDVI and DEM in the Qaidam Basin of China. *Remote Sensing of Environment* **115** (12), 3069–3079.

Kling, H., Fuchs, M. & Paulin, M. 2012 Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *Journal of Hydrology* **424–425**, 264–277.

Kneis, D., Chatterjee, C. & Singh, R. 2014 Evaluation of TRMM rainfall estimates over a large Indian river basin (Mahanadi). *Hydrology and Earth System Sciences* **18** (7), 2493–2502.

Lakew, H. B., Moges, S. A. & Asfaw, D. H. 2017 Hydrological evaluation of satellite and reanalysis precipitation products in the Upper Blue Nile Basin: a case study of Gilgel Abbay. *Hydrology* **4** (3), 39.

Lange, S. 2016 *EartH2Observe, WFDEI and ERA-Interim Data Merged and Bias-Corrected for ISIMIP (EWEMBI).* GFZ Data Services. https://doi.org/10.5880/pik.2019.004.

Lange, S. 2018 Bias correction of surface downwelling longwave and shortwave radiation for the EWEMBI dataset. *Earth System Dynamics* **9** (2), 627–645.

Lemma, E., Upadhyaya, S. & Ramsankaran, R. 2019 Investigating the performance of satellite and reanalysis rainfall products at monthly timescales across different rainfall regimes of Ethiopia. *International Journal of Remote Sensing* **40** (10), 4019–4042.

Li, D., Christakos, G., Ding, X. & Wu, J. 2018 Adequacy of TRMM satellite rainfall data in driving the SWAT modeling of Tiaoxi catchment (Taihu lake basin, China). *Journal of Hydrology* **556**, 1139–1152.

Lo Conti, F., Hsu, K.-L., Noto, L. V. & Sorooshian, S. 2014 Evaluation and comparison of satellite precipitation estimates with reference to a local area in the Mediterranean Sea. *Atmospheric Research* **138**, 189–204.

Mair, A. & Fares, A. 2011 Comparison of rainfall interpolation methods in a mountainous region of a Tropical Island. *Journal of Hydrologic Engineering* **16** (4), 371–383.

Meng, J., Li, L., Hao, Z., Wang, J. & Shao, Q. 2014 Suitability of TRMM satellite rainfall in driving a distributed hydrological model in the source region of Yellow River. *Journal of Hydrology* **509**, 320–332.

Mengistu, A. G., van Rensburg, L. D. & Woyessa, Y. E. 2019 Techniques for calibration and validation of SWAT model in data scarce arid and semi-arid catchments in South Africa. *Journal of Hydrology: Regional Studies* **25**, 100621.

Nadew, B. 2018 Stream flow and sediment yield modeling: a case study of Beles watershed, Upper Blue Nile Basin. *Irrigation & Drainage Systems Engineering* **07** (03), 216.

Neitsch, S. L., Arnold, J. R., Kiniry, J. R. & Williams, J. R. 2012 *Soil and Water Assessment Tool (SWAT) Model Version 2012, Texas Water Resources Institute, Technical Report No. 406*. Texas A&M University System, Collage Station, TX, USA.

Ouermi, K. S., Paturel, J.-E., Adounpke, J., Lawin, A. E., Goula, B. T. A. & Amoussou, E. 2019 Comparison of hydrological models for use in climate change studies: a test on 241 catchments in West and Central Africa. *Comptes Rendus Geoscience* **351** (7), 477–486.

Raimonet, M., Oudin, L. & Thieu, V. 2017 Evaluation of gridded meteorological datasets for hydrological modeling. *Journal of Hydrometeorology* **18** (11), 3027–3041.

Rajulapati, C. R., Papalexiou, S. M., Clark, M. P., Razavi, S., Tang, G. & Pomeroy, J. W. 2020 Assessment of extremes in global precipitation products: how reliable are they? *Journal of Hydrometeorology* **21** (12), 2855–2873.

Reda, K. W., Liu, X., Tang, Q. & Gebremicael, T. G. 2021 Evaluation of global gridded precipitation and temperature datasets against gauged observations over the Upper Tekeze River Basin, Ethiopia. *Journal of Meteorological Research* **35** (4), 673–689.

Roy, A., Thakur, P. K., Pokhriyal, N., Aggarwal, S. P., Nikam, B. R., Garg, V. & Choksey, A. 2018 Intercomparison of different rainfall products and validation of WRF modelled rainfall estimation in N-W Himalya during monsoon period. In: *ISPRS TC V Mid-Term Symposium Geospatial Technology – Pixel to People 4-5* (Kumar, A. S., Saran, S. & Padalia, H., eds), ISPRS, Hannover, Germany, pp. 351–358.

Satgé, F., Defrance, D., Sultan, B., Bonnet, M.-P., Seyler, F., Rouché, N. & Paturel, J.-E. 2020 Evaluation of 23 gridded precipitation datasets across West Africa. *Journal of Hydrology* **581**, 124412.

Schuol, J., Abbaspour, K. C., Srinivasan, R. & Yang, H. 2008 Estimation of freshwater availability in the West African sub-continent using the SWAT hydrologic model. *Journal of Hydrology* **352** (1–2), 30–49.

Singh, A., Imtiyaz, M., Isaac, R. K. & Denis, D. M. 2014 Assessing the performance and uncertainty analysis of the SWAT and RBNN models for simulation of sediment yield in the Nagwa watershed, India. *Hydrological Sciences Journal* **59** (2), 351–364.

Sperna Weiland, F. C., Vrugt, J. A., van Beek, R. P. H., Weerts, A. H. & Bierkens, M. F. P. 2015 Significant uncertainty in global scale hydrological modeling from precipitation data errors. *Journal of Hydrology* **529**, 1095–1115.

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S. & Hsu, K. L. 2018 A review of global precipitation data sets: data sources, estimation, and intercomparisons. *Reviews of Geophysics* **56** (1), 79–107.

Tan, G., Philip, W. & Cracknell, A. P. 2017 Assessment of three long-term gridded climate products for hydro-climatic simulations in Tropical River Basins. *Water* **9** (3), 229.

Tan, X., Wu, Y., Liu, B. & Chen, S. 2020 Inconsistent changes in global precipitation seasonality in seven precipitation datasets. *Climate Dynamics* **54** (5–6), 3091–3108.

Thiemig, V., Rojas, R., Zambrano-Bigiarini, M. & De Roo, A. 2013 Hydrological evaluation of satellite-based rainfall estimates over the Volta and Baro-Akobo Basin. *Journal of Hydrology* **499**, 324–338.

Wang, J., Liu, G. & Zhu, C. 2020 Evaluating precipitation products for hydrologic modeling over a large river basin in the Midwestern USA. *Hydrological Sciences Journal* **65** (7), 1221–1238.

Wang, T., Tu, X., Singh, V. P., Chen, X. & Lin, K. 2021 Global data assessment and analysis of drought characteristics based on CMIP6. *Journal of Hydrology* **596**, 126091.

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.

Welde, K. & Gebremariam, B. 2017 Effect of land use land cover dynamics on hydrological response of watershed: case study of Tekeze Dam watershed, northern Ethiopia. *International Soil and Water Conservation Research* **5** (1), 1–16.

Wu, H. & Chen, B. 2015 Evaluating uncertainty estimates in distributed hydrological modeling for the Wenjing River watershed in China by GLUE, SUFI-2, and ParaSol methods. *Ecological Engineering* **76**, 110–121.

Xu, H., Xu, C.-Y., Chen, H., Zhang, Z. & Li, L. 2013 Assessing the influence of rain gauge density and distribution on hydrological model performance in a humid region of China. *Journal of Hydrology* **505**, 1–12.

Xue, X., Hong, Y., Limaye, A. S., Gourley, J. J., Huffman, G. J., Khan, S. I. & Chen, S. 2013 Statistical and hydrological evaluation of TRMM-based multi-satellite precipitation analysis over the Wangchu Basin of Bhutan: are the latest satellite precipitation products 3B42V7 ready for use in ungauged basins? *Journal of Hydrology* **499**, 91–99.

Yang, J., Reichert, P., Abbaspour, K. C., Xia, J. & Yang, H. 2008 Comparing uncertainty analysis techniques for a SWAT application to the Chaohe Basin in China. *Journal of Hydrology* **358** (1–2), 1–23.

Yue, T. X., Wang, Y. F., Du, Z. P., Zhao, M. W., Zhang, L. L., Zhao, N. & Wilson, J. P. 2016 Analysing the uncertainty of estimating forest carbon stocks in China. *Biogeosciences* **13** (13), 3991–4004.

Zenebe, A. 2009 *Assessment of Spatial and Temporal Variability of River Discharge, Sediment Yield and Sediment-Fixed Nutrient Export in Geba River Catchment, Northern Ethiopia*. PhD, Stoch. Env. Res. Risk. A., KU Leuven, Belgium.

Zeng, Q., Chen, H., Xu, C.-Y., Jie, M.-X., Chen, J., Guo, S.-L. & Liu, J. 2018 The effect of rain gauge density and distribution on runoff simulation using a lumped hydrological modelling approach. *Journal of Hydrology* **563**, 106–122.

Zhao, H., Yang, S., Wang, Z., Zhou, X., Luo, Y. & Wu, L. 2015 Evaluating the suitability of TRMM satellite rainfall data for hydrological simulation using a distributed hydrological model in the Weihe River catchment in China. *Journal of Geographical Sciences* **25** (2), 177–195.

Zhou, J., Liu, Y., Guo, H. & He, D. 2014 Combining the SWAT model with sequential uncertainty fitting algorithm for streamflow prediction and uncertainty analysis for the Lake Dianchi Basin, China. *Hydrological Processes* **28** (3), 521–533.