A High-Accuracy Rainfall Dataset by Merging Multi-Satellites and Dense Gauges over Southern Tibetan Plateau for 2014-2019 Warm Seasons

Kunbiao Li[1], Fuqiang Tian[1], Mohd Yawar Ali Khan[2], Ran Xu[1], Zhihua He[3], Long Yang[4], Hui Lu[5], Yingzhao Ma[6]

[1] Department of Hydraulic Engineering, Tsinghua University, Beijing, China

[2] Department of Hydrogeology, King Abdul-Aziz University, Jeddah, Saudi Arabia

[3] Centre for Hydrology, University of Saskatchewan, Saskatoon, SK S7N 5C8, Canada

[4] School of Geography and Ocean Science, Nanjing University, Nanjing, China

[5] Department of Earth System Science, Tsinghua University, Beijing, China

[6] Cooperative Institute for Research in the Atmosphere, Colorado State University, 1375 Campus Delivery, Fort Collins, CO 80523, United States.

Correspondence: Fuqiang Tian (tianfq@mail.tsinghua.edu.cn)

Abstract

Tibetan Plateau (TP) is well known as the Asia's water tower from where many large rivers originate. However, due to complex spatial variability of climate and topography, there is still a lack of high-quality rainfall dataset for hydrological modelling and flood prediction. This study, therefore, aims to establish a high-accuracy daily rainfall product through merging rainfall estimates from three satellites, i.e., GPM-IMERG, GSMaP, and CMORPH, based on the likelihood measurements of a high-density rainfall gauge network. The new merged daily rainfall dataset with a spatial resolution of 0.1°, focuses on warm seasons (June 10th - October 31st) from 2014 to 2019. Statistical evaluation indicated that the new dataset outperforms the raw satellite estimates, especially in terms of rainfall accumulation and the detection of ground-based rainfall events. Hydrological evaluation in the Yarlung Zangbo River Basin demonstrated high performance of the merged rainfall dataset in providing accurate and robust forcings for streamflow simulations. The new rainfall dataset additionally shows superiority to several other products of similar types, including MSWEP and CHIRPS. This new rainfall dataset is publicly accessible at https://doi.org/10.11888/Hydro.tpdc.271303 (Li et al.,2021).

1. Introduction

Precipitation, linking atmospheric and hydrological processes, serves as a crucial component of the water cycle (Eltahir & Bras, 1996; Trenberth et al., 2003). Gridded precipitation datasets become more and more popular with the advent of satellite precipitation measurement. Most famous satellite gridded precipitation datasets include Tropical Rainfall Measuring Mission (TRMM) (Huffman et al., 2007) and its successor the Integrated Multi-satellite Retrievals for Global Precipitation Measurement mission (GPM-IMERG) (Hou et al., 2014), the Global Satellite Mapping of Precipitation

(GSMaP) (Ushio et al., 2009), the Climate Prediction Centre (CPC) MORPHing technique (CMORPH) (Joyce et al., 2004), etc. These products have been successfully applied in various hydrometeorological studies and water resources management practices (Kidd, C., & Levizzani, V., 2011; Jiang et al., 2012; Tong et al., 2014; Yang et al., 2015; Sun et al., 2016; Wang et al., 2017).

However, all existing precipitation datasets show insufficient accuracy in high mountainous regions (Yilmaz et al., 2016; Derin et al., 2018; Derin et al., 2019; Anagnostou & Zhang, 2019), which hinders our understanding of climate and hydrological processes over these areas. This can be attributed to the complex physical nature of electromagnetic transmission and precipitation forming processes (Hong et al., 2007; Bitew & Gebremichael 2010; Dinku et al., 2010), and harsh environments in high mountains that lead to very limited deployment of in-situ rain gauges with insufficient representation of ground observations for training satellite-based precipitation retrieval algorithms. For instance, the Tibetan Plateau (TP) as the roof of the world is surrounded by imposing mountain ranges with an average elevation exceeding 4000 m. It generates several large rivers in Asia and provides invaluable freshwater resources for more than 1.4 billion people living downstream (Immerzeel et al., 2010). However, this vast plateau has very limited number of precipitation gauges across its 2.5 million km^2 area. The precipitation gauge network operated by China Meteorological Agency (CMA) contains only 86 gauges over the entire TP (Figure 1). These gauges are essential to correct satellite precipitation datasets. For example, GPM-IMERG 'Final' Run dataset uses Global Precipitation Climatology Centre (GPCC) database, GSMap_Gauge and CMORPH use NOAA Climate Prediction Centre (CPC) database. Although both GPCC and CPC databases received data through Global Telecommunication System (GTS), only part of the above-mentioned gauges in TP were utilized (Xie et al., 2007; Becker 2013). Previous evaluations over the TP indicated that most products present dependence on topography to varying degrees, and products adjusted by gauge observations shows better performance than satellite-only

70   products (Gao et al.,2013; Lu et al., 2018). Therefore, a better spatial coverage of rain

71   gauges is critical to correct satellite products in high mountains.

72   In 2014, the Ministry of Water Resources of China (MWR) launched the flash

73   flood monitoring and alarming campaign. A large number of rain gauges is now

74   accessible over the TP, especially in the southern TP. There are 440 new rain gauges

75   totally involved in 6 years, and are available since 2014, independent of the existing

76   CMA precipitation gauge network (Figure 1). These gauges provide measurements of

77   precipitation in liquid phase (i.e., rainfall) at event time scale. A couple of recent studies

78   have demonstrated the utility of this rain gauge network (Xu et al., 2017; He et al., 2017;

79   Tian et al., 2018; Wang et al., 2020). For instance, Xu et al. (2017) evaluated the

80   performance of TRMM and GPM and the dependence on topography and rainfall

81   intensity based on the network. Their results demonstrated that the data quality of this

82   dense gauge network is strictly controlled, serving as the currently highest gauge dense

83   for satellite product evaluation on TP. Wang et al. (2020) used the gauge data to validate

84   their reproduced precipitation dataset. However, there is not a merging product that

85   assimilate the observations from this dense rain gauge network. This is apparently a

86   unique opportunity to improve the performance of existing satellite-based precipitation

87   datasets for its highest density and quality.

88   This study aims to provide a high-accuracy rainfall dataset by merging all

89   available ground gauges and three good-quality satellite precipitation datasets over the

90   southern TP for the warm seasons (June 10th - October 31st) from 2014 to 2019. The

91   remainder of this paper is organized as follows: Section 2 describes the study area and

92   the source data. Section 3 provides details of the data merging method and the methods

93   adopted to evaluate the quality of dataset. Results are presented in Section 4. The data

94   availability and summary are provided in Section 5 and Section 6, respectively.

2. Study Area and Source Data

2.1. Southern Tibetan Plateau

Tibetan Plateau, known as the Asian water tower, mainly covers parts of China, India, Myanmar, Bhutan, Nepal and Pakistan. Various climate systems affect the plateau, including westerly winds in winter and the Indian monsoon in summer (Yao et al., 2012). Many Asian large rivers originate from this vast area, including the Yellow River, the Yangtze River, the Yarlung Zangbo River (YZR), Jinsha River (JR), Lancang River (LR), Salwen River (SR), Irrawaddy River (IR), Ganges River (GR), and Indus River (IDR). This study is focused on the southern part of TP (Figure 1), including the upper YZR Basin (YZRB) as a major basin.

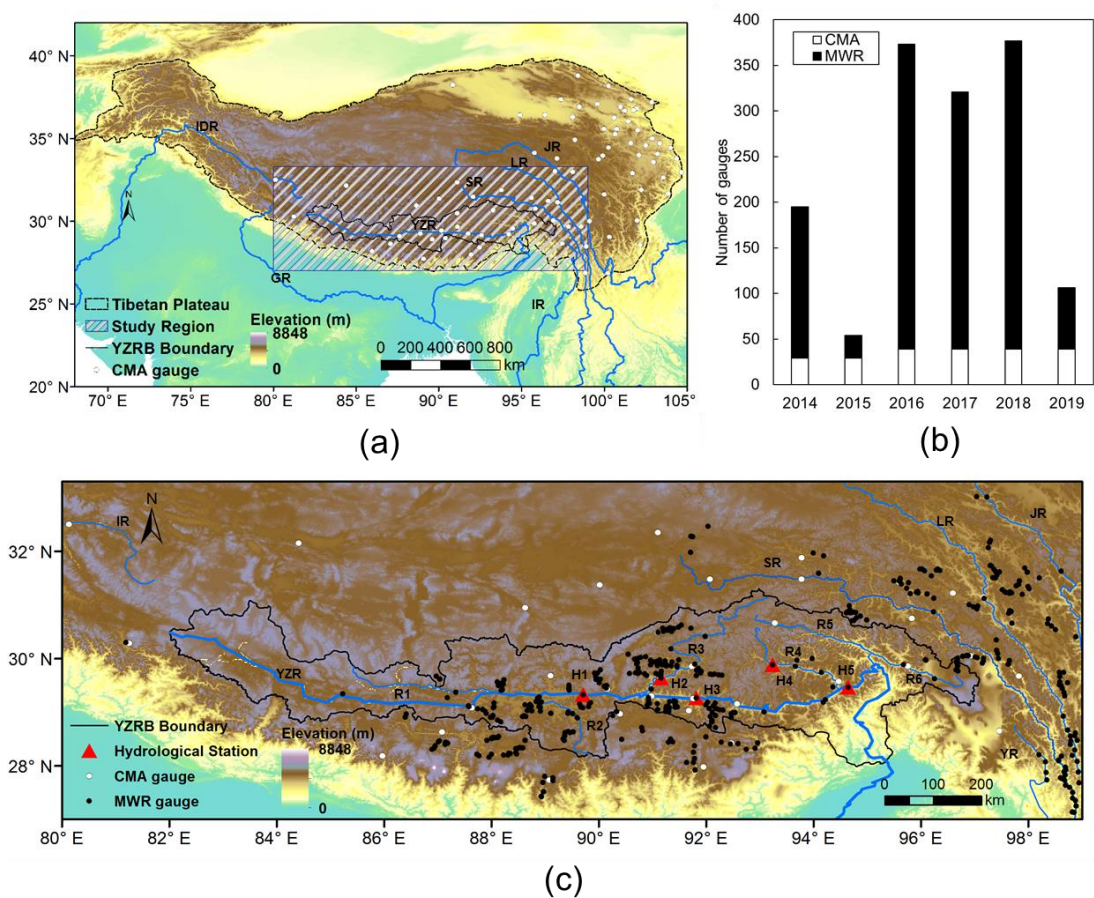



Figure 1. (a) The location and topography of the TP and the spatial distributions of CMA gauges. (b) Numbers of ground gauges installed by CMA and MWR in southern TP during 2014-2019, (c)

108   Locations of CMA and MWR rain gauges and main hydrological stations in southern TP. The names

109   of hydrological stations are labelled as H1-Yangcun, H2-Lhasa, H3-Nugesha, H4-Gongbujiangda,

110   H5-Nuxia. The names of tributary rivers are labelled as R1-Duoxiong Zangbo, R2-Nianchu River,

111   R3-Lhasa River, R4-Niyang River, R5-Yigong Zangbo, R6-Parlung Zangbo.

2.2. Ground gauged rainfall

113   We combined two rain gauge networks managed by MWR and CMA to obtain a

114   high-quality ground reference dataset up to date. The number of rain gauge is presented

115   in Figure 1b, and varies across different years. The spatial distribution of all gauges is

116   presented in Figure 1c. The gauges are mainly located in the middle reaches of YZRB

117   and the east part of the study area. Despite the high density, we can see these rain gauges

118   are not evenly distributed across the space. This makes satellite rainfall products over

119   varying altitudes and aspects important. Daily rainfall observations during the warm

120   seasons of 2014-2019 were accumulated from the original event scale measurements.

121   Total number of the CMA and MWR gauges ranges from 53 in 2015 to 377 in 2018,

122   forming the densest rain gauge network up till now.

123   The CMA gauge data has been widely demonstrated as reliable and accurate in

124   previous studies (Zhai et al., 2005; Su et al., 2020; He et al., 2020). Gauge data used in

125   this study has been manufactured under strict quality control procedures, including (1)

126   internal consistency check, (2) extreme values check (0~85mm/h), and (3) spatial

127   consistency check (Ren et al., 2010). Rain gauges with erroneous values (e.g.

128   enormously large values) were discarded from the entire records. Through strict quality

129   control methods, the warm seasons from June 10th to October 31st were selected as the

130   study period to maintain the high quality of outcome rainfall dataset.

2.3. Satellite Precipitation Datasets

132   Three satellite precipitation products were chosen for the data merging procedure

133   (Lu et al., 2019; Derin et al., 2019; Tang et al., 2020), including GPM-IMERG 'Final'

run (here after referred to as IMERG) from the National Aeronautics and Space Administration (NASA) (https://disc.gsfc.nasa.gov/), the GSMaP_Gauge (here after referred to as GSMaP) from Japan Aerospace Exploration Agency (JAXA) (http://sharaku.eorc.jaxa.jp) and the CMORPH v1.0 from NOAA CPC (ftp://ftp.cpc.ncep.noaa.gov/precip/CMORPH_V1.0/). Spatial resolutions and temporal frequency of the satellite datasets are listed in Table 1. To be consistent, IMERG and GSMaP data were accumulated to daily scale (08:00-08:00 of local time, i.e. UTC+8) and CMORPH was bilinearly interpolated to the grid resolution of 0.1°.

The merged dataset was further compared with two popular merged rainfall datasets of Climate Hazards Group InfraRed Precipitation with Stations (CHIRPS) (Funk et al., 2015) and Multi-Source Weighted-Ensemble Precipitation (MSWEP) (Beck et al., 2019). CHIRPS was originated by merging CHPClim, thermal infrared, TRMM3B42, NOAA CFSv2 precipitation data, and ground observation precipitation data. MSWEP was merged from multiple datasets including CPC, GPCC, CMORPH, GSMaP-MVK, GPM-IMERG, ERA5, and JRA-55. CHIRPS and MSWEP showed great potentials in rainfall estimates in previous studies (Liu et al., 2019).

Table 1. Multiple satellite precipitation datasets used in this study.

Datasets	Resolution	Frequency	Source	Reference
GPM IMERG	0.1 °x 0.1 °	0.5 hourly	NASA	(Hou et al., 2014)
GSMaP_Gauge	0.1 °x 0.1 °	1 hourly	JAXA	(Ushio et al., 2009)
CMORPH v1.0	0.25 °x 0.25 °	daily	CPC	(Joyce et al., 2004)
CHIRPS v2.0	0.25 °x 0.25 °	daily	USGS and CHC	(Funk et al., 2015)
MSWEP v2	0.1 °x 0.1 °	3 hourly	-	(Beck et al., 2019)

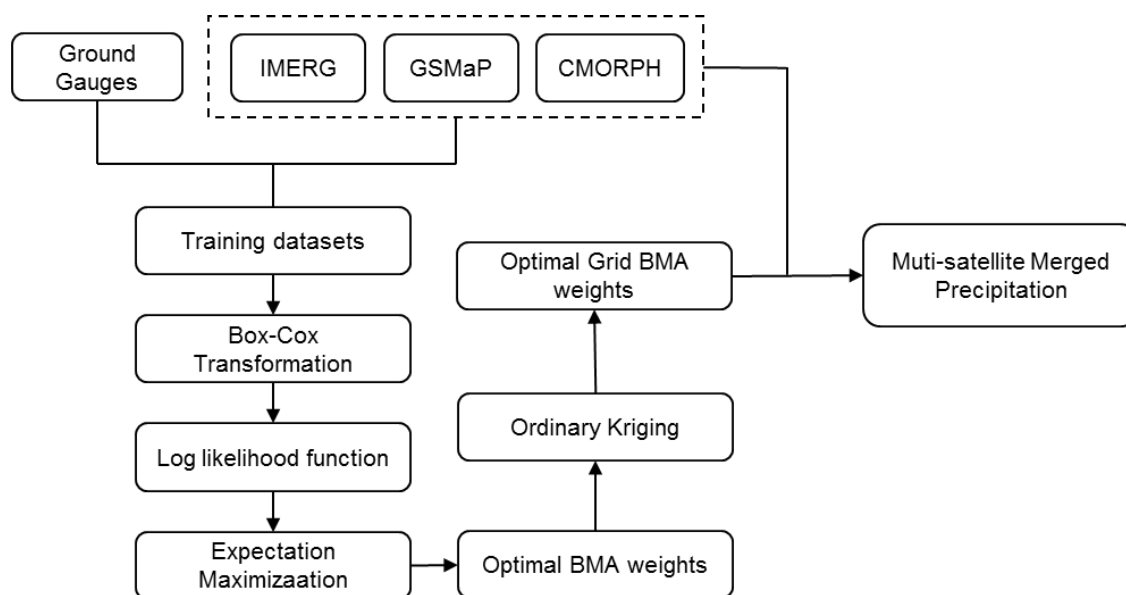
3. Methodology

We used the Dynamic Bayesian Model Averaging (DBMA) method (Ma et al., 2017) to merge the satellite datasets with in-situ rain gauges. To evaluate the quality of the new dataset, we carried out statistical and hydrological evaluations and comparisons

155 with CHIRPS and MSWEP in southern TP.

3.1. Dynamic Bayesian Model Averaging method

157 The Dynamic Bayesian Model Averaging (DBMA) method developed by Ma et
158 al. (2018) was utilized in this work. A flow chart of the merging method is shown in
159 Figure 2. In the first step, a training dataset was formed by selecting samples from the
160 ground gauged data and three original satellite datasets. The training period was set as
161 40 days. Increasing the length of the training period did not lead to obvious
162 improvement of the merging method (Ma et al., 2018). In the second step, the training
163 dataset was transformed by the Box-Cox Gaussian distribution, and the optimal weights
164 for each of the original satellite datasets on a specific grid where a ground gauge is
165 located on each training day were estimated by a logarithmic likelihood equation and
166 the optimal expectation algorithm. In the third step, an ordinary Kriging interpolation
167 method was applied to spatially interpolate the daily weights onto grids with no gauges.
168 Finally, posterior spatiotemporal weights were used to obtain the final merged rainfall
169 dataset. The DBMA-merged data has been proved in Ma et al. (2017) to outperform
170 original satellite data during 2007-2012 over TP.

171

172 Figure 2. Flowchart of the DBMA merging method (adapted from Ma et al., 2018).

173 For statistical evaluation of the merged data against ground gauges, around 85%

174   of the gauges were randomly selected to form a training gauge set for the merging

175   approach in each year during 2014-2019, and the remaining 15% were used for test.

176   Training method DBMA of 40 days was only conducted in training dataset. Table 2 lists

177   the numbers of training and test gauges in each of the warm seasons. The spatial

178   distributions of gauges in each year are presented in Figure S1. Data from all gauges

179   were involved in the training procedure of the final released version of the merged data.

180   Table 2. Number of rain gauges for training and test in 2014-2019.

Year	Total number of rain gauges	Number of training gauges	Number of test gauges
2014	195	166	29
2015	54	46	8
2016	373	317	56
2017	321	273	48
2018	377	320	57
2019	106	90	16

181   **3.2. Statistical Evaluation**

182   Performance of the multiple datasets were statistically evaluated by comparing

183   with ground observations on the corresponding statelite grids. Relative bias (RB) and

184   normalized root mean square error (RMSE) were adopted to measure the amount

185   difference between the gridded rainfall and the gauged rainfall. Correlation Coefficient

186   (CC) was used to evaluate the consistency between satellite estimates and gauge

187   observations. The skill of rainfall data on detecting rainfall occurrence (rainfall events

188   higher than zero) was evaluated through a set of metrics (similarly to Wilks, 2006): i.e.

189   the probability of detection (POD) assessing how good the multiple rainfall datasets are

190   at detecting the occurrence of rainfall, false alarm ratio (FAR) measuring how often the

191   gridded rainfall datasets detect rainfall when there actually is not rainfall, and critical

192   success index (CSI) measuring the ratio of rainfall events that are correctly detected by

193   the gridded datasets to the total number of observed or detected events. Equations for

194   the above metrics are shown in Table 3.

195 Table 3. Statistical indices that were used to assess the performance of the gridded rainfall

196 datasets.

Statistical Indicators	Equation	Optimal Value	Equation number
Relative Bias (RB)	$Bias = \dfrac{\sum_{i=1}^{n}(S_i - G_i)}{\sum_{i=1}^{n} G_i}$	0	(1)
Correlation Coefficient (CC)	$CC = \dfrac{[\sum_{i=1}^{n}(S_i - \bar{S}) \cdot (G_i - \bar{G})]^2}{\sum_{i=1}^{n}(S_i - \bar{S})^2 \cdot \sum_{i=1}^{n}(G_i - \bar{G})^2}$	1	(2)
Root Mean Square Error (RMSE)	$RMSE = \sqrt{\dfrac{1}{n}\sum_{i=1}^{n}(S_i - G_i)^2}$	0	(3)
Probability of Detection (POD)	$POD = \dfrac{a}{a + c}$	1	(4)
False Alarm Ratio (FAR)	$FAR = \dfrac{b}{a + b}$	0	(5)
Critical success index (CSI)	$CSI = \dfrac{a}{a + b + c}$	1	(6)

197 For the equations listed in Table 3, n is the total number of gridded product data

198 and gauge observation data; i is the ith of satellite product data and gauge

199 observation data; G_i means gauge observation and $\bar{G}$ is the average of gauge

200 observation. S_i and $\bar{S}$ are gridded estimates and their average, respectively. a

201 represents hit (i.e., event was detected to occur and observed to occur), b represents

202 false alarm (i.e., event was detected to occur but not observed to occur), and c

203 represents miss (i.e., event was not detected to occur but observed to occur).

204 Triple Collocation (TC) technique provides a platform for quantifying the root

205 mean square errors of three products that estimate the same geophysical variable

206 (Stoffelen, 1998). Roebeling et al. (2012) successfully applied the TC technique to

207 estimate errors of three rainfall products across Europe. An extended Triple Collocation

208 (ETC) introduced in Kaighin et al. (2014), which is able to estimate errors and

209 correlation coefficients with respect to an unknown target was used in this study to

210 compare the performance of the DBMA-merged data and two previous merged datasets

211   of CHIRPS and MSWEP.

3.3. Hydrological Evaluation

213   In addition to the statistical assessments against rain gauges, hydrological

214   assessment was used as a tool to test the performance of merged rainfall datasets on

215   forcing hydrological modelling in the study area (similarly see Yong et al, 2012; Xue

216   et al, 2013; Yong et al, 2014; Li et al, 2014). In this section, a semi-distributed

217   hydrological model developed by Tian (2006), namely Tsinghua Hydrological Model

218   based on Representative Elementary Watershed (THREW), was adopted for the

219   hydrological assessment of rainfall datasets in the YZRB. YZRB has a drainage area of

220   approximately 240,480 km^2 within China's boarder. The basin elevation ranges from

221   143 to 7,261 m, with an average of around 4,600 m. YZR is one of the most important

222   transboundary rivers in South Asia and the highest river in the world, which is

223   characterized by a dynamic fluvial regime with exceptional physiographic setting

224   spreading along the eastern Himalayan region (Goswami, 1985). Due to complex

225   terrain and strongly varying elevation, the YZRB is under control of a variety of climate

226   systems, such as the semi-arid plateau climate prevailing in the upper and middle

227   reaches, and the mountainous subtropical and tropical climates prevailing in the lower

228   reaches. In the cold upper reaches, the mean annual rainfall is less than 300 mm. In the

229   warm middle reaches, the mean annual rainfall falls between 300 mm and 600 mm.

230   The whole basin area above the Nuxia hydrological station was divided into 63

231   Representative Elementary Watersheds (REWs). Model parameters were calibrated by

232   daily discharges measured at the Nuxia station. The calibration period is scheduled to

233   run in the warm seasons from June 10th to October 31st in 2014- 2017, encompassing a

234   period length of 576 days. The validation period includes two warm seasons in 2018

235   and 2019 with a total duration of 288 days. Descriptions of the calibrated model

236   parameters can be found in Table 4. An automatic algorithm pySOT developed by D.

237   Eriksson et al (2019) was used to optimize the parameter values based on an objective-

238 function of *NSE* (Nash and Sutcliffe, 1970) in Eq. 7. To conduct a continuous

239 hydrological simulation in the study period, the datasets of daily grid-based

240 precipitation over China (Zhao et al., 2014) were used as model inputs in the non-warm

241 seasons when merged rainfall is not available.

242
Table 4. Calibrated parameters of the THREW model.

Symbol	Description	Unit	Value Range
kv	Fraction of potential transpiration rate over potential evaporation	-	0.001-0.8
n^t	Manning roughness coefficient for hillslope	-	0.0001-0.2
GaIFL	Spatial heterogeneous coefficient for infiltration capacity	-	0.0001-0.7
GaEFL	Spatial heterogeneous coefficient for exfiltration capacity	-	0.0001-0.7
GaETL	Spatial heterogeneous coefficient for evapotranspiration capacity	-	0.0001-0.7
WM	Tensor water storage capacity	cm	0.1-10
B	Shape coefficient to calculate the saturation excess runoff area	-	0.01-1
Gaus	Coefficient representing spatial heterogeneity of exchange term between t-zone and r-zone	-	0.001-10
KKA	Exponential coefficient to calculate subsurface flow	-	0.01-6
KKD	Linear coefficient to calculate subsurface flow	-	0.001-0.5
MM	Snow melting degree-day factor	mm/day	0.001-10
MMG	Ice melting degree-day factor	mm/day	0.001-10
C1+C2	Muskingum parameter	-	0.0001-1
C1/(C1+C2)	Muskingum parameter	-	0.0001-1

$$NSE = 1 - \frac{\sum_{n=1}^{N}(Q_{obs}^n - Q_{sim}^n)^2}{\sum_{n=1}^{N}(Q_{obs}^n - \bar{Q}_{obs})^2} \quad (7)$$

243 where, N is the total number of days in the evaluation period, Q_{obs}^n and Q_{sim}^n

244 represent the observed and simulated runoff on the nth day, respectively. $\bar{Q}_{obs}$

245 represents the average of observed runoff in the evaluation period.

4. Results and Discussions

4.1. Spatiotemporal Patterns

Based on the merging method, a new daily rainfall dataset with spatial resolution of 0.1°×0.1° in the warm seasons from June 10th to October 31st (144 days in each year) in 2014-2019 (864 days in six years) was generated. Figure 3 presents the spatial pattern of the mean rainfall over the six warm seasons of the merged data in southern TP. It is shown that extremely high summer rainfall centres concentrate in the south-eastern and south-western of the study area where is known as a world-famous heavy rainfall centre (see Biskop et al., 2015; Bookhagen & Burbank, 2006; Kumar et al., 2010).

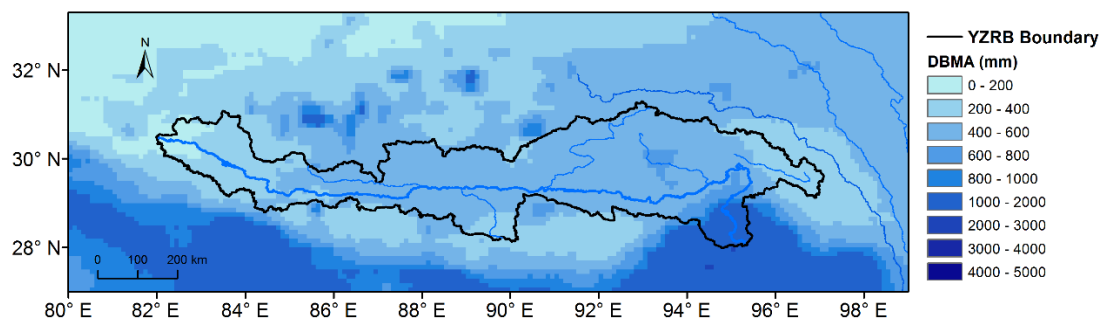
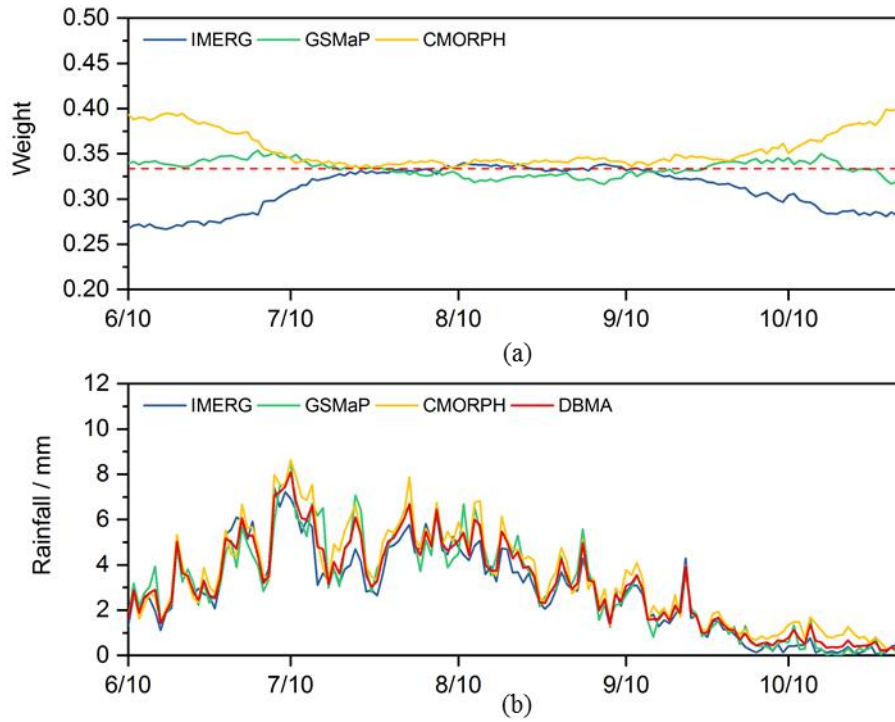



Figure 3. Spatial pattern of mean rainfall over six warm seasons in 2014-2019 of the DBMA-merged data in southern TP.

In addition, Figure 4 compares the time series of average daily weight and rainfall over the YZRB basin derived from the DBMA-merged data and the original satellite datasets. As expected, the DBMA-merged daily rainfall in general fall in the envelope ranges of the three satellite datasets. Merged data is closer to CMORPH in June, September and October, while showing equal closeness to all the three source satellite data in July and August. It indicates that CMORPH is closer to the in-situ gauges than IMERG at basin scale when rainfall value is small, especially for light rainfall events smaller than 2 mm, but this difference tends to be small for heavy rainfall events.

266

267 Figure 4. Seasonal variations in basin-averaged (a) weights and (b) rainfall estimates of the

268 multiyear daily values of IMERG, GSMaP, CMORPH and DBMA.

269 **4.2. Statistical Evaluation**

270 Figure 5 shows the statistical evaluation of the merged and original datasets in the

271 warm seasons. The statistical indices were calculated for three gauge groups including

272 the training gauges, the test gauges and all gauges at different elevation bands. The

273 datasets in general presented comparable performance for the training and test gauge

274 groups, indicating that the sampling procedure of ground gauges is adequately random.

275 The comparable performance of merged data in training and test gauge groups

276 demonstrated robustness of the merging method on varying gauges. In terms of RSME,

277 CC, and POD, the DBMA-merged data shows much better performance on all gauge

278 groups and elevation bands than the original satellite datasets. The smallest RSME of

279 merged data indicate that the total rainfall amount of the merged data during the

280 evaluation period showed the lowest difference from the total amount of gauged rainfall.

281 The highest CC and POD highlight the best consistency between merged data and

282   ground gauge data on days when most regions in the basin were rainy. The RB of

283   DBMA-merged data is at an intermediate level among the satellite datasets as it is the

284   weighted average of those three datasets. The higher FAR and lower CSI of DBMA-

285   merged data could be attributed to that the merging method detected rainfall events

286   when rainfall estimate is higher than zero in any one of the three satellite datasets and

287   thus resulted in overestimated rainfall occurrence. The overestimated rainfall

288   occurrence might have small effects on the estimation of rainfall amount, as most of

289   the falsely alarmed events were tiny. It is noteworthy that the performance of the

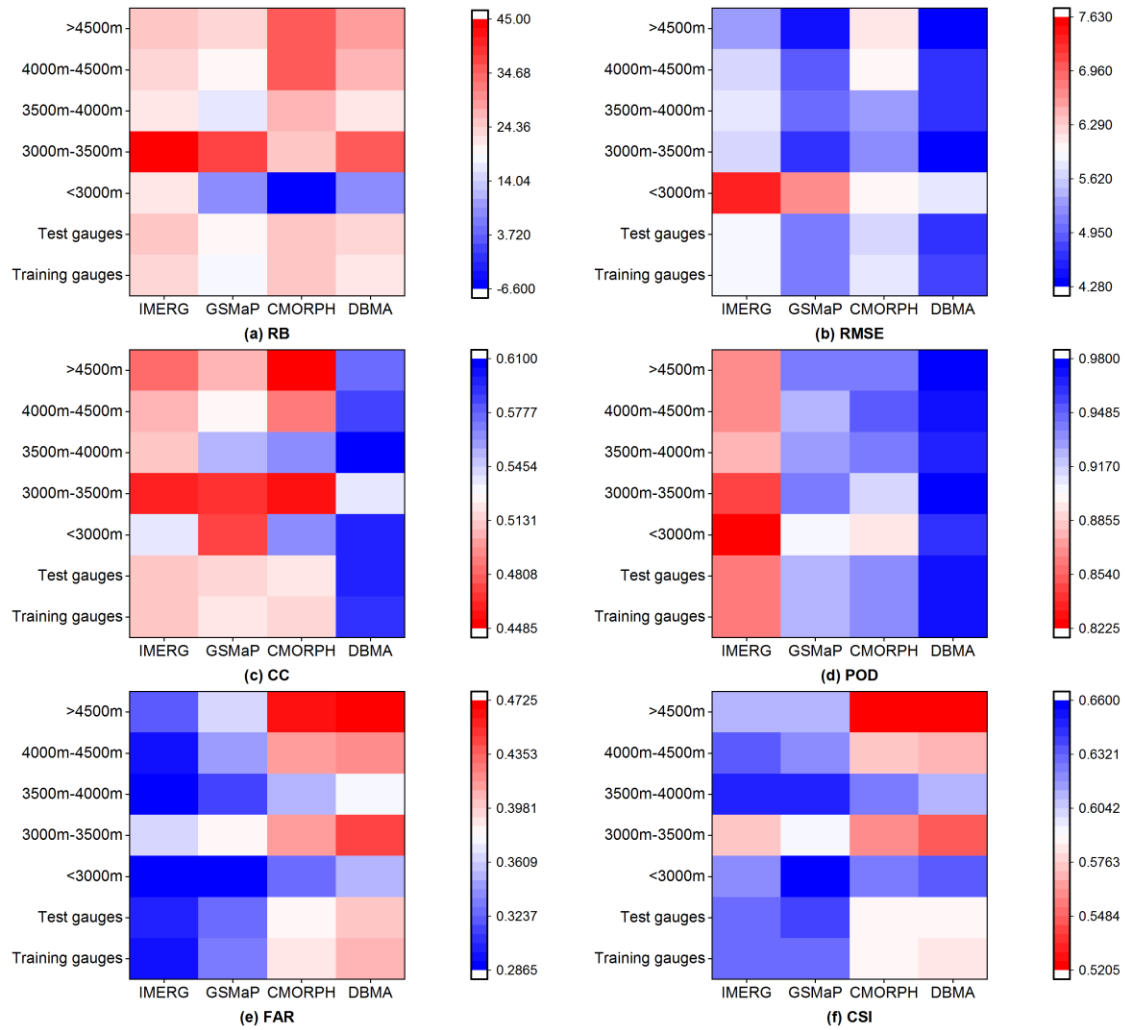
290   merged data shows smaller variance across elevation bands than that of the original

291   satellite datasets. This is most likely benefiting from the spatially dynamic optimal

292   weights for the original satellite data. However, the merged data presented the largest

293   difference from gauged data at the altitudes of 3000-3500 m, because there are much

294   less gauges on this elevation zone.

295

296   Figure 5. Comparisons of the statistical indices of (a) RB, (b)RMSE, (c) CC, (d) POD, (e) FAR

297   and (f) CSI for training gauges, test gauges and all gauges at five elevation bands.

298   Figure 6 shows CC of different datasets on specific gauges. The merged data

299   presents higher CC values in regions where are densely gauged, i.e., the middle reaches

300   of YZRB and the east part of the study region, which can be expected as the dense

301   ground gauges provided strongly informative benchmark likelihoods for the estimation

302   of satellite data weights. On most of the gauges (Figure 6a), the merged-data presented

303   higher CC values than the IMERG data, which is consistent with Figure 5c. On contrary,

304   the merged-data showed reduced CC than GSMaP and CMORPH on more gauges

305   (Figures 6b-c), indicating that involving IMERG data in the merging procedure on these

306   gauges lead to deteriorated consistence performance.

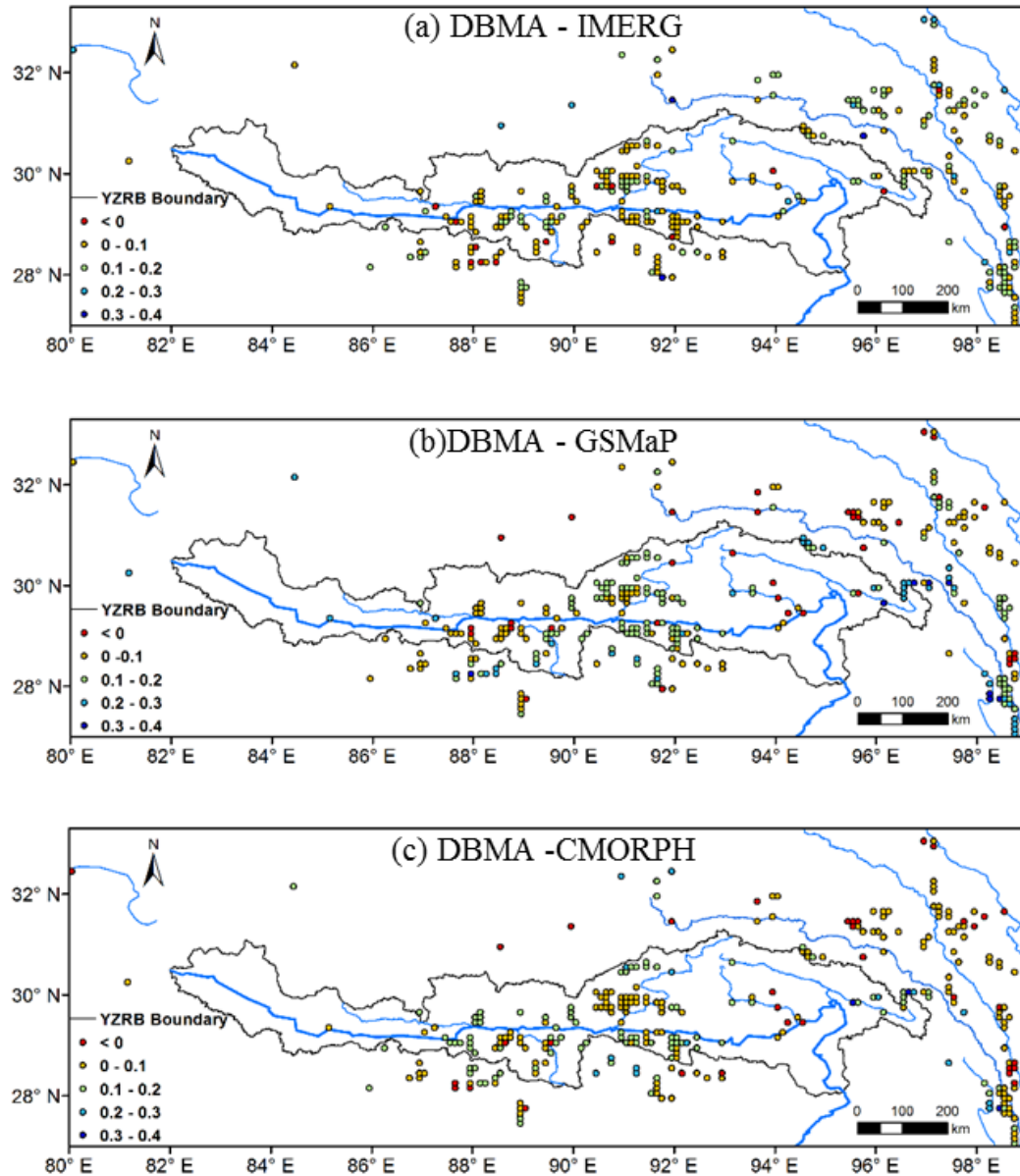



Figure 6. Spatial distributions of CC difference between (a) DBMA and IMERG, (b) DBMA and GSMaP, (c) DBMA and CMORPH

4.3. Hydrological Evaluation

(a) Hydrological simulation

Performance of the THREW model forced by different rainfall datasets are compared in Table 5. The DBMA-merged dataset achieved the best runoff simulation among all rainfall inputs, with NSE reaching 0.93 and 0.86 in calibration and validation period, respectively, indicating an excellent agreement between simulated and observed

316 hydrographs. Both IMERG and GSMaP underestimated the measured daily discharge,

317 but the DBMA-merged dataset improved such underestimations (see *RB* values in Table

318 4).

Table 5. Evaluation metrics of hydrological simulations forced by IMERG, GSMaP, CMORPH and DBMA.

Parameters	IMERG	GSMaP	CMORPH	DBMA
NSEcal	0.91	0.90	0.90	0.93
NSEval	0.75	0.57	0.81	0.86
RB	-0.07	-0.10	0.02	-0.05

321 (b) Uncertainty analysis

322 The automatic algorithm pySOT was ran 200 times to investigate the modelling

323 uncertainty caused by parameter calibration. Figure 7 presents the distributions of *NSE*

324 values estimated by the ensemble parameter sets of the merged and original rainfall

325 forces. It is shown that streamflow simulated by the DBMA data at the Nuxia station

326 presented higher NSEs and smaller uncertainty ranges than that simulated by the

327 original satellite datasets, indicating that streamflow simulations driven by the merged

328 dataset showed stronger robustness and were less affected by uncertainty of parameter

329 calibration.

330 In addition to the Nuxia hydrological station, model performance on simulating

331 streamflow at the interior hydrological stations of Yangcun, Nugesha, Gongbujiangda

332 and Lhasa (Figure 1) were evaluated in Figure 7. It shows that the IMERG forced

333 simulations presented poor NSE outliers lower than zero at the Lhasa station, in spite

334 of their good performance at the Yangcun and Nugesha stations; the GSMaP forced

335 simulations presented large uncertainty ranges in calibration period at Nugesha and

336 Lhasa, and in validation period at Nuxia and Gongbujiangda; the CMORPH forced

337 simulations showed the worst performance in validation period at the interior

338 hydrological stations, despite their sound good performance in calibration period at

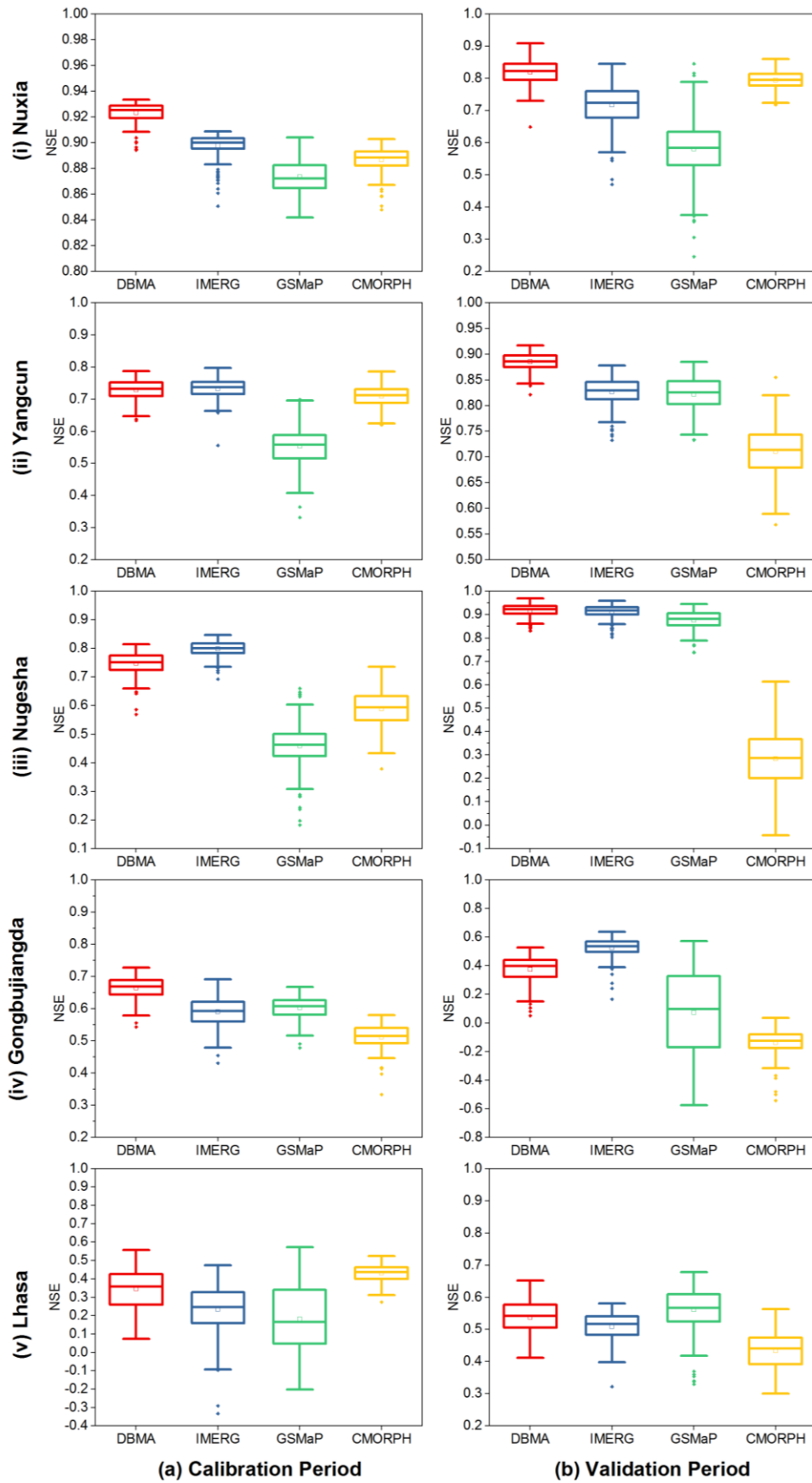
339 Yangcun and Nugesha. In comparison to the satellite datasets, the DBMA forced

340 simulations tend to perform consistently better with smaller uncertainties at all the

341   hydrological stations, which can be attributed to that the merged data incorporated the

342   advantages of different datasets in different regions and temporal periods and thus

343   better captured the spatial variability of rainfall inputs in sub-basins.

344

345   Figure 7. Runoff simulations at Nuxia, Yangcun, Nugesha, Gongbujiangda and Lhasa stations

346   forced by multiple rainfall inputs.

4.4. Comparisons with other datasets

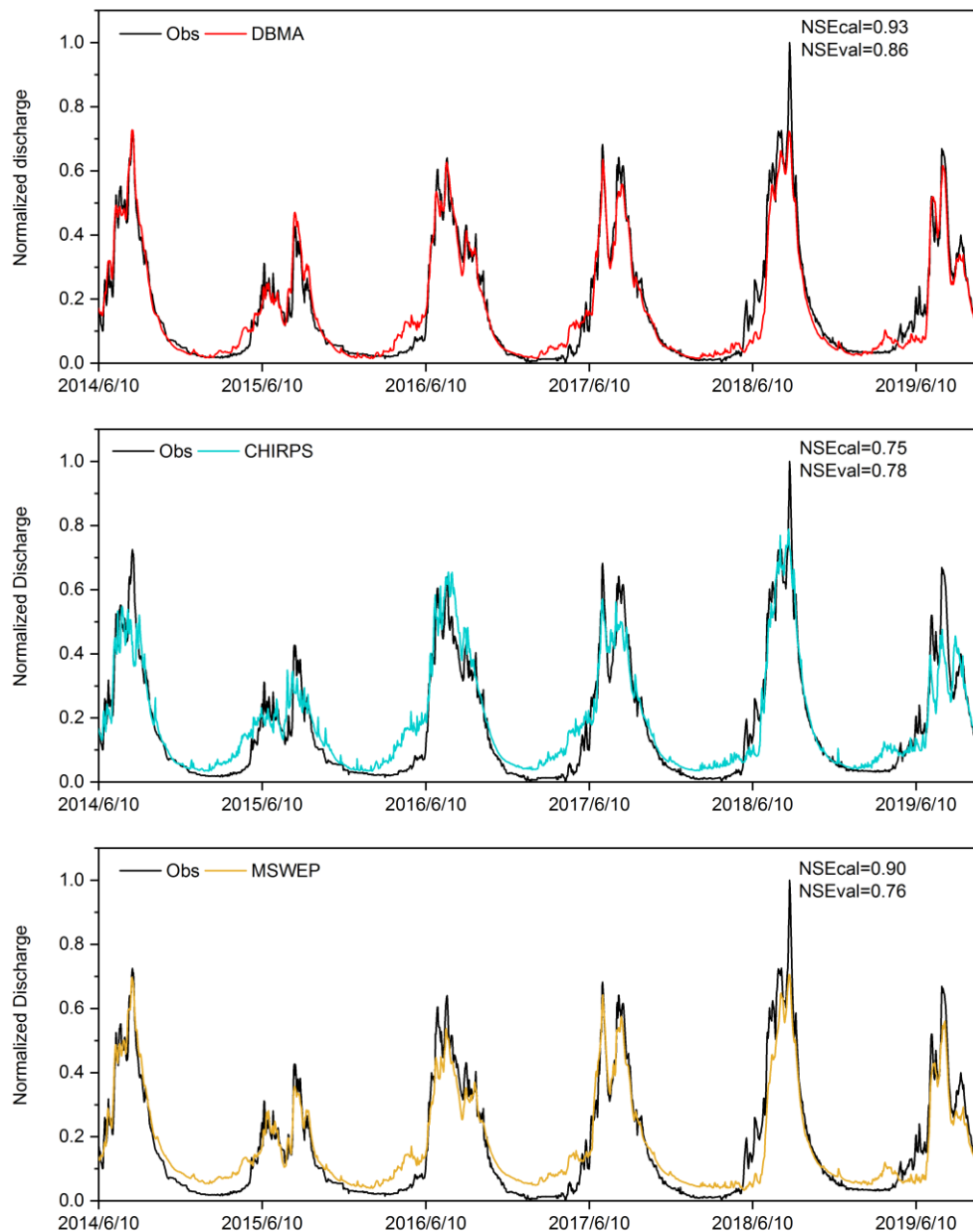
To avoid interference of ground gauge data that merged in the DBMA dataset, the ETC method introduced in Section 3.2 was applied to compare the three merged datasets in Table 6. The RMSE and CC of DBMA calculated by ETC were 1.11 and 0.80, respectively, both of which are obviously superior compared to the corresponding values estimated by CHIRPS and MSWEP, indicating that DBMA data is closer to the true value of rainfall in the study region.

Table 6. Statistical RMSE and CC of merged datasets calculated by the ETC method.

Datasets	DBMA	CHIRPS	MSWEP
RMSE-ETC	1.11	7.15	2.82
CC-ETC	0.80	0.28	0.62

Runoff simulations forced by the three merged datasets during June 10th 2014 to October 31st 2019 estimated by the corresponding optimal parameter sets were presented in Figure 8. Note that the daily runoff is normalized as Eq. 8 for data security reasons. Simulation by the CHIRPS data presented the lowest performance with NSE values of 0.75 and 0.78 in the calibration and validation periods, respectively. The DBMA forced simulation showed the highest performance with NSE values of 0.93 and 0.86 in the calibration and validation periods, followed by the MSWEP forced simulation which estimated NSE values of 0.9 in the calibration period and 0.76 in the validation period. The performance of streamflow forced by the merged datasets are consistent with the agreements between the merged rainfall estimates and ground truth shown in Table 6.

$$Q_{Normalized}^{n} = \frac{Q_{sim}^{n} - \min(Q_{obs})}{\max(Q_{obs}) - \min(Q_{obs})} \tag{8}$$

366

367   Figure 8. Simulated daily runoff at Nuxia station forced by DBMA, CHIRPS, and MSWEP.

5 Data Availability

369   The high-accuracy rain dataset by merging multi-satellite and dense ground

370   gauges over southern Tibetan Plateau for the warm seasons in 2014-2019 is freely

371   accessible at the National Tibetan Plateau Data Center

372   https://doi.org/10.11888/Hydro.tpdc.271303 (Li et al.,2021).

6. Summary

We collated ground-based rainfall observations from a dense gauge network over southern TP. The gauged data provides crucial ground references of measured rainfall. Based on this rain gauge network and three satellite rainfall datasets of IMERG, GSMaP, and CMORPH, a merged rainfall dataset in six warm seasons from June 10th to October 31st during 2014-2019 over the southern TP was established. The DBMA method was used to estimate weights varying in space and time of the three satellite datasets for the merged data. The merged rainfall dataset presented improved performance on representing the total amount of rainfall and detecting the occurrence of gauged rainfall events, and provide a more reliable forcing for hydrological simulations in the YZRB, compared to the original satellite datasets. Comparisons with previous merged rainfall datasets of CHIRPS v2.0 and MSWEP v2 that used relatively sparse rain gauges in the study area demonstrated high values of the newly installed rain gauges for providing robust ground reference for the merging of current satellite datasets. Our results indicated that the merged datasets can meet the critical needs of accurate forcing inputs for the simulations of warm season floods and the robustness calibration of hydrological models. Based on this high-accuracy rainfall data and reliable hydrological modelling, much further research in this region then could be conducted, for example, fluvial sediment transport modelling through coupling sediment and hydrology, validation and correction of precipitation from Global Climate Model, and future runoff projections based on reliable modelling calibration in history.

Author contribution

TF and LK designed the research. LK, XR, MY developed the approach and datasets. LK downloaded the datasets and performed most of the computation and analysis work. YL, HZ, LH and MY contributed to the revising of the paper.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

Ground gauge data from the hydrological bureau of MWR is acknowledged here.

Financial support

This research has been financially supported by the National Natural Science Foundation of China (grant no. 92047301) and National Key R&D Program of China (2018YFC1508102).

References

Anagnostou, E. N., & Zhang, X.: Evaluation of Numerical Weather Model–Based Satellite Precipitation Adjustment in Tropical Mountainous Regions, Journal of Hydrometeorology, 20(3), 431-445, http://doi.org/10.1175/JHM-D-18-0008.1, 2019.

Beck, H. E., Wood, E. F., Pan, M., Fisher, C. K., Miralles, D. G., van Dijk, A. I. J. M., Adler, R. F.: MSWEP V2 Global 3-Hourly 0.1° Precipitation: Methodology and Quantitative Assessment, Bulletin of the American Meteorological Society, 100(3), 473-500, http://doi.org/10.1175/bams-d-17-0138.1, 2019.

Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., Schamm, K., Schneider, U., & Ziese, M.: A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901–present, Earth System Science Data, 5(1),

419 71-99, http://doi.org/10.5194/essd-5-71-2013, 2013.

420 Biskop, S., Maussion, F., Krause, P., & Fink, M.: Differences in the water-balance

421   components of four lakes in the southern-central Tibetan Plateau, Hydrology and

422   Earth System Sciences, 20(1), 209-225, http://doi.org/10.5194/hess-20-209-2016,

423   2016.

424 Bitew, M. M., & Gebremichael, M.: Evaluation through independent measurements:

425   Complex terrain and humid tropical region in Ethiopia, Satellite rainfall

426   applications for surface hydrology (pp. 205-214), Springer, Dordrecht,

427   http://doi.org/ 10.1007/978-90-481-2915-7_12, 2010.

428 Bookhagen, B., & Burbank, D. W.: Topography, relief, and TRMM-derived rainfall

429   variations along the Himalaya, Geophysical Research Letters, 33(8),

430   http://doi.org/10.1029/2006gl026037, 2006.

431 David Eriksson, David Bindel, Christine A. Shoemaker.: pySOT and POAP: An event-

432   driven asynchronous framework for surrogate optimization. *arXiv* preprint

433   arXiv:1908.00420, 2019.

434 Derin, Y., Anagnostou, E., Anagnostou, M. N., Kalogiros, J., Casella, D., Marra, A. C.,

435   Sano, P.: Passive Microwave Rainfall Error Analysis Using High-Resolution X-

436   Band Dual-Polarization Radar Observations in Complex Terrain, IEEE

437   Transactions on Geoscience and Remote Sensing, 56(5), 2565-2586,

438   http://doi.org/10.1109/tgrs.2017.2763622, 2018.

439 Derin, Y., Anagnostou, E., Berne, A., Borga, M., Boudevillain, B., Buytaert, W.,

440   Yilmaz, K.: Evaluation of GPM-era Global Satellite Precipitation Products over

441   Multiple Complex Terrain Regions, Remote Sensing, 11(24),

442   http://doi.org/10.3390/rs11242936, 2019.

443 Dinku, T. C., Stephen J.; Ceccato, Pietro.: Comparison of CMORPH and TRMM-3B42

444   over mountainous regions of Africa and South America, Satellite rainfall

445   applications for surface hydrology (pp. 193-204): Springer, Dordrecht,

446   http://doi.org/ 10.1007/978-90-481-2915-7_11, 2010.

447 Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Michaelsen,

448   J.: The climate hazards infrared precipitation with stations--a new environmental

449   record for monitoring extremes, Scientific Data, 2, 150066,

450   http://doi.org/10.1038/sdata.2015.66, 2015.

451 Gao, Y. C., & Liu, M. F.: Evaluation of high-resolution satellite precipitation products

452   using rain gauge observations over the Tibetan Plateau, Hydrology and Earth

453   System Sciences, 17(2), 837-849, http://doi.org/10.5194/hess-17-837-2013, 2013.

454 Goswami, D. C.: Brahmaputra River, Assam, India: Physiography, basin denudation,

455   and channel aggradation, Water Resources Research, 21(7), 959-978,

456   http://doi.org/ 10.1029/WR021i007p00959, 1985.

457 He, J., Yang, K., Tang, W., Lu, H., Qin, J., Chen, Y., & Li, X.: The first high-resolution

458   meteorological forcing dataset for land process studies over China, Scientific data,

459   7(1), 1-11, http://doi.org/ 10.1038/s41597-020-0369-y, 2020.

460 He, Z., Yang, L., Tian, F., Ni, G., Hou, A., & Lu, H.: Intercomparisons of Rainfall

461   Estimates from TRMM and GPM Multisatellite Products over the Upper Mekong

462   River Basin, Journal of Hydrometeorology, 18(2), 413-430,

463   http://doi.org/10.1175/jhm-d-16-0198.1, 2017.

464 Hong, Y., Gochis, D., Cheng, J.-t., Hsu, K.-l., & Sorooshian, S.: Evaluation of

465   PERSIANN-CCS Rainfall Measurement Using the NAME Event Rain Gauge

466   Network, Journal of Hydrometeorology, 8(3), 469-482,

467   http://doi.org/10.1175/jhm574.1, 2007.

468 Hou, A. Y., Kakar, R. K., Neeck, S., Azarbarzin, A. A., Kummerow, C. D., Kojima, M.,

469   Iguchi, T.: The Global Precipitation Measurement Mission, Bulletin of the

470   American Meteorological Society, 95(5), 701-722, http://doi.org/10.1175/bams-d-

471   13-00164.1, 2014.

472 Hou, S., Tian, F., Yang, L., Hu, H., & Hou, A.: How Does the Evaluation of the GPM

473   IMERG Rainfall Product Depend on Gauge Density and Rainfall Intensity?

474   Journal of Hydrometeorology, 19(2), 339-349, http://doi.org/10.1175/jhm-d-17-

475   0161.1, 2018.

476   Huang, J., Kang, S., Zhang, Q., Jenkins, M. G., Guo, J., Zhang, G., & Wang, K.: Spatial

477   distribution and magnification processes of mercury in snow from high-elevation

478   glaciers in the Tibetan Plateau, Atmospheric Environment, 46, 140-146,

479   http://doi.org/10.1016/j.atmosenv.2011.10.008, 2012.

480   Immerzeel, W. W., Van Beek, L. P., & Bierkens, M. F.: Climate change will affect the

481   Asian   water   towers,   Science,   328(5984),   1382-1385,   http://doi.org/

482   10.1126/science.1183188, 2010.

483   Jiang, S., Ren, L., Hong, Y., Yong, B., Yang, X., Yuan, F., & Ma, M.: Comprehensive

484   evaluation of multi-satellite precipitation products with a dense rain gauge

485   network and optimally merging their simulated hydrological flows using the

486   Bayesian model averaging method, Journal of Hydrology, 452-453, 213-225,

487   http://doi.org/10.1016/j.jhydrol.2012.05.055, 2012.

488   Joyce, R. J., JE; Arkin, PA; Xie, PP.: CMORPH: A Method that Produces Global

489   Precipitation Estimates from Passive Microwave and Infrared Data at High Spatial

490   and Temporal Resolution, Journal of Hydrometeorology, 5(3), 487-503, 2004.

491   Kidd, C., & Levizzani, V.: Status of satellite precipitation retrievals, Hydrology and

492   Earth System Sciences, 15(4), 1109-1116, http://doi.org/10.5194/hess-15-1109-

493   2011, 2011.

494   Kumar, V., Jain, S. K., & Singh, Y.: Analysis of long-term rainfall trends in India.

495   Hydrological   Sciences   Journal,   55(4),   484-496,

496   http://doi.org/10.1080/02626667.2010.481373, 2010.

497   Liu, J., Shangguan, D., Liu, S., Ding, Y., Wang, S., & Wang, X.: Evaluation and

498   comparison of CHIRPS and MSWEP daily-precipitation products in the Qinghai-

499   Tibet Plateau during the period of 1981–2015, Atmospheric Research, 230,

500   http://doi.org/ 10.1016/j.atmosres.2019.104634, 2019.

501   Lu, D., & Yong, B.: Evaluation and Hydrological Utility of the Latest GPM IMERG

502   V5 and GSMaP V7 Precipitation Products over the Tibetan Plateau, Remote

503   Sensing, 10(12), http://doi.org/10.3390/rs10122022, 2018.

504   Ma, Y., Hong, Y., Chen, Y., Yang, Y., Tang, G., Yao, Y., Liu, R.: Performance of

505   Optimally Merged Multisatellite Precipitation Products Using the Dynamic

506   Bayesian Model Averaging Scheme Over the Tibetan Plateau, Journal of

507   Geophysical   Research: Atmospheres,   123(2),   814-834,

508   http://doi.org/10.1002/2017jd026648, 2018.

509   McColl, K. A., Vogelzang, J., Konings, A. G., Entekhabi, D., Piles, M., & Stoffelen, A.:

510   Extended triple collocation: Estimating errors and correlation coefficients with

511   respect to an unknown target, Geophysical Research Letters, 41(17), 6229-6236,

512   http://doi.org/10.1002/2014gl061322, 2014.

513   Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part

514   I – A discussion of principles, Journal of hydrology, 10, 282–290,

515   https://doi.org/10.1016/0022-1694(70)90255-6, 1970.

516   Parsons, D. B., Rasmussen, R. M., Dai, A., & Trenberth, K. E.: The Changing Character

517   of Precipitation, Bulletin of the American Meteorological Society, 84(9), 1205-

518   1218, http://doi.org/10.1175/bams-84-9-1205, 2003.

519   Ren, Z. H., Zhao, P., Zhang, Q., Zhang, Z. F., Cao, L. J., Yang, Y. R., & Chen, Z.:

520   Quality control procedures for hourly precipitation data from automatic weather

521   stations in China, Meteorol. Mon, 36(7), 123-132, 2010.

522   Roebeling, R. A., Wolters, E. L. A., Meirink, J. F., & Leijnse, H.: Triple Collocation of

523   Summer Precipitation Retrievals from SEVIRI over Europe with Gridded Rain

524   Gauge and Weather Radar Data, Journal of Hydrometeorology, 13(5), 1552-1566,

525   http://doi.org/10.1175/jhm-d-11-089.1, 2012.

526   Stoffelen, A.: Toward the true near-surface wind speed: Error modeling and calibration

527   using triple collocation, Journal of Geophysical Research: Oceans, 103(C4), 7755-

528   7766, http://doi.org/10.1029/97jc03180, 1998

529   Su, Y., Zhao, C., Wang, Y., & Ma, Z.: Spatiotemporal Variations of Precipitation in

530   China Using Surface Gauge Observations from 1961 to 2016, Atmosphere, 11(3),

531   http://doi.org/10.3390/atmos11030303, 2020.

532   Sun, R., Yuan, H., Liu, X., & Jiang, X.: Evaluation of the latest satellite–gauge

533   precipitation products and their hydrologic applications over the Huaihe River

534   basin, Journal of Hydrology, 536, 302-319,

535   http://doi.org/10.1016/j.jhydrol.2016.02.054, 2016.

536   Tang, G., Ma, Y., Long, D., Zhong, L., & Hong, Y.: Evaluation of GPM Day-1 IMERG

537   and TMPA Version-7 legacy products over Mainland China at multiple

538   spatiotemporal scales, Journal of Hydrology, 533, 152-167,

539   http://doi.org/10.1016/j.jhydrol.2015.12.008, 2016.

540   Tang, G., Clark, M. P., Papalexiou, S. M., Ma, Z., & Hong, Y.: Have satellite

541   precipitation products improved over last two decades? A comprehensive

542   comparison of GPM IMERG with nine satellite and reanalysis datasets, Remote

543   Sensing of Environment, 240, http://doi.org/10.1016/j.rse.2020.111697, 2020.

544   Tong, K., Su, F., Yang, D., & Hao, Z.: Evaluation of satellite precipitation retrievals and

545   their potential utilities in hydrologic modeling over the Tibetan Plateau, Journal of

546   Hydrology, 519, 423-437, http://doi.org/10.1016/j.jhydrol.2014.07.044, 2014.

547   Tian, F., Hu, H., Lei, Z., & Sivapalan, M.: Extension of the Representative Elementary

548   Watershed approach for cold regions via explicit treatment of energy related

549   processes, Hydrology and Earth System Sciences, 10(5), 619-644, http://doi.org/

550   10.5194/hess-10-619-2006, 2006.

551   Ushio, T., Sasashige, K., Kubota, T., Shige, S., Okamoto, K. i., Aonashi, K., Kawasaki,

552   Z.-I.: A Kalman Filter Approach to the Global Satellite Mapping of Precipitation

553   (GSMaP) from Combined Passive Microwave and Infrared Radiometric Data,

554   Journal of the Meteorological Society of Japan, 87A, 137-151,

555   http://doi.org/10.2151/jmsj.87A.137, 2009.

556   Wang, Y., Wang, L., Li, X., Zhou, J., & Hu, Z.: An integration of gauge, satellite, and

557   reanalysis precipitation datasets for the largest river basin of the Tibetan Plateau.

558   Earth System Science Data, 12(3), 1789-1803, http://doi.org/10.5194/essd-12-

559   1789-2020, 2020.

560   Wang, Z., Zhong, R., Lai, C., & Chen, J.: Evaluation of the GPM IMERG satellite-
561   based precipitation products and the hydrological utility, Atmospheric Research,
562   196, 151-163, http://doi.org/10.1016/j.atmosres.2017.06.020, 2017.

563   Wilks, D. S.: Statistical methods in the atmospheric sciences (Vol. 100). Academic press,
564   2011.

565   Wolff, D. B., Nelkin, E. J., Bolvin, D. T., Huffman, G. J., Adler, R. F., Gu, G., Stocker,
566   E. F.: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global,
567   Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales, Journal of
568   Hydrometeorology, 8(1), 38-55, http://doi.org/10.1175/jhm560.1, 2007.

569   Xu, R., Tian, F., Yang, L., Hu, H., Lu, H., & Hou, A.: Ground validation of GPM
570   IMERG and TRMM 3B42V7 rainfall products over southern Tibetan Plateau
571   based on a high-density rain gauge network, Journal of Geophysical Research:
572   Atmospheres, 122(2), 910-924, http://doi.org/10.1002/2016jd025418, 2017.

573   Xue, X., Hong, Y., Limaye, A. S., Gourley, J. J., Huffman, G. J., Khan, S. I., Chen, S.:
574   Statistical and hydrological evaluation of TRMM-based Multi-satellite
575   Precipitation Analysis over the Wangchu Basin of Bhutan: Are the latest satellite
576   precipitation products 3B42V7 ready for use in ungauged basins? Journal of
577   Hydrology, 499, 91-99, http://doi.org/10.1016/j.jhydrol.2013.06.042, 2013.

578   Yang, D., Li, Z., Gao, B., Jiao, Y., Hong, Y., & Xu, T.: Multiscale Hydrologic
579   Applications of the Latest Satellite Precipitation Products in the Yangtze River
580   Basin using a Distributed Hydrologic Model, Journal of Hydrometeorology, 16(1),
581   407-426, http://doi.org/10.1175/jhm-d-14-0105.1, 2015

582   Yang, S., Chen, M., Xie, P., Yatagai, A., Hayasaka, T., Fukushima, Y., & Liu, C.: A
583   Gauge-Based Analysis of Daily Precipitation over East Asia, Journal of
584   Hydrometeorology, 8(3), 607-626, http://doi.org/10.1175/jhm583.1, 2007.

585   Yao, T., Thompson, L. G., Mosbrugger, V., Zhang, F., Ma, Y., Luo, T., & Fayziev, R.:
586   Third pole environment (TPE). Environmental Development, 3, 52-64, 2012

587   Yilmaz, K. K., Vergara, H. J., Rodríguez-Sánchez, J.-P., Salerno, F., Sahlu, D.,

588   Nikolopoulos, E. I., Derin, Y.: Multiregional Satellite Precipitation Products

589   Evaluation over Complex Terrain, Journal of Hydrometeorology, 17(6), 1817-

590   1836, http://doi.org/10.1175/jhm-d-15-0197.1, 2016.

591   Yong, B., Chen, B., Gourley, J. J., Ren, L., Hong, Y., Chen, X., Gong, L.:

592   Intercomparison of the Version-6 and Version-7 TMPA precipitation products over

593   high and low latitudes basins with independent gauge networks: Is the newer

594   version better in both real-time and post-real-time analysis for water resources and

595   hydrologic extremes? Journal of Hydrology, 508, 77-87,

596   http://doi.org/10.1016/j.jhydrol.2013.10.050, 2014.

597   Yong, B., Hong, Y., Ren, L.-L., Gourley, J. J., Huffman, G. J., Chen, X., Khan, S. I.:

598   Assessment of evolving TRMM-based multisatellite real-time precipitation

599   estimation methods and their impacts on hydrologic prediction in a high latitude

600   basin, Journal of Geophysical Research: Atmospheres, 117(D9),

601   http://doi.org/10.1029/2011jd017069, 2012.

602   Zhai, P., Zhang, X., Wan, H., & Pan, X.: Trends in total precipitation and frequency of

603   daily precipitation extremes over China. Journal of climate, 18(7), 1096-1108,

604   http://doi.org/ 10.1175/JCLI-3318.1, 2005.

605   Zhao, Y., Zhu, J., & Xu, Y.: Establishment and assessment of the grid precipitation

606   datasets in China for recent 50 years, J. Meteor. Sci, 34(4), 414-420, 2014.