Drought and Flood Characterization and Connection to Climate Variability in the Pearl River Basin in Southern China Using Long-Term GRACE and Reanalysis Data

ZHIYONG HUANG,[a,b,c] JIU JIMMY JIAO,[a,b,c] XIN LUO,[a,b,c] YUN PAN,[d,e] AND TAOYONG JIN[f]

[a] *Department of Earth Sciences, The University of Hong Kong, Hong Kong, China*
[b] *The University of Hong Kong–Shenzhen Research Institute, Shenzhen, China*
[c] *The University of Hong Kong–Zhejiang Institute of Research and Innovation, Hangzhou, China*
[d] *Beijing Laboratory of Water Resources Security, Capital Normal University, Beijing, China*
[e] *College of Resource Environment and Tourism, Capital Normal University, Beijing, China*
[f] *School of Geodesy and Geomatics, Wuhan University, Wuhan, China*

(Manuscript received 11 May 2020, in final form 2 December 2020)

ABSTRACT: Drought and flood are investigated in the Pearl River basin (PRB) using long-term terrestrial water storage anomaly (TWSA) data from the mascon (mass concentration) solutions based on Gravity Recovery and Climate Experiment (GRACE) satellite measurements (2002–19) and reanalysis data (1980–2019). The GRACE mascon solutions capture two major drought periods (2003–06 and 2009–12) with similar onsets and endings over the last two decades, but show considerable differences in quantifying total drought severity. The reanalysis data significantly overestimate drought duration and severity during 1980–2000 owing to overestimated negative TWSA forced by underestimated precipitation. The GRACE mascon solutions identify four major flood events in August 2002, June 2008, and July in 2006 and 2019. The flood potential is influenced by the precipitation in both the current and antecedent months. The flood potential index of the most recent flood in 2008 showed a similar spatial pattern compared to precipitation at monthly and subbasin scales. The precipitation and TWSA in the PRB are mainly influenced by El Niño–Southern Oscillation (ENSO). TWSA exhibits a lag of 1–3 months responding to ENSO during 1980–2019. This study emphasizes the significance of removing water storage changes in new large reservoirs before long-term drought and flood characterization. The inclusion of reservoir water storage would expand (shrink) the drought duration and overestimate (underestimate) drought severity for the period before (after) reservoir impoundment and overestimate flood potential for the period after reservoir impoundment. This study highlights the intensifying drought conditions in the PRB over the last four decades under the circumstances of more frequent human activities (reservoir construction and regulation) and the complex changing climate system.

KEYWORDS: Drought; Flood events; ENSO; Pacific decadal oscillation; Remote sensing; Reanalysis data

1. Introduction

Monitoring drought and flood is necessary to forecast the development of hazards and to assist in planning strategies for the prevention and mitigation of the negative impact of the hazards. Traditionally, drought and flood monitoring rely on in situ measurement of hydrometeorological variables like rainfall, soil moisture, and streamflow. However, these in situ measurements are limited to point and local scales and require lots of investment of labor, money, and time. With the development of remote sensing technology, an unprecedented breakthrough at spatiotemporal scales has been achieved for drought and flood monitoring. The advantages of remote sensing relative to traditional ground observations include consistent data records, improved spatial resolution, and global, near-real-time observations (AghaKouchak et al. 2015; Heumann 2011). However, traditional satellite-based optical sensors retrieve high-quality data susceptible to weather effects. Radar sensors do not contain spectral features, and there is scarce dense spatiotemporal data sampling in many regions (Sun et al. 2017). Among the many satellite missions, the Gravity Recovery and Climate Experiment (GRACE; see the appendix for expansions of acronyms) twin-satellite mission is the only one that can acquire total terrestrial water storage anomaly (TWSA) information at a monthly and even denser temporal scale despite weather effect (Tapley et al. 2004). Until now, GRACE satellite data have been widely used to detect regional or basin-scale drought and flood events worldwide, including for the Amazon River basin (Chen et al. 2009, 2010), La Plata basin (Abelen et al. 2015), Yangtze River basin (Sun et al. 2017, 2018; Z. Zhang et al. 2015; Zhou et al. 2017), Liao River basin (X. Chen et al. 2018), and the southwest karst plateau (Long et al. 2014) in China.

Drought characterization is generally performed using various drought indices, which are effectively continuous functions of precipitation and other hydrometeorological variables (Morid et al. 2006). To date, more than 100 drought indices have been proposed (Zargar et al. 2011). Among the reported drought indices, the standardized runoff index (Shukla and Wood 2008), the standardized precipitation index (McKee 1993), and the Palmer drought severity index (Palmer 1965) are the most widely used indices. In recent years, GRACE TWSA data have been integrated into the research on drought

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/JCLI-D-20-0332.s1.

Corresponding author: Jiu Jimmy Jiao, jjiao@hku.hk

DOI: 10.1175/JCLI-D-20-0332.1

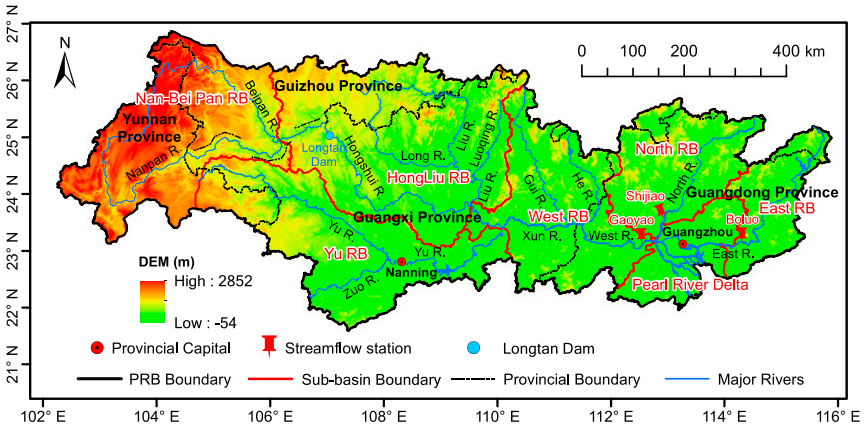

FIG. 1. Study region and its digital elevation model (DEM). Also shown is the location of streamflow stations and the Longtan Dam.

indices. Yirdaw et al. (2008) initially introduced the total storage deficit index, which is a renaming of the soil moisture deficit index developed by Narasimhan and Srinivasan (2005). Thomas et al. (2014) designed a GRACE-derived water storage deficit (WSD) for measuring hydrological drought occurrence and severity. Sinha et al. (2017) developed Thomas's method into the water storage deficit index (WSDI), which has been used by Sun et al. (2018) for drought evaluation in the Yangtze River basin, China. Zhao et al. (2017) designed a new GRACE-based drought severity index that makes it capable of comparing drought features across different regions and periods. Hosseini-Moghari et al. (2019) introduced a modified total storage deficit index using the residual signal rather than the total signal. Sinha et al. (2019) proposed a novel combined climatologic deviation index by combining precipitation anomalies and TWSA.

The world is facing higher flood risk under climate change and socioeconomic development (Hirabayashi et al. 2013; Winsemius et al. 2016). However, there have been limited studies on flood risk indexes since it is difficult to capture flood due to its relatively small extent and short duration (Sun et al. 2017). To highlight a flood-potential application, Reager and Famiglietti (2009) designed a monthly global flood index, namely Reager's flood potential index (FPI). This index utilizes the difference between GRACE-derived TWS and precipitation to explore flood potential. The index has been used by other researchers for flood detection. For example, Long et al. (2014) examined the severe flood in 2008 in Yun-Gui Plateau in southwestern China using Reager's FPI. Molodtsova et al. (2016) evaluated the efficacy of the Reager's FPI for flood risk assessment over the continental United States and found consistency between Reager's FPI flood risks and the observed floods on regional and even local scales. Sun et al. (2017) found that GRACE data are very useful for monitoring large floods in the Yangtze River basin. Idowu and Zhou (2019) assessed the flood potential in the lower Niger River basin in Nigeria and found that the GRACE-based FPI agrees well with the water budget–based FPI and with the flood reported by the Dartmouth Flood Observatory.

Drought and flood are influenced by climate variability like El Niño–Southern Oscillation (ENSO) and the Indian Ocean dipole (IOD). Understanding the relationship between TWSA (or drought and flood) and climate variability is essential for better management of water resources at a regional or basin scale. Phillips et al. (2012) and Ni et al. (2018) investigated the connection between TWS and ENSO on a global scale. Forootan et al. (2019) explored the global hydrological drought and connection to ENSO, the IOD, and the North Atlantic Oscillation (NAO). Awange et al. (2014) analyzed TWSA and assessed the influence of anthropogenic activities as well as climate variability (ENSO and IOD) in the Nile Basin. Z. Zhang et al. (2015) used GRACE-based TWSA to study the 2006 summer drought and 2011 spring drought in the Yangtze River basin in China and found a connection between drought and ENSO. Ndehedehe et al. (2017) examined the connection between TWS and three global climate indices—the Atlantic multidecadal oscillation (AMO), IOD, and ENSO over West Africa.

In addition to GRACE data, the development of the data-assimilation-based reanalysis system offers an alternative way to monitor and characterize drought and flood as well as the climatic diagnosis. Previous studies (e.g. Awange et al. 2016; Ndehedehe et al. 2018) have used reanalysis data like the MERRA-2 (Modern-Era Retrospective Analysis for Research and Applications, version 2) TWSA to characterize drought and flood. However, both GRACE and MERRA-2 reanalysis data have uncertainty owing to the different strategies of GRACE data processing centers and the bias in the long-term forcing data in the reanalysis system. Therefore, an evaluation of the reliability of both GRACE data from different centers and the reanalysis data from a long-term perspective is necessary for an accurate drought and flood characterization and teleconnection analysis.

The Pearl River basin (PRB) in southern China (Fig. 1) is the second-largest river regarding the average annual runoff in China (Niu 2013; Pearl River Water Conservancy Commission 2005). Droughts and floods occur frequently in the PRB because of the high spatiotemporal variation of precipitation/runoff (Cui et al. 2007; Niu 2010). A previous study (e.g., Luo

et al. 2016) analyzed the TWSA in the PRB. However, Luo et al. (2016) did not focus on the drought and flood in this basin based on GRACE data. Han et al. (2019) explored the propagation from meteorological drought to groundwater drought considering the impact of climate variability. However, they used the output from the Global Land Data Assimilation System (GLDAS) for the isolation of groundwater storage components from TWSA. As GLDAS does not include reservoir storage (RESS) changes, their estimates of groundwater drought may be biased. Besides, they focused on a relatively short period of GRACE data (2002–15), which is mostly a dry period. Therefore, their characterization of groundwater drought may be even more biased relative to the long-term time series. Considering the research gap in the above literature, and the significance of drought and flood studies for basin-scale water resource management, this study aims at monitoring and characterizing drought and flood using WSDI and FPI based on long-term TWSA from GRACE (2002–19) and a reanalysis product (MERRA-2; 1980–2019) in the PRB. The reliability of GRACE TWSA and long-term reanalysis data will be evaluated using a water balance approach, long-term observed precipitation data, and an ancillary drought index before drought and flood characterization. The teleconnection between TWSA and climate indices will be investigated. The contribution of RESS changes to drought and flood characterization and teleconnection analysis will be quantified and discussed.

2. Study area and data

a. Study area

The PRB (Fig. 1; 442 748 km^2) is located in southern China (102°12′–115°54′E, 21°32′–26°55′N). The Pearl River originates from the Yunnan Plateau and flows eastward through mountains and country to the South China Sea. It has three main subbasins, namely the West River [2075 km; drainage area: 341 878 km^2, including Nan-Bei Pan River basin (RB), Hongliu RB, Yu RB, and West RB], the North River (468 km; drainage area: 46 834 km^2), and the East River (520 km; drainage area: 27 611 km^2) (Zhang et al. 2009).

The PRB has a tropical to subtropical climate with an annual mean temperature of 14°–22°C, relative humidity of 71%–80%, and annual average precipitation of 1200–2000 mm (Liu et al. 2012). The wet season is April–September, accounting for 72%–88% of the annual total precipitation. The PRB receives 6–7 times more annual precipitation in the wet years than in dry years, resulting in the occurrence of flood or drought disasters (Liu et al. 2012).

b. GRACE TWSA data

This study uses three level-3 mascon (mass concentration) solutions from GRACE and its Follow-On mission, including the Release 06 mascons from Center for Space Research (CSR) at The University of Texas at Austin and Jet Propulsion Laboratory (JPL), and the mascon from the Goddard Space Flight Center (GSFC). These three mascon solutions deal with C20 (degree 2 order 0) and the degree-1 coefficients (geocenter) similarly. The C20 coefficients are replaced with the solutions from Satellite Laser Ranging (Cheng et al. 2011, 2013) and the degree-1 coefficients are estimated using the method from Swenson et al. (2008). A glacial isostatic adjustment correction has been applied for JPL mascon based on the ICE6G-D model from Peltier et al. (2018), for CSR mascon based on the model from A et al. (2012), and for GSFC mascon based on the model from ICE5G (VM2) (Peltier 2004). The CSR RL05 mascon is longitude–latitude grids at 0.25°, but they represent the equal-area geodesic grid at 1° × 1° at the equator (i.e., the current native resolution; Save et al. 2016). JPL RL05 mascons are longitude–latitude grids at 0.5°, but they represent the 3° × 3° equal-area caps (i.e., the current native resolution; Watkins et al. 2015). The GSFC mascons are estimated with 10-day and 1-arc-degree equal-area (~12 390 km^2) sampling, applying anisotropic constraints (Luthcke et al. 2013). This study does not apply additional smoothing or filtering to these data and no further processing is needed. This study uses the CSR and JPL mascon solutions from April 2002 to December 2019 with 32 months of missing data, including an 11-month data gap between July 2017 and May 2018. The GSFC mascon data from 2003 to 2014 are used. The missing values (excluding the 11-month data gap) are conservatively estimated through cubic-spline interpolation. The three mascon solutions (2003–14) are intercompared and validated against water balance–based total water storage change (TWSC) estimates. Considering the longer period, only the CSR and JPL mascon solutions are used for long-term analysis with a comparison with MERRA-2 data.

c. MERRA-2 TWSA data

This study uses the latest atmospheric reanalysis data, MERRA-2, which has been used for atmospheric research and climate monitoring (Gelaro et al. 2017). MERRA-2 uses the same variation analysis algorithm and observation handling as MERRA. Significant improvements have been made for MERRA-2 compared to previous MERRA, MERRA-Land, and Integrated Earth System Analysis (IESA) (Reichle et al. 2017). Through data assimilation using a three-dimensional variational algorithm and simulation using the catchment model (Bosilovich et al. 2015; Reichle et al. 2017), MERRA-2 produces long-term historical time series of atmospheric fields, land surface variables, and water fluxes at a global scale and different time scales (hourly, diurnal, or monthly) (Bosilovich et al. 2015). This study uses the long-term monthly TWSA output from MERRA-2 land surface diagnostics from 1980 to 2019 with a resolution of 0.5° × 0.625°. Before using the data for drought and flood characterization, the reliability of MERRA-2 TWSA data is validated using GRACE TWSA during 2002–19. The accuracy of historical TWS data in the 1980s and 1990s is also assessed by comparing its precipitation forcing data with observed precipitation data. The MERRA-2 data can be accessed through the EarthData search engine.

d. RESS data

According to the annual statistics of constructed reservoirs in the China Agriculture Yearbook (Editorial Board of China Agriculture Yearbook 1980–2017) and China Water Conservancy Yearbook (Editorial Board of China Water

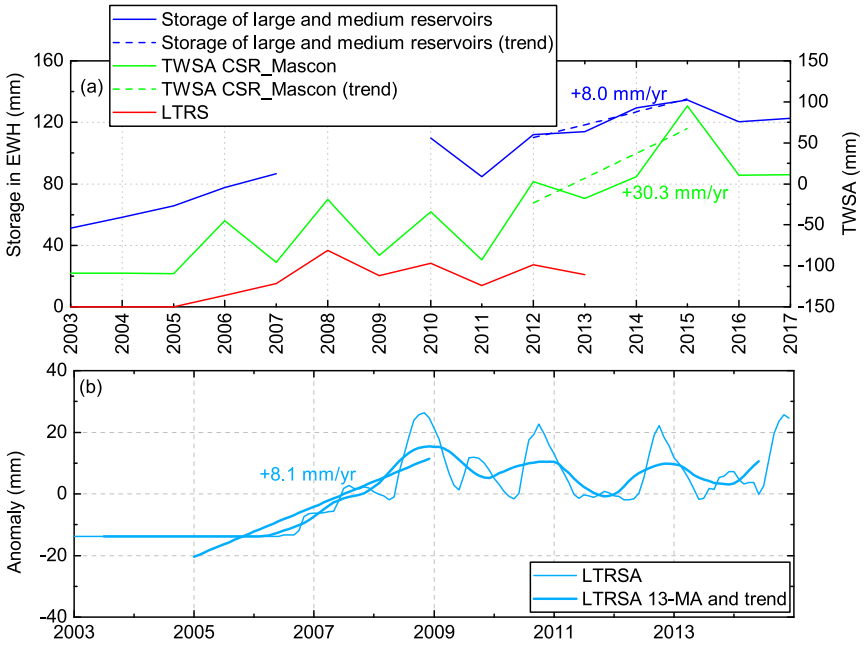

FIG. 2. (a) Actual storage data (available on 1 Jan and 1 Dec during 2003–18, with missing values in 2009) for large and medium reservoirs in the PRB from Water Resources Bulletin of the PRB (Pearl River Water Conservancy Commission 2003–2018). The averaged data on 1 Dec in one year and on 1 Jan the next year are used to be compared with GRACE TWSA and LTRS in the same months. (EWH is equivalent water height.) (b) The monthly and 13-point moving average (13-MA) smoothed anomaly time series of LTRS during 2003–14.

Conservancy Yearbook 2018–2019), there is an increasing number of large reservoirs [storage capacity (SC) $\geq 0.1\,km^3$] in Guangxi province since 2000, with turning points occurring in 2005 and 2011. As shown in Fig. S1 in the online supplemental material, the total SC of large reservoirs in Guangxi was increased by $13.3\,km^3$ from 2005 to 2008 and increased by $27.8\,km^3$ from 2011 to 2012. An increasing trend of SC for medium reservoirs ($0.01 < SC \leq 0.1\,km^3$) can be noted in both Guangxi and Guangdong since 1990 with the total SC increased by $4.5\,km^3$ during 1990–2018. There were over 10 000 small reservoirs ($SC < 0.01\,km^3$) in Guangxi and Guangdong, but the accumulated SC is around $1\,km^3$, and there is only a small increasing trend since 2005 for small reservoirs. An increase of storage by $1\,km^3$ is equivalent to 2.3-mm water height on basin average in the PRB. The newly constructed large and medium reservoirs mainly over the last two decades would result in considerable positive mass anomalies contributing to TWSA and their impacts on drought and flood characterization should be quantified. The contribution of storage anomalies from small reservoirs is minor and can be neglected.

This study further collected the reported actual storage data for large and medium reservoirs in the PRB from the Water Resources Bulletin of the PRB (Pearl River Water Conservancy Commission 2003–2018). The data are only available on 1 January and 1 December each year from 2003 to 2018 with missing values in 2009. Here, the averaged data on 1 December in one year and on 1 January the next year are used to be compared with GRACE TWSA in the same months. As shown in Fig. 2, the RESS represents a remarkable increasing trend since 2003. For the period 2012–15, the average RESS in December and January increased at a rate of $8.0\,mm\,yr^{-1}$, accounting for 26.4% of the trend ($30.3\,mm\,yr^{-1}$) of TWSA from GRACE CSR mascon solution. For the period 2003–08, the increasing trend is mainly induced by the water impoundment of the Longtan Reservoir (LTR) with the storage at a normal water level being approximately $16.2\,km^3$. Here, the observed water-level data of the Longtan Reservoir from 2006 to 2014 (Luo et al. 2016) are used. The water level is converted to storage using the fitted exponential water level–storage relationship estimated by Huang et al. (2019a). The anomaly time series of Longtan Reservoir storage (LTRS) is shown in Fig. 2b, exhibiting an increasing rate of $8.1\,mm\,yr^{-1}$ during 2005–08.

To quantify the impact of RESS change on drought and flood characterization from a long-term perspective, a synthetic long-term RESS time series is constructed using the available LTRS and reported RESS data. The LTRS was assumed to be zero for the period (1980–2006) before dam construction and to stay at the dead water level (330 m with storage of $4.8\,km^3$) during 2015–19. Due to the lack of observed monthly data for all large and medium reservoirs, this study assumes that the trend contribution of RESS to GRACE TWSA based on the data in winter (December and January) is equal to the overall trend contribution of the RESS to GRACE TWSA based on all months. The estimated ratio of 26.4% between storage trends in large and medium reservoirs and GRACE TWSA in winter (December–January) is adopted as

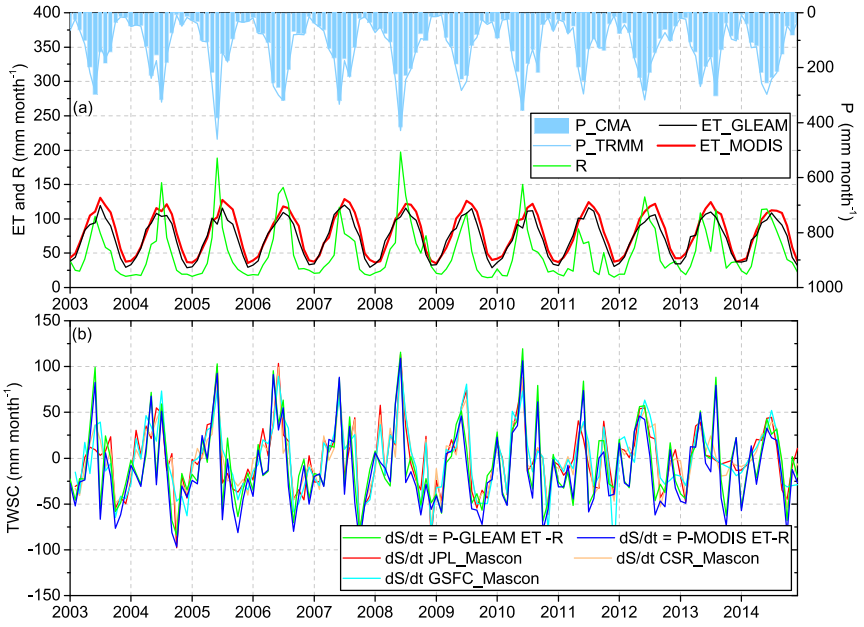

FIG. 3. (a) Monthly time series of precipitation *P* from China Meteorological Administration (CMA) and TRMM, evapotranspiration (ET) from MODIS and GLEAM, and in situ streamflow *R*. (b) Monthly TWSC during 2003–14 estimated from TWSA using three GRACE mascon solutions and terrestrial water balance.

the ratio of the overall linear fitted trend using the entire monthly time series. As the overall fitted trend of GRACE TWSA using the CSR mascon solution during 2012–15 is 17.8 mm yr^{-1} (see Fig. 4a), the RESS trend (26.4%) is assumed to be 4.7 mm yr^{-1}. This rate results in a total of 18.8-mm water storage increment in the PRB by the end of 2015, which is mainly induced by water impoundment at the early stage of newly constructed reservoirs and should also be considered after 2015. Hence, the assumed and observed LTRS during 1980–2019, the assumed trend during 2012–15, and a constant water storage equivalent to 18.8 mm during 2015–19 are added together to construct the synthetic long-term RESS time series, whose anomalies are estimated and removed from the long-term TWSA from GRACE and MERRA-2 before drought and flood characterization.

e. Hydrometeorological data

The monthly precipitation dataset (0.5°) in China from January 1980 to August 2018 is obtained from the China Meteorological Administration (CMA). This gridded precipitation dataset was produced through a thin-plate spline interpolation using the ground-based precipitation observation data from 2472 meteorological stations in China (http://www.nmic.cn/). This study compares the in situ precipitation with the satellite precipitation product (Tropical Rainfall Measuring Mission, TRMM 3B43, 1998–2019) (Huffman et al. 2007) (Fig. 3a or Fig. S2). High consistency with a correlation coefficient of 0.99 and root-mean-square error (RMSE) of 13.6 mm month^{-1} is found between in situ and satellite-based precipitation (Fig. S2). This gives the confidence in the in situ precipitation over the PRB. The CMA precipitation data will be used for results and discussion. The CMA precipitation data during September 2018–December 2019 are unavailable, and TRMM data are used as a surrogate. In addition to the precipitation data from CMA and TRMM, the MERRA-2 precipitation data are also used for comparison. The use of observation-corrected precipitation data is one of many advances relevant to previous MERRA. However, there are still some limitations of this precipitation data, such as the unrealistic representation of the time of maximum precipitation inherited from MERRA (Gelaro et al. 2017). Therefore, the reliability of the MERRA-2 precipitation data should be evaluated to better interpret and assess the historical TWSA dynamics when GRACE TWSA data are unavailable for comparison.

Monthly streamflow measurements at the outlet of the three tributaries of the Pearl River (i.e., the Gaoyao station at the West River, the Boluo station at the East River, and the Shijiao station at the North River) are obtained from the China River Sediment Bulletin (2003–19) for use in estimation of net flow of the study region.

Evapotranspiration (ET) is obtained from the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor and the Global Land Evaporation Amsterdam Model (GLEAM). For the MODIS ET, the MOD16A2 global ET dataset at 0.5° resolution is used. The MOD16 ET datasets are estimated using Mu et al.'s improved ET algorithm (Mu et al. 2011) based on their previous work (Mu et al. 2007). The ET algorithm is based on the Penman–Monteith equation (Monteith 1965). The GLEAM is a set of algorithms developed for the estimation of root-zone soil moisture and terrestrial evaporation from

satellite data. This study uses version 3 of GLEAM ET data (0.25°). Relative to the previous versions, GLEAM V3 includes a novel soil moisture data assimilation system, an improved drainage algorithm, and a modified formulation of the evaporative stress (Martens et al. 2017). The monthly MODIS and GLEAM ET data from 2003 to 2014 are displayed in Fig. 3a. Both ET datasets are used for estimating water balance based TWSC, which is further compared with GRACE TWSC.

f. Self-calibrated Palmer drought severity index

The self-calibrated Palmer drought severity index (scPDSI) is developed based on the original PDSI (Palmer 1965), to make results from different climate regimes more comparable. The scPDSI automatically adjusts the climate pattern and calculates the duration factor using the climate pattern at a given location. In this way, scPDSI considers the precipitation variation at different locations and allows for more consistent and accurate comparisons of the index at various locations than PDSI (van der Schrier et al. 2013; Wells et al. 2004). As the methodology of calculating scPDSI is not significantly different from PDSI, the two indices have the same issue in terms of time lag, frozen precipitation, and frozen soils. For a better analysis of drought using scPDSI, the precipitation data (CRU TS 4.03) used to calculate scPDSI are also collected and validated against CMA precipitation. The scPDSI and CRU precipitation data at a resolution of $0.5° \times 0.5°$ are available from 1901 to 2018 through the Climatic Research Unit (CRU), the University of East Anglia, and are freely accessible from https://crudata.uea.ac.uk/cru/data/drought/.

g. Climate indices

As a climate proxy, El Niño–Southern Oscillation indicates the occurrence of abnormally warm (El Niño) and cold (La Niña) sea surface temperature anomalies in the eastern Pacific (Ndehedehe et al. 2017; Phillips et al. 2012). The magnitude of ENSO is generally expressed as the Niño-3.4 index (Ahi and Jin 2019). The North Atlantic Oscillation is one of the most prominent and recurrent patterns of atmospheric circulation variability (Hurrell et al. 2003). The Indian Ocean dipole is a coupled ocean and atmosphere phenomenon in the equatorial Indian Ocean that mostly affects the climate of countries around the Indian Ocean (e.g., Awange et al. 2014; Ndehedehe et al. 2017; Saji et al. 1999). The Pacific decadal oscillation (PDO) is defined as the leading empirical orthogonal function of North Pacific sea surface temperature anomalies (Schneider and Cornuelle 2005). It is a long-lived ENSO-like pattern of Pacific climate variability (Mantua and Hare 2002).

3. Method

a. GRACE-derived and water balance–based total water storage change

To validate GRACE TWS data, a water balance–based estimate of TWSC is compared with GRACE-based TWSC. Due to the shorter period of MODIS ET data, the validation is only performed for the period 2003–14. The water balance estimation of TWSC can be expressed as the following equation:

$$\frac{dS}{dt} = P + R_{\text{in}} - R_{\text{out}} - \text{ET}, \tag{1}$$

where dS/dt is monthly TWSC (mm month^{-1}); P stands for monthly precipitation (mm month^{-1}), and R_{in} and R_{out} represent the inflow and outflow of the target basin. The inflow is assumed to be negligible for a basin. The outflow is measured at the three streamflow stations located at the outlets of the three tributaries of the Pearl River (i.e., the Gaoyao, Shijiao, and Boluo stations; see Fig. 1). ET is monthly evapotranspiration (mm month^{-1}). In this study, we neglected the impact of human activities on TWSC. GRACE-based TWSC is calculated as the backward difference of TWSA (Long et al. 2014):

$$\frac{dS}{dt} = \frac{\text{TWSA}(t) - \text{TWSA}(t-1)}{t}. \tag{2}$$

b. Water storage deficit index

The water storage deficit is estimated as the residuals by subtracting the climatology from the GRACE-derived TWSA time series (Thomas et al. 2014). Here, the climatology for the period from January 1980 to December 2019 and for the period from April 2002 to December 2019 was calculated for the MERRA-2 and GRACE TWSA time series separately by averaging the TWSA of each month of the GRACE record (e.g., averaging the values of all Januaries in the whole data record). For comparison with GRACE data, WSD and WSDI based on MERRA-2 data during 2002–19 were also calculated. The estimation of WSD can be expressed as follows:

$$\text{WSD}_{i,j} = \text{TWSA}_{i,j} - \overline{\text{TWSA}}_j, \tag{3}$$

where TWSA_{ij} is the GRACE-derived TWSA time series for the jth month in the year i, and $\overline{\text{TWSA}}_j$ is the climatology of TWSA. For a better characterization of drought using WSD and to compare WSD with another drought index, the WSD is normalized into WSDI using zero-mean normalization method as follows:

$$\text{WSDI} = \frac{\text{WSD} - \mu}{\sigma}, \tag{4}$$

where μ is the mean of WSD and σ is the standard deviation of the WSD. The negative sign of WSDI indicates drought conditions, while its magnitude represents the intensity (Sinha et al. 2017). The detailed categorization of drought intensity using WSDI is shown in Table 1.

Thomas et al. (2014) further defined the total severity $S(t)$ to capture the combined effect of deficits and duration of water storage:

$$S(t) = \overline{M}(t)D(t), \tag{5}$$

where $\overline{M}$ is the average deficit since the onset of the deficit period, D is the duration of a drought event, and t denotes the number of drought events; S, $\overline{M}$, and D are determined after the termination of the drought event. A drought event is identified as the negative WSDI persisting for consecutive three or more months (Thomas et al. 2014).

TABLE 1. Category of drought severity based on drought indices (Sun et al. 2018; Wells et al. 2004).

Category	Drought condition	scPDSI	WSDI
D0	No drought	$-1.0 < P$	$0 < W$
D1	Mild drought	$-2.0 < P \leq -1.0$	$-1.0 < W \leq 0$
D2	Moderate drought	$-3.0 < P \leq -2.0$	$-2.0 < W \leq -1.0$
D3	Severe drought	$-4.0 < P \leq -3.0$	$-3.0 < W \leq -2.0$
D4	Extreme drought	$P \leq -4.0$	$W \leq -3.0$

c. Flood potential index

This study follows the method proposed by Reager and Famiglietti (2009) to calculate the flood potential index. First, the storage deficit S_{DEF} is estimated as follows:

$$S_{DEF}(t) = S_{MAX} - S(t-1), \tag{6}$$

where S_{MAX} is the maximum value of GRACE TWSA from 1980 to 2019 (or for the period 2002–19), $S(t-1)$ is the TWS of the previous month, and S_{DEF} represents the highest allowable relative storage change. Second, by combing monthly precipitation P_{MON} and storage deficit S_{DEF}, flood potential amount (FPA) is calculated as follows:

$$\mathrm{FPA}(t) = P_{MON}(t) - S_{DEF}(t), \tag{7}$$

Finally, the FPI is estimated by normalizing the FPA by its historical maximum in each grid cell:

$$\mathrm{FPI}(t) = \frac{\mathrm{FPA}(t)}{\max\left[\mathrm{FPA}(t)\right]}. \tag{8}$$

The FPI value should be less than or equal to 1. The flood will be more likely to occur if FPI is closer to 1.

4. Results and discussion

a. Validation of GRACE TWSC

GRACE TWSA from CSR, JPL, and GSFC mascon solutions are intercompared with each other during 2003–14, showing good agreement [$r > 0.95$, root-mean-square difference (RMSD) = 15–18 mm] (see Fig. S3). A further comparison of TWSA from CSR, JPL mascons, and MERRA-2 at a longer time scale is shown in Fig. 4a (see section 4b). The monthly TWSC from GRACE data is compared with the water balance estimation from ground-based precipitation, streamflow, and MODIS/GLEAM-based ET (Fig. 3b). The TWSC estimated using CSR, JPL, and GSFC mascons indicates an acceptable agreement ($r = 0.69$, 0.70, and 0.68, and RMSD = 33.8, 34.0, and 34.9 mm month^{-1}, respectively) comparing to the water balance estimations using MODIS ET. The comparison of TWSC between CSR, JPL, and GSFC mascons and water balance estimations using GLEAM ET also indicates good consistency in both timing ($r = 0.68$, 0.68, and 0.66, respectively) and magnitude with slightly lower RMSD (31.8, 32.3, and 33.2 mm month^{-1}, respectively) due to better ET estimates. As seen, the TWSC from CSR and JPL mascon shows a slightly higher correlation and lower RMSD than GSFC mascon compared to water balance estimations. The overall agreement between GRACE-based and water balance–based TWSC verifies the reliability of GRACE TWS data to be used for drought and flood characterization. Considering the better performance, only the CSR and JPL mascons are used for further study.

b. Long-term variability of TWSA from GRACE and MERRA-2

Figure 4a compares the TWSA from GRACE and MERRA-2 during 2002–19, showing generally good agreement on both timing and amplitude. The difference between GRACE and MERRA-2 TWSA is a bit larger than that between the two GRACE mascon solutions. The major discrepancy lies in the overestimation of TWSA by MERRA-2 in the wet season (June–September; see the climatology in Fig. 7e). The three TWSA time series show a close linear-fitted trend based on the 13-point moving average (13-MA) during 2002–16 with a difference of less than 2 mm yr^{-1} (Figs. 4a,b). Some larger differences are found among the fitted trend during two short periods. During 2005–08, the trend of TWSA from the JPL mascon solution is ~7 mm yr^{-1} larger than that of CSR mascon and MERRA-2. During 2012–15, MERRA-2 exhibits a larger trend of 6–7 mm yr^{-1} compared to JPL and CSR mascon solutions (Figs. 4a,b). It should be noted that these two periods correspond to the time of building new large dams in the PRB (see section 2d). The increasing trends shown in Figs. 4a and 4b should be partly attributed to reservoir impoundment (e.g., LTRS accounts for 34%–50% of increment of TWSA during 2005–08). The overall fitted trend of RESS during 2002–16 according to observation-based synthetic data is 2.8 mm yr^{-1}. Meanwhile, the increasing trend of TWSA during 2002–16 (also for periods 2005–08 and 2012–15) is partly influenced by the increased precipitation with a rate of 16.7 mm yr^{-1} (Fig. 4c).

It is worth mentioning that the MERRA-2 reanalysis system does not simulate reservoir and groundwater storage change, which all can be measured by GRACE. Nevertheless, the general consistency between MERRA-2 and GRACE TWSA indicates the effectiveness of the total land water budget used by MERRA-2 for modeling TWSC in the PRB (Bosilovich et al. 2016). However, the consistency between MERRA-2 and GRACE TWSA cannot guarantee the reliability of MERRA-2 TWSA during the period before the launch of the GRACE mission. Comparison of precipitation data from MERRA-2 with those from CMA and TRMM at annual scale shows that MERRA-2 significantly underestimates the precipitation for the period 1982–97 (before the TRMM mission). This is mainly owing to the lack of sufficient satellite or observation data to force the reanalysis system (see Fig. 2-1 in Bosilovich et al. 2015). As precipitation

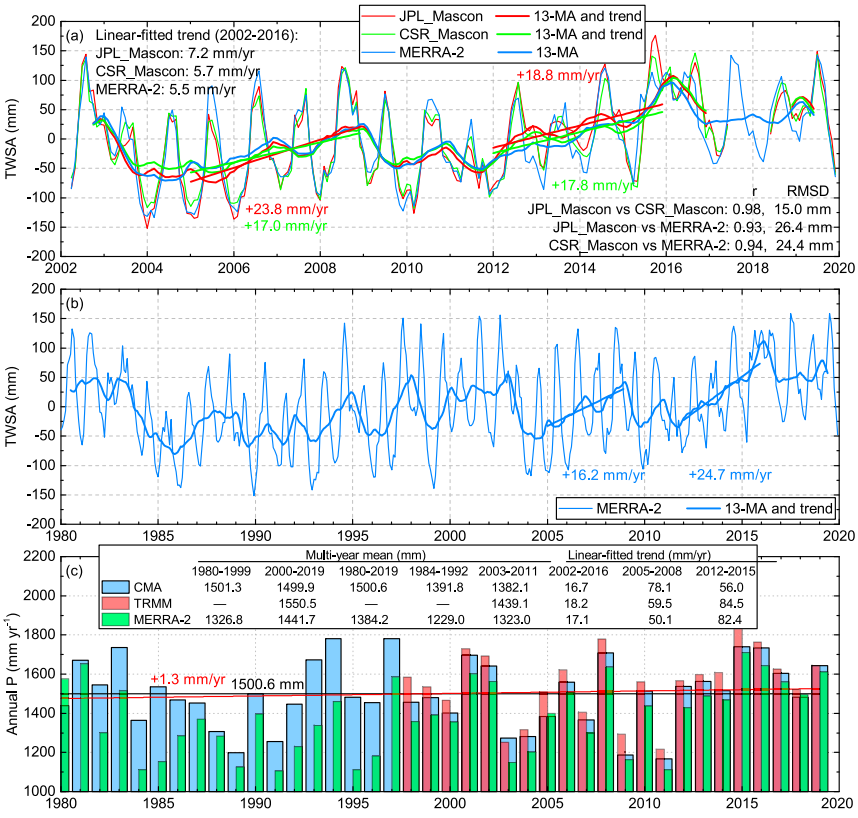

FIG. 4. (a) Monthly and 13-MA smoothed TWSA from two GRACE mascon solutions and MERRA-2. The linear-fitted trends of TWSA from the two GRACE mascon solutions for the periods 2005–08 and 2012–15 are plotted and noted. The linear-fitted trends of TWSA from the three datasets for the period 2002–16 are shown with notation. The comparison [correlation and root-mean-square difference (RMSD)] between the three TWSA time series are shown with notation. (b) Long-term monthly and 13-MA smoothed TWSA from MERRA-2 since 1980. The linear-fitted trends of TWSA for the periods 2005–08 and 2012–15 are plotted and noted. The TWSA shown in (a) and (b) is the original data without removing RESS. (c) Annual precipitation from CMA, TRMM, and MERRA-2. The inserted legend with a table shows the multiyear mean annual precipitation for five different periods and the linear-fitted trends for three periods. The black and red lines, respectively, show the long-term mean and linear-fitted trend of precipitation over the past four decades.

plays a critical role in regulating TWS dynamics, the underestimated precipitation in the 1980s and 1990s in the MERRA-2 system would result in incorrect TWSA simulation. In particular for the years 1985–87, the MERRA-2 precipitation exhibits an opposite trend compared to the in situ–based precipitation from CMA. Although the MERRA-2 TWSA before 1998 may be unreliable, it provides a test bed for examining the impact of RESS change on drought and flood characterization from a long-term perspective. Besides, further comparison between MERRA-2-based WSDI and scPDSI verifies the overall reliability of TWSA in timing during 1987–2001 [see section 4c(2)].

c. *Characterization of drought*

1) DROUGHT BEFORE AND AFTER 2000

Figure 5 shows the time series of WSD, drought severity, precipitation anomaly, and cumulative WSD and precipitation anomaly. Here, the results with and without RESS contribution are compared. A statistical summary of drought characteristics is shown in Table 2 and Tables S1–S3. The results without RESS contribution will be discussed first. As shown in Table 2, there are considerable differences in drought characteristics using different datasets. During 2003–19, CSR mascon, JPL mascon, and MERRA-2 data identify 14, 11, and 9 drought events with a total duration (total severity) of 110 (−3525), 98 (−4014), and 97 (3342) months (mm), respectively. There have been two major drought periods over the last two decades: 2003–06 and 2009–12. For the period 2003–06, all three datasets identify a similar onset in June/July 2003 and the end in May 2006 with a total duration of 33–35 months (excluding the one-month intervals). However, there is a large difference in the total severity for this period with the largest severity of 1691.3 mm from JPL mascon, followed by MERRA-2 (1386.9 mm) and CSR mascon (1118.9 mm). CSR mascon might underestimate the drought severity during

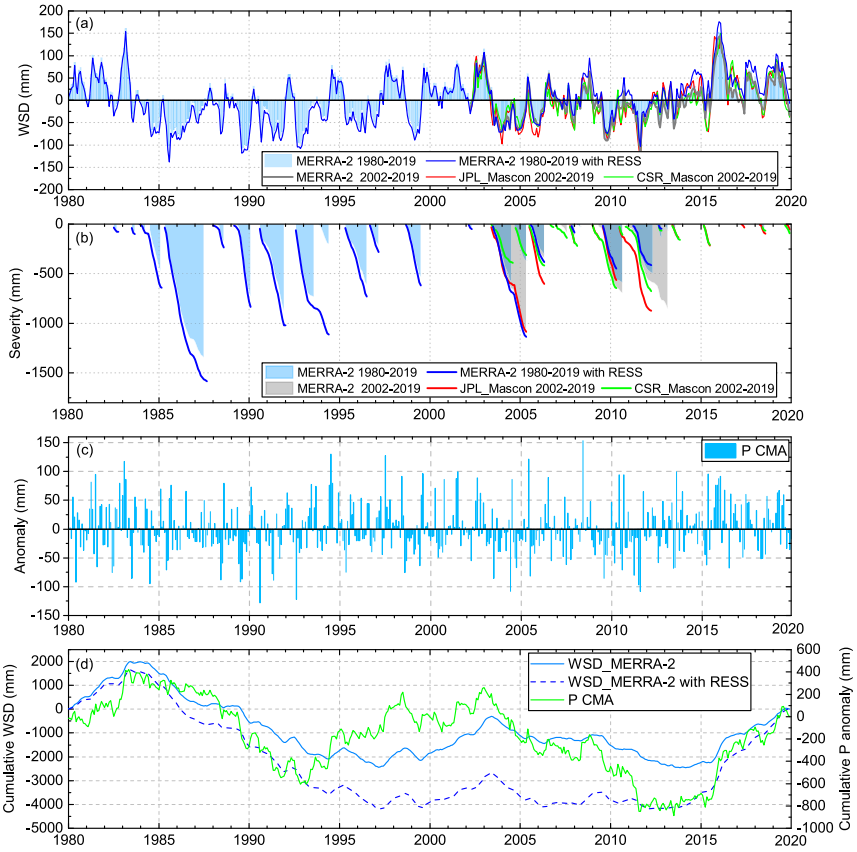

FIG. 5. (a) Monthly time series of water storage deficit (WSD) estimated using TWSA from MERRA-2 (1980–2019) and two GRACE mascon solutions (2002–19). For comparison with GRACE results, the WSD estimated using the MERRA-2 TWSA only during 2002–19 is also plotted. (b) Water storage severity for each drought event estimated using the data shown in (a). (c) Monthly time series of precipitation P anomalies using the data from CMA. (d) Cumulative WSD using MERRA-2 data and cumulative precipitation anomaly using CMA data. It is noted that in (a), (b), and (d), except for the notation "with RESS", all the other legends are the results without RESS. RESS means reservoir storage based on synthetic data (see section 2d).

2003–06. As shown in Figs. S3 and S4a, the TWSA values from GSFC and JPL mascons and MERRA-2 all represent lower negative amplitude than CSR mascon during 2003–06, indicating the possible underestimation of drought severity by CSR mascon. For the period 2009–12, JPL and CSR mascon identify the same drought duration (38 months, excluding the one-month intervals), which starts from January 2009 and ends in April 2012, with less divergence (166 mm) in total severity than the period 2003–06. Comparatively, MERRA-2 may fail to well capture this drought event with the onset occurring later by 4 months (in May 2009) and lasting a longer duration covering the near-normal water year in 2012 (see Fig. 4c) and ending in March 2013 (Table 2). The different results estimated from the GRACE mascons are caused by the sensitivity of time-variable spatial mass distribution on the mascon solutions considering their differences in spatial constraints with different spatial resolution and mass inversion strategies (Save et al. 2016; Watkins et al. 2015). The distinctions between GRACE and MERRA-2 are linked to their intrinsic physically different theory and methodology with the former relying on hydrogeodesy and the latter based on land water budget and data assimilation.

Over the past four decades, the cumulative rainfall anomaly (Fig. 5d) exhibits a decreasing trend during two periods, 1984–92 and 2003–11, indicating the occurrence of meteorological drought. The period 2003–11 experienced a more serious meteorological drought as revealed by the lower turning point of cumulative rainfall anomaly in August 2011 (−848.6 mm) than that in March 1993 (−606.0 mm). The cumulative WSD based on MERRA-2 without RESS contribution also shows decreasing trends during those two drought periods. However, unlike the period 2003–11 when the cumulative WSD and precipitation anomaly behave in a similar trend, the decreasing trend that began in 1984 lasted for a longer duration than the cumulative precipitation anomaly with the turning point occurring later by five years in 1997. This result is problematic. MERRA-2 significantly overestimates the WSD before 2000 since it used poor precipitation data with large underestimation (Fig. 4c). In particular, for the period May 1985–July 1987 the

TABLE 2. A statistical summary of characteristics of hydrological drought events identified using GRACE and MERRA-2 data without RESS contribution during 2002–19.

No.	Period	Duration (months)	Avg deficit (mm month^{-1})	Peak deficit (mm)	Peak deficit date	Total severity (mm)	Peak WSDI (category)	Peak scPDSI (category)	Combined category
				MERRA-2 2002–19					
1	Jul 2003–May 2005	23	−45.6	−72.0	Jan 2004	−1049.7	−1.6 (D2)	−3.5 (D3)	D2/D3
2	Aug 2005–May 2006	10	−33.7	−57.7	Jan 2006	−337.2	−1.3 (D2)	−3.3 (D3)	D2/D3
3	Apr 2007–Aug 2007	5	−16.9	−26.5	Aug 2007	−84.3	−0.6 (D1)	−2 (D2)	D1/D2
4	Nov 2007–Jan 2008	3	−41.0	−50.9	Jan 2008	−123.1	−1.1 (D2)	−2.4 (D2)	D2/D2
5	Mar 2008–May 2008	3	−8.9	−18.8	May 2008	−26.6	−0.4 (D1)	−1.2 (D1)	D1
6	May 2009–Sep 2010	17	−40.8	−90.0	Oct 2009	−693.3	−2 (D3)	−3.3 (D3)	D3
7	Nov 2010–Mar 2013	29	−29.3	−115.4	Sep 2011	−850.2	−2.6 (D3)	−4 (D4)	D3/D4
8	Jun 2013–Aug 2013	3	−42.1	−63.8	Jul 2013	−126.2	−1.4 (D2)	−1.1 (D1)	D2/D1
9	Apr 2018–Jul 2018	4	−12.9	−21.2	Jul 2018	−51.5	−0.5 (D1)	1.4 (D1)	D1
1–9	2003–18	97	−34.5			−3342.2			
1–2	2003–06	33	−42.0			−1386.9			
6–7	2009–13	46	−33.6			−1543.5			
				JPL mascon 2002–19					
1	Jun 2003–May 2005	24	−45.3	−98.1	Jan 2004	−1087.3	−2 (D3)	−3.5 (D3)	D3
2	Jul 2005–May 2006	11	−54.9	−82.8	Jan 2006	−604.1	−1.6 (D2)	−3.3 (D3)	D2/D3
3	Oct 2007–Jan 2008	4	−47.3	−65.0	Jan 2008	−189.4	−1.3 (D2)	−2.4 (D2)	D2
4	Jan 2009–Jun 2009	6	−25.9	−43.0	Feb 2009	−155.3	−0.9 (D1)	−1.4 (D1)	D1
5	Aug 2009–May 2010	10	−56.5	−88.0	Nov 2009	−565.2	−1.8 (D2)	−3.3 (D3)	D2/D3
6	Jul 2010–Apr 2012	22	−39.7	−130.7	Sep 2011	−874.0	−2.6 (D3)	−4 (D4)	D3/D4
7	Jun 2013–Oct 2013	5	−27.2	−41.5	Aug 2013	−136.1	−0.8 (D1)	−1.1 (D1)	D1
8	Mar 2015–Jul 2015	5	−43.4	−70.2	Jun 2015	−217.0	−1.4 (D2)	0.3 (D0)	D2/D0
9	Apr 2017–Jun 2017	3	−11.5	−23.4	Jun 2017	−34.4	−0.5 (D1)	2.2 (D2)	D1/D2
10	Apr 2018–Aug 2018	5	−19.4	−39.4	Jun 2018	−97.1	−0.8 (D1)	1.4 (D1)	D1
11	Oct 2019–Dec 2019	3	−18.0	−36.7	Dec 2019	−53.9	−0.7 (D1)	—	—
1–11	2003–18	98	−41.0			−4013.8			
1–2	2003–06	35	−48.3			−1691.3			
4–6	2009–12	38	−42.0			−1594.5			
				CSR mascon 2002–19					
1	Jun 2003–Aug 2004	15	−26.0	−66.3	Jan 2004	−390.2	−1.5 (D2)	−2.7 (D2)	D2
2	Oct 2004–May 2005	8	−38.9	−64.5	Jan 2005	−311.3	−1.5 (D2)	−3.5 (D3)	D2/D3
3	Jul 2005–May 2006	11	−38.0	−59.4	Jan 2006	−417.5	−1.3 (D2)	−3.3 (D3)	D2/D3
4	Sep 2006–Nov 2006	3	−10.2	−13.5	Oct 2006	−30.5	−0.3 (D1)	−2.4 (D2)	D1/D2
5	Jan 2007–Aug 2007	8	−17.3	−35.6	Feb 2007	−138.4	−0.8 (D1)	−2 (D2)	D1/D2
6	Oct 2007–Mar 2008	6	−36.7	−67.3	Jan 2008	−220.1	−1.5 (D2)	−2.4 (D2)	D2
7	Jan 2009–May 2010	17	−38.0	−77.7	Nov 2009	−646.7	−1.8 (D2)	−3.3 (D3)	D2/D3
8	Jul 2010–Sep 2010	3	−35.5	−50.3	Aug 2010	−106.4	−1.1 (D2)	−2.9 (D2)	D2
9	Nov 2010–Apr 2012	18	−37.5	−115.1	Sep 2011	−675.5	−2.6 (D3)	−4 (D4)	D3/D4
10	Sep 2012–Nov 2012	3	−21.7	−32.4	Oct 2012	−65.2	−0.7 (D1)	−1.3 (D1)	D1
11	Jun 2013–Nov 2013	6	−26.3	−40.2	Sep 2013	−157.5	−0.9 (D1)	−1.1 (D1)	D1
12	Mar 2015–Jul 2015	5	−42.0	−68.8	May 2015	−209.9	−1.6 (D2)	0.3 (D0)	D2/D0
13	Jun 2018–Aug 2018	3	−21.7	−32.4	Jul 2018	−65.1	−0.7 (D1)	1.5 (D1)	D1
14	Sep 2019–Dec 2019	4	−22.7	−37.7	Dec 2019	−91.0	−0.9 (D1)	—	D1/—
1–14	2003–19	110	−32.0			−3525.1			
1–3	2003–06	34	−32.9			−1118.9			
7–9	2009–12	38	−37.6			−1428.6			

drought duration and severity quantified using MERRA-2 reach 27 months and 1338.7 mm respectively (Fig. 5b and Table S1), which exceed the values for 2003–06 and 2009–12 despite less rainfall deficit (Fig. 4c). Besides, using the problematic long-term MERRA-2 data from 1980 to 2019 leads to underestimated duration and severity for the drought events after 2000 (Fig. 5b, Table 2, and Table S1). Although this study lacks reliable historical TWS data for quantifying the hydrological drought before 2000, it can be inferred that the total severity of the hydrological drought for that period should be less than that of the period after 2000 considering the propagation from less severe meteorological drought.

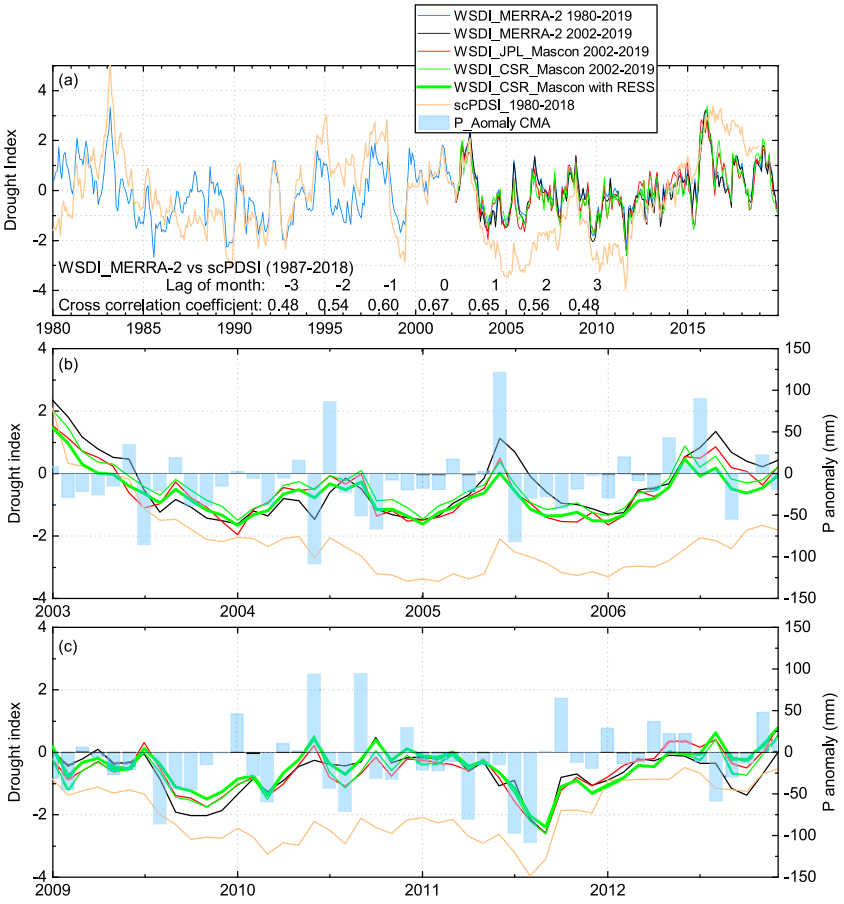

FIG. 6. (a) Comparison of scPDSI and WSDI estimated using TWSA from MERRA-2 (1980–2019) and two GRACE mascon solutions (2002–19). For comparison with GRACE results, the WSDI estimated using the MERRA-2 TWSA only during 2002–19 is also plotted. (b) As in (a), but only for the period 2003–06. (c) As in (a), but only for the period 2009–12. Precipitation anomalies are also shown in (b) and (c) for comparison.

This will be further demonstrated by the scPDSI in the next section.

2) WSDI VERSUS SCPDSI

Figure 6 compares the time series of WSDI and scPDSI over the PRB from 1980 to 2018. Agreement on timing can be noted between WSDI and scPDSI, except for the period before 1987. During 2002–19, the WSDI derived using GRACE mascons represents a slightly higher correlation ($r = 0.73$) with scPDSI than MERRA-2 ($r = 0.69$). For the period 1980–2001 (or 1987–2001), the WSDI based on MERRA-2 also has a good correlation with scPDSI with a correlation coefficient of 0.65 (or 0.72). As the scPDSI originated from PDSI, which is designed to measure meteorological drought (Palmer 1965), a time lag would be expected between WSDI and scPDSI considering the propagation from meteorological drought to hydrological drought. However, there is no apparent time lag between WSDI and scPDSI as indicated by the maximum cross-correlation coefficients at the lag of month zero in Fig. 6a. This is because scPDSI involves the balance between soil moisture supply and demand. Hence, scPDSI can be also regarded as a hydrological drought index (Heddinghaus and Sabol 1991; Karl 1986). A further comparison between the drought indices and precipitation anomalies during the two drought periods 2003–06 and 2009–12 is conducted to reveal the lag effect between meteorological and hydrological drought. As shown in Figs. 6b and 6c, the troughs of WSDI always occur later (at least one month) than the largest rainfall deficit. This is more apparent when the rainfall deficit occurs in summer (e.g., July in 2003 and 2005 and August in 2009 and 2011). Similar lag responses also happen to scPDSI except for the summer drought in 2011 when the peak rainfall deficit and minimum scPDSI value co-occur in August. The results indicate that scPDSI does have the characteristics of hydrological drought index as WSDI, but it may not be suitable for all drought conditions like the severe drought in 2011.

Despite the overall good agreement on timing, WSDI and scPDSI have large differences in characterizing drought duration as well as the drought intensity, in particular during 2004–12. For the droughts since 2003, both WSDI and scPDSI

TABLE 3. A summary of the drought duration when scPDSI is below −1 and −3 and WSDI is below −2, and the peak value of drought index during each event.

No.	Period	Duration (months)	Peak index	Peak date
		scPDSI < −1		
1	Jan 1980–Jun 1980	6	−1.6	Jan 1980
2	Jan 1986–Mar 1986	3	−1.4	Mar 1986
3	Nov 1986–Oct 1987	12	−2.0	Apr 1987
4	Dec 1987–Jul 1988	8	−2.0	Jul 1988
5	Jun 1989–Jan 1990	8	−3.1	Nov 1989
6	Aug 1990–Dec 1990	5	−1.5	Aug 1990
7	Apr 1991–Dec 1991	9	−2.3	May 1991
8	Aug 1992–Jan 1993	6	−2.3	Nov 1992
9	Sep 1998–Jun 1999	10	−2.9	Jun 1999
10	Jul 2003–Feb 2008	56	−3.5	Feb 2005
11	Feb 2009–Dec 2011	35	−4.0	Aug 2011
12	Aug 2012–Oct 2012	3	−1.3	Oct 2012
		scPDSI < −3		
1	Oct 2004–May 2005	8	−3.5	Feb 2005
2	Oct 2005–Jan 2006	4	−3.3	Jan 2006
3	Jul 2011–Sep 2011	3	−4.0	Aug 2011
		WSDI < −2		
1	Oct 2009–Nov 2009	2	−2.0	Oct 2009
2	Aug 2011–Sep 2011	2	−2.6	Sep 2011

indicate the similar onset of drought in June/July 2003 (or January/February 2009). However, the drought events reflected by scPDSI are more continuous with a longer duration than those reflected by WSDI, which are divided into several shorter droughts (Table 2 and Figs. 6a–c). In particular, for the mirroring on severe drought (scPDSI ≤ −3, WSDI ≤ −2), scPDSI reveals three dry spells that reach severe drought with a duration of 3–8 months (i.e., the droughts during October 2004–May 2005, October 2005–January 2006, and the summer drought in 2011), whereas only two short severe droughts with a 2-month duration are detected by WSDI in 2009 and 2011 (Table 3). Over 40% of the drought events identified by WSDI are characterized to be less serious by one category compared to scPDSI (Table 2). For instance, all the three datasets (MERRA-2 and two GRACE mascon solutions) capture the same lowest WSDI in September 2011 when the drought is classified as severe level (D3), one level lower than the extreme drought (D4) identified by scPDSI (Table 2). Such differences in drought characteristics quantified between WSDI and scPDSI are reasonable. First, unlike WSDI, which encompasses all the water storage components, the storage contribution in lakes, reservoirs, rivers, and aquifers is not included in the Palmer drought model. Besides, the scPDSI is calibrated based on a longer period of data starting from 1901, whereas the WSDI is estimated using a shorter period of data. The parameters (e.g., potential evapotranspiration) used by scPDSI are sensitive to the calibration period. This will cause changes in the scPDSI values from a longer perspective.

Although scPDSI is different in measuring hydrological drought comparing to WSDI, its longer period provides an insight into the long-term variability of drought conditions over the last four decades. A comparison between the CRU precipitation data used by scPDSI and the data from CMA shows good agreement ($r = 0.99$, RMSE = 15.2 mm month^{-1}; see Fig. S2), further stressing confidence in using scPDSI as an alternative to interpreting historical drought features when no reliable TWSA data for calculating WSDI are available. For the period before 2000, compared to scPDSI, MERRA-2 overestimates the drought intensity during 1983–86 and underestimates the wet condition during 1993–97. The hydrological drought condition is less serious before 2000 than after. The total drought duration identified by scPDSI during 1980–99 is 67 months including only one month with severe drought (November 1989) compared to the 94 months of drought duration with 15 months being severe drought during 2000–19 (Table 3).

3) CONTRIBUTION OF RESS TO DROUGHT CHARACTERIZATION

The above two subsections only illustrate the results after removing the contribution of storage change in the new reservoirs over the last two decades (the LTR constructed in 2008 and some large and medium reservoirs during 2012–15). Here, the original results considering RESS contribution are also presented. Figure 5d compares the cumulative WSD before and after removing the reservoir contribution. When considering the RESS contribution, the cumulative WSD stays at a stable level during 2006–12, reflecting no serious droughts. This is distinguishable from the decreasing trend of cumulative WSD during 2009–12 when reservoir contribution is removed. Tables S2 and S3 and Table 2 respectively summarize the duration and severity of each drought event identified before and after removing the RESS contribution. The inclusion of RESS contribution results in expanded (shrunk) drought duration and overestimated (underestimated) severity for the period 2003–06 (2009–12). Taking CSR mascon as an example, the duration (severity) considering the RESS contribution is 2 months longer (586.4 mm larger) during 2003–06, but 7 months shorter

(135.2 mm lower) during 2009–12 than those removing the reservoir contribution. For the period 1980–99, the drought duration (severity) quantified using MERRA-2 data is also overestimated by 17 months (1211.4 mm) when considering reservoir contribution.

The impact of RESS change on drought characterization can be divided into two stages. The first is the early water-filling stage as the new dam is built, and the second is the seasonal regulation stage for hydropower generation or flood control purposes, etc. The water-filling stage mainly results in an increasing trend of TWSA, while the seasonal regulation alters the amplitude and timing of TWSA. Both stages influence the quantification of drought duration and severity. Notably, the water-filling stage only increases the constant water storage in a reservoir, which usually remains insignificant change even during drought after the reservoir reaches dead water level (e.g., the lowest level of LTRS during summer 2011 in Fig. 2b). The increasing trend induced by the water-filling stage cannot reflect hydrological drought and should be removed when focusing on a long-term hydrological drought analysis, especially surface water drought.

d. Characterization of flood

1) Flood potential before and after 2000

The FPI calculated using the TWSA and precipitation directly indicates the flood potential. Figures 7a and 7b show the long-term monthly, smoothed time series of FPI over the past four and two decades, respectively. Here, the impact of RESS change on FPI estimation has been removed. Figures 7c and 7d show the monthly and smoothed time series of rainfall and streamflow. FPI (Figs. 7a,b) exhibits the same interannual variability and trend as GRACE TWSA (Figs. 4a,b) as well as rainfall and streamflow (Figs. 7c,d). There are increasing trends for FPI, rainfall, and streamflow from 2005 to 2008 and from 2012 to 2015 and decreasing trends in 2003, 2009, 2011, and 2016–17. Figures 7e–g compare the climatologies of TWSA, FPI, *P*, and streamflow before and after the GRACE mission (2002). The FPI is featured by an apparent seasonal cycle with the high flood potential occurs in summer (June–August) (Fig. 7f). Notably, as a result of considering precipitation the climatology of FPI during 2002–19 peaks one month earlier in June (the same as the peak of precipitation) than the TWSA (Figs. 7e–g). This indicates the effectiveness of GRACE- and precipitation-based FPI on capturing the hydrometeorological extremes.

For better analysis of the capability of FPI to capturing flood events using different datasets, this study defines a high flood potential when the FPI is equal to or above 0.7. A summary of the flood events with corresponding FPI, precipitation, and streamflow is shown in Table 4. During 2002–19, both JPL and CSR mascons identify four flood events with the highest flood potential that occurred in August 2002 (FPI = 1.0), followed by June 2008 (FPI > 0.80), July 2006 (FPI ≥ 0.79), and July 2019 (FPI = 0.76). Notably, the flood potential is relatively higher in August 2002 although the precipitation (293.3 mm) is 29.5% less than that in June 2008 (415.8 mm). This difference is caused by the antecedent precipitation that controls the degree of soil saturation and hence influences the runoff generation process. In 2002, May and June are normal months with almost identical rainfall compared to climatology, which determined the initial status of soil saturation. The following July is a wet month. As indicated in Table 4, both July and August in 2002 received more rainfall (over 21% and 43%) than the climatology. With continuous rainfall infiltration in the antecedent three months, the soil is saturated, and substantial overland flow generated when peak rainfall came in August. As seen, the streamflow in August 2002 was 100.2% more than the climatology (Table 4). The flood potential in June 2008 was relatively lower than that in August 2002 considering the relatively drier antecedent spring with the total rainfall below the long-term mean (by 4%). As the soil layers are drier, they need more water to become saturated and to lead to runoff generation.

The FPI based on MERRA-2 should be overestimated in summer owing to the higher TWSA (Figs. 7e,f). This can be further confirmed by considering the rainfall intensity, particularly for June 2016 and August 2019 when no abnormal high rainfall occurred (Table 4). It is still difficult to verify the reliability of the FPI based on MERRA-2 before 2002 when no GRACE data or other reliable flood index is available for validation. Nevertheless, the high flood potential identified by MERRA-2 in July 1994 and July 1997 may be realistic considering the high rainfall intensity (over 50% more than the climatology). It should be noted that the FPI is sensitive to the study period, which influences the determination of the maximum flood potential amount for normalization. Table 4 shows that five more flood events are captured by MERRA-2 when focusing on a shorter period during 2002–19 than the long-term period 1980–2019. Besides, during the gap between GRACE and the Follow-On mission, MERRA-2 identifies a flood event in August 2017 when the FPI estimated using the MERRA-2 data on a long-term basis during 1980–2019 is 0.78, much lower than that estimated (0.94) based on the period 2002–19. The former 0.78 is more acceptable considering the rainfall and streamflow amount, which exceed the climatology by only 22.9% and 31.1%, respectively.

2) Spatial variability of the flood in 2008

This section will focus on the spatial variability of the most recent severe flood event in 2008. The monthly spatial map of FPI from March to November 2008 is derived using gridded CSR mascon solution (0.25°), which is compared with gridded precipitation data from TRMM at the same resolution (Fig. 8). In March and April, the entire PRB was at a relatively dry status with low FPI (<0), especially for the upper stream of the western PRB (i.e., the Nan-Bei Pan RB). In May, the western and northern parts of Nan-Bei Pan RB were still dry, but the other regions of the PRB became wet, especially for the northern PRB (the Hongliu RB) with the highest FPI of 0.82. In June, the high flood potential region moved to the eastern parts of the PRB, including the West RB, North RB, East RB, and Pearl River Delta with the FPI above 0.9 in most of these regions. In July, the western PRB became wet, and the highest flood potential (FPI > 0.8) located in the Hongliu RB. In August, the whole PRB was in wet condition, in particular for the Hongliu RB and Yu RB where most regions had an FPI over 0.9. The severity of the flood started to alleviate in most parts of the PRB in September except for the Yu RB, most of

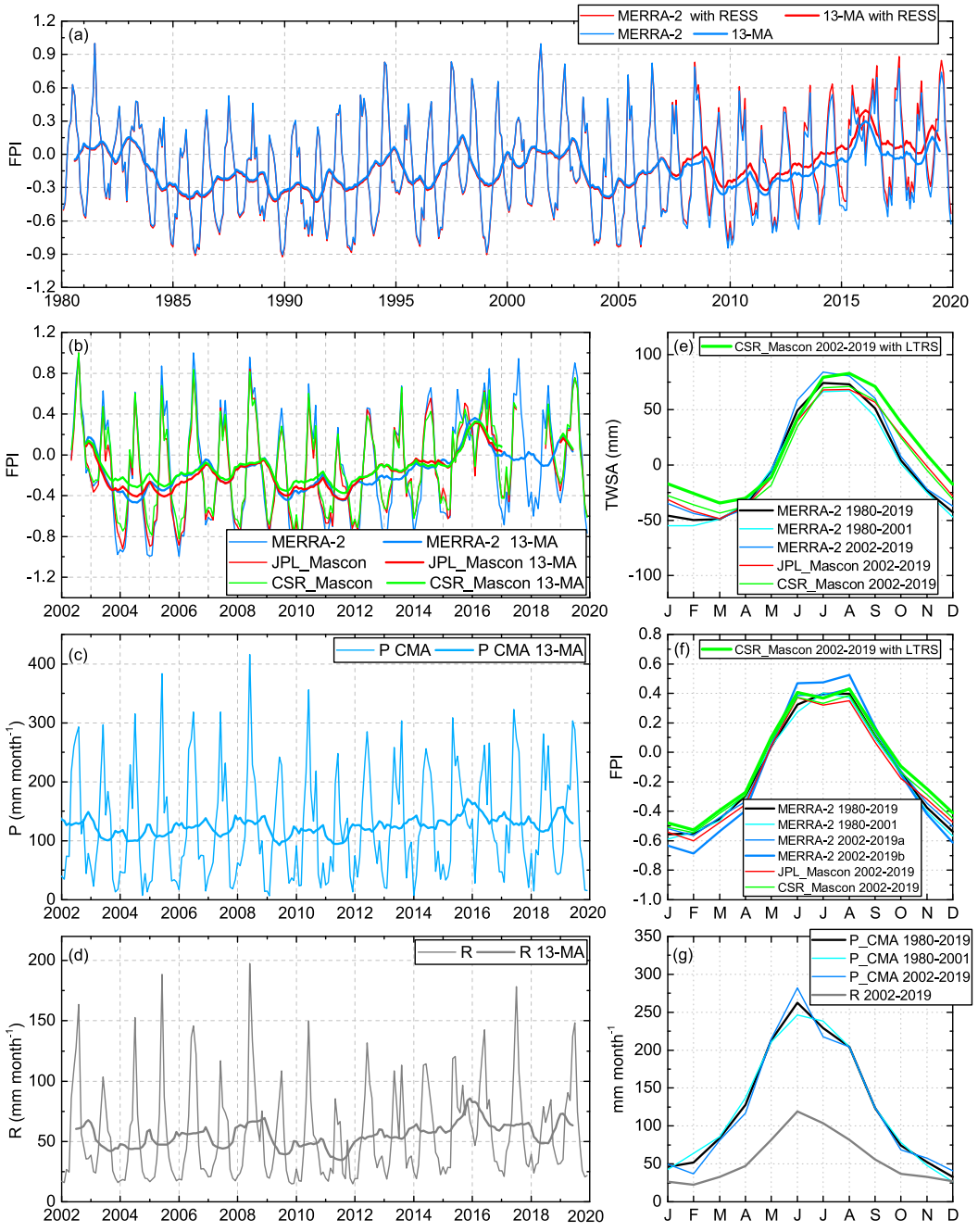

FIG. 7. (a) Monthly and 13-MA smoothed time series of flood potential index (FPI) estimated using TWSA from MERRA-2 (1980–2019). (b) Monthly and 13-MA smoothed time series of FPI estimated using TWSA from MERRA-2 and two GRACE mascon solutions during 2002–19. (c) Monthly and 13-MA smoothed time series of precipitation *P*. (d) Monthly and 13-MA smoothed time series of streamflow *R*. (e) Climatology (i.e., multiyear mean annual cycle) of TWSA for the periods before and after 2002. (f) Climatology of FPI for the periods before and after 2002 using the data in (a) and (b). (g) Climatology of precipitation and streamflow for the periods before and after 2002 using the data in (c) and (d). It is noted that in (a), (b), (e), and (f), except for the notation "with RESS", all the other legends are the results without RESS. RESS means reservoir storage based on synthetic data (see section 2d).

TABLE 4. A summary of flood events identified with FPI ≥ 0.7. The results estimated with and without RESS are compared. Also shown is the precipitation (P_CMA) and streamflow (R) of the corresponding flooding month. PAC means the percentage of precipitation or streamflow above the climatology. The italic and bold numbers are FPI equal to or greater than 0.7.

		Without RESS				With RESS								
		FPI 1980–2019	FPI 2002–19			FPI 1980–2019	FPI 2002–19			P_CMA	PAC for P		PAC for	
No.	Period	MERRA-2	MERRA-2	JPL mascon	CSR mascon	MERRA-2	MERRA-2	JPL mascon	CSR mascon	(mm)	1980–2019	R (mm)	R 2002–19	CMA
1	Jul 1981	***1.00***				***1.00***				323.8	41.5%			
2	Jul 1994	***0.83***				***0.83***				358.9	56.8%			
3	Aug 1994	***0.80***				***0.80***				284.0	39.0%			
4	Jul 1997	***0.83***				***0.83***				356.5	55.7%			
5	Aug 1997	***0.75***				***0.75***				242.8	18.8%			
6	Jun 2001	***0.75***				***0.75***				349.6	33.2%			
7	Jul 2001	***1.00***				***1.00***				328.2	43.4%			
8	Aug 2001	***0.71***				***0.71***				207.3	1.5%			
9	Jul 2002	0.61	***0.74***	0.62	0.62	0.60	0.68	0.57	0.61	277.9	21.4%	127.7	23.4%	
10	Aug 2002	***0.81***	***0.99***	***1.00***	***1.00***	***0.81***	***0.92***	***1.00***	***1.00***	293.3	43.6%	163.4	100.2%	
11	Jun 2005	***0.72***	***0.87***	0.65	0.68	***0.71***	***0.81***	0.61	0.67	383.6	46.2%	188.7	58.5%	
12	Jul 2006	***0.82***	***1.00***	***0.79***	***0.84***	***0.82***	***0.93***	***0.77***	***0.83***	318.6	39.2%	145.8	40.9%	
13	Jun 2008	***0.79***	***0.96***	***0.84***	***0.81***	***0.83***	***0.94***	***0.88***	***0.86***	415.8	58.4%	197.5	65.9%	
14	Aug 2008	0.54	0.66	0.55	0.54	0.63	***0.71***	0.61	0.62	202.3	−1.0%	96.6	18.3%	
15	Aug 2013	0.56	0.68	0.61	0.66	0.62	***0.70***	0.64	0.72	303.5	48.6%	113.2	38.7%	
16	Aug 2014	0.54	0.66	0.56	0.43	0.63	***0.72***	0.62	0.51	213.6	4.6%	82.2	0.7%	
17	Jul 2015	0.53	0.64	0.16	0.21	0.62	***0.71***	0.17	0.29	243.6	6.4%	78.5	−24.1%	
18	Jun 2016	0.58	**0.70**	0.63	0.61	0.68	***0.77***	***0.71***	***0.71***	262.3	0.0%	142.6	19.7%	
19	Aug 2016	0.69	***0.84***	0.56	0.63	***0.80***	***0.91***	0.63	***0.74***	265.4	29.9%	84.7	3.7%	
20	Aug 2017	***0.78***	***0.94***	—	—	***0.88***	***1.00***	—	—	251.2	22.9%	107.0	31.1%	
21	Aug 2018	0.56	0.68	0.45	0.50	0.66	***0.75***	0.51	0.60	261.3	27.9%	74.1	−9.2%	
22	Jun 2019	0.61	***0.75***	0.65	0.63	***0.72***	***0.81***	***0.74***	***0.73***	303.2	15.5%	137.2	15.2%	
23	Jul 2019	***0.74***	***0.90***	***0.76***	***0.76***	***0.84***	***0.96***	***0.85***	***0.86***	288.6	26.1%	148.5	43.5%	
24	Aug 2019	0.62	***0.76***	0.65	0.66	***0.73***	***0.83***	***0.74***	***0.76***	207.3	1.5%	74.5	−8.8%	

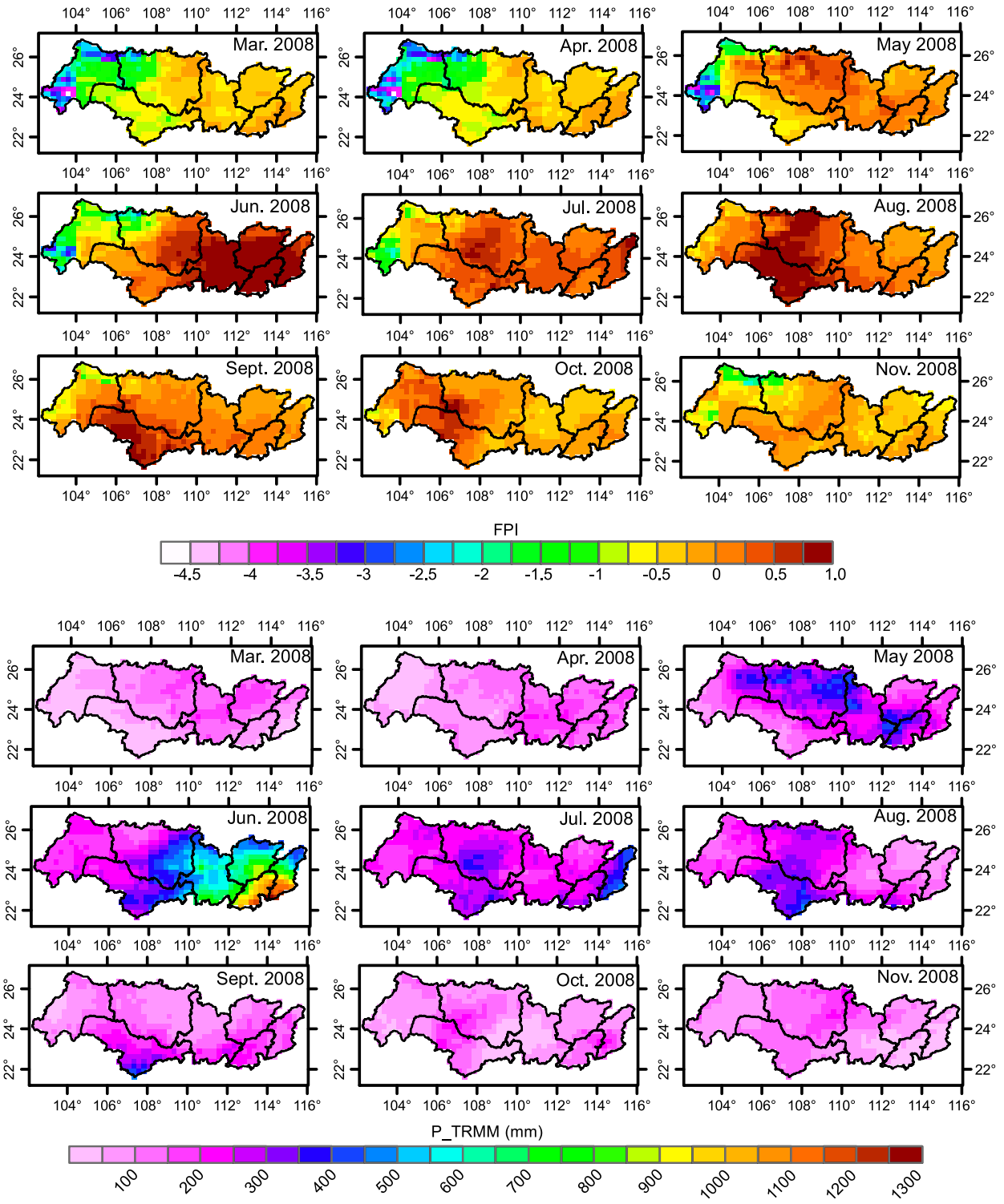

FIG. 8. (top) Spatial evolution of flood in the PRB based on flood potential index (FPI) estimated using CSR mascon TWSA and TRMM precipitation data during March–November 2008. Here, the contribution of RESS was not removed from TWSA. (bottom) Spatial evolution of precipitation in the PRB based on the TRMM data during March–November 2008.

which shows the FPI above 0.70. In October, the eastern PRB became dry, while the western part had high flood potential with the center at the southwestern Hongliu RB where the highest FPI was above 0.8. In November, the entire PRB returned to the dry status again.

The lower panel of Fig. 8 displays the gridded precipitation. It is obvious to find the consistency of the spatial pattern between FPI and precipitation with the occurrence of high and low values locating in the same region (e.g., the heavy rainfall and large FPI in the Hongliu RB in May and July, in the eastern PRB in June, in the Hongliu RB and Yu RB in August, etc.). This consistency is directly caused by the inclusion of precipitation data for calculating FPI. It is worth mentioning that the spatial pattern of FPI revealed by the CSR mascon solution can highlight the high flood potential at the subbasin scale below the typical GRACE footprint (90 000 km^2; Huang et al. 2019b; Vishwakarma et al. 2016; Watkins et al. 2015) as indicated above. This result is partly attributed to the use of high-resolution TRMM precipitation data for calculating FPI. Therefore, it does not implicate the suitability of the CSR mascon solution for small- or local-scale hydrological studies below the typical GRACE footprint like drought characterization and TWS variability considering the intrinsic larger leakage effects at smaller scales. Besides, although the CSR mascon-based FPI shows subscale spatial variability, its spatial resolution at several tens of thousands of square kilometers and the monthly temporal scale still cannot meet the demand for small- or even local-scale near-real time flood monitoring.

3) Impact of RESS change on flood characterization

The impact of RESS change on FPI estimation is different from that on drought characterization. As shown in Fig. 7a and Table 4, there is no apparent change in timing and amplitude between the FPI estimated with and without RESS contribution before 2007, but RESS enlarged the amplitude FPI after 2007 when RESS represented large seasonal amplitude due to human regulation. As shown in Table 4, the FPI estimated using both CSR and JPL mascon with RESS contribution identifies three more flood events (June 2016, June 2019, and August 2019) than that without RESS contribution. For these three flood events, the FPI should be overestimated considering the insignificant increment (<20%) of precipitation/streamflow over the climatology.

e. Drought and flood in response to climate variability

1) Cross-correlation between interannual precipitation/TWSA and climate indices

This section uses the interannual TWSA data for teleconnection analysis. To highlight the interannual variability of TWSA and climate indices, TWSA is deseasonalized and detrended through least squares harmonic analysis (Jin and Feng 2013; Pawlowicz et al. 2002). The deseasonalized and detrended (i.e., interannual) TWSA and the four climate indices are smoothed using a 5-month moving averaging similar to Ni et al. (2018) and Z. Zhang et al. (2015). A comparison is made between using the smoothed and nonsmoothed interannual TWSA to calculate their correlation with the four climate indices. As shown in Fig. S4, the smoothed interannual TWSA has a relatively higher correlation with the climate indices than the interannual TWSA without smoothing. This study also compares interannual TWSA and WSDI and their correlation with climate indices. Despite some differences in the correlation coefficients between TWSA and WSDI when compared with IOD and PDO, TWSA and WSDI exhibit the same capability on capturing the time lag in response to the four climate indices (Fig. S4). The cross-correlation analysis is conducted for the periods before and after 2002, and also the entire period during 1980–2019. Although MERRA-2 has overestimated the drought condition before 2002, the timing of MERRA-2 TWSA during 1987–2001 should be reliable considering the overall consistency between MERRA-2-based WSDI and scPDSI [see section 4c(2)]. Therefore, MERRA-2 TWSA before 2002 can be used for cross-correlation analysis. Here, monthly precipitation data are also deseasonalized, detrended, and smoothed to be used for cross-correlation analysis with climate indices.

The smoothed TWSA, precipitation, and the four climate indices are shown in Fig. 9. Figure 10 plots the cross-correlation coefficients between smoothed TWSA (also precipitation) and climate indices considering a 30-month lag. The time lag is determined at the month with maximum (minimum) positive (negative) correlation. A positive (negative) time lag means that TWSA lags behind (leads ahead) climate index. A summary of the time lag and the corresponding correlation coefficient is shown in Table 5. Precipitation leads one month ahead of ENSO for the period 1980–2001, whereas it lags one month behind ENSO for the period 2002–19. TWSA lags behind ENSO by one to three (two) months longer than precipitation with the lagged correlation coefficient above 0.4 (0.3) for the period 2002–19 (1980–2001) (Figs. 10a,e, Table 5). There are three peaks for the cross-correlation coefficients between PDO and precipitation during 2002–19, with the highest correlation of 0.32 at the lag of −13 months (Fig. 10b, Table 5). Also, significant correlation can be noted at the lag of −3 months (lagged $r = 0.29$) and 10 months (lagged $r = 0.30$). There are some discrepancies in the maximum correlation (i.e., time lag) between different TWSA data and PDO during 2002–19. TWSA from CSR mascon shows a time lag of 14 months after PDO with a significant correlation of 0.46, while TWSA from JPL mascon and MERRA-2 shows a time lag of −2 and −1 month(s) relative to PDO with the correlation of 0.41 and 0.42, respectively. Considering the hysteresis relationship between TWSA and precipitation, the TWSA should exhibit a longer time lag relative to PDO than precipitation. Therefore, both the lag of 14 months and the lag of −2 or −1 month(s) are possible for TWSA compared to the 10- and −3-month lags for precipitation in response to PDO. For the period 1980–2001, precipitation represents a lag of −6 months in response to PDO with a weaker significant correlation (lagged $r = 0.21$) than during the period 2002–19, while there is an insignificant relationship between PDO and TWSA (Figs. 10b,f, Table 5).

For the period 1980–2001, a significant correlation is found between the IOD and precipitation (MERRA-2 TWSA) with a maximum correlation of 0.34 (0.45) at a lag of 2 (5) months. For the period 2002–19, a long lag of over 30 months is found

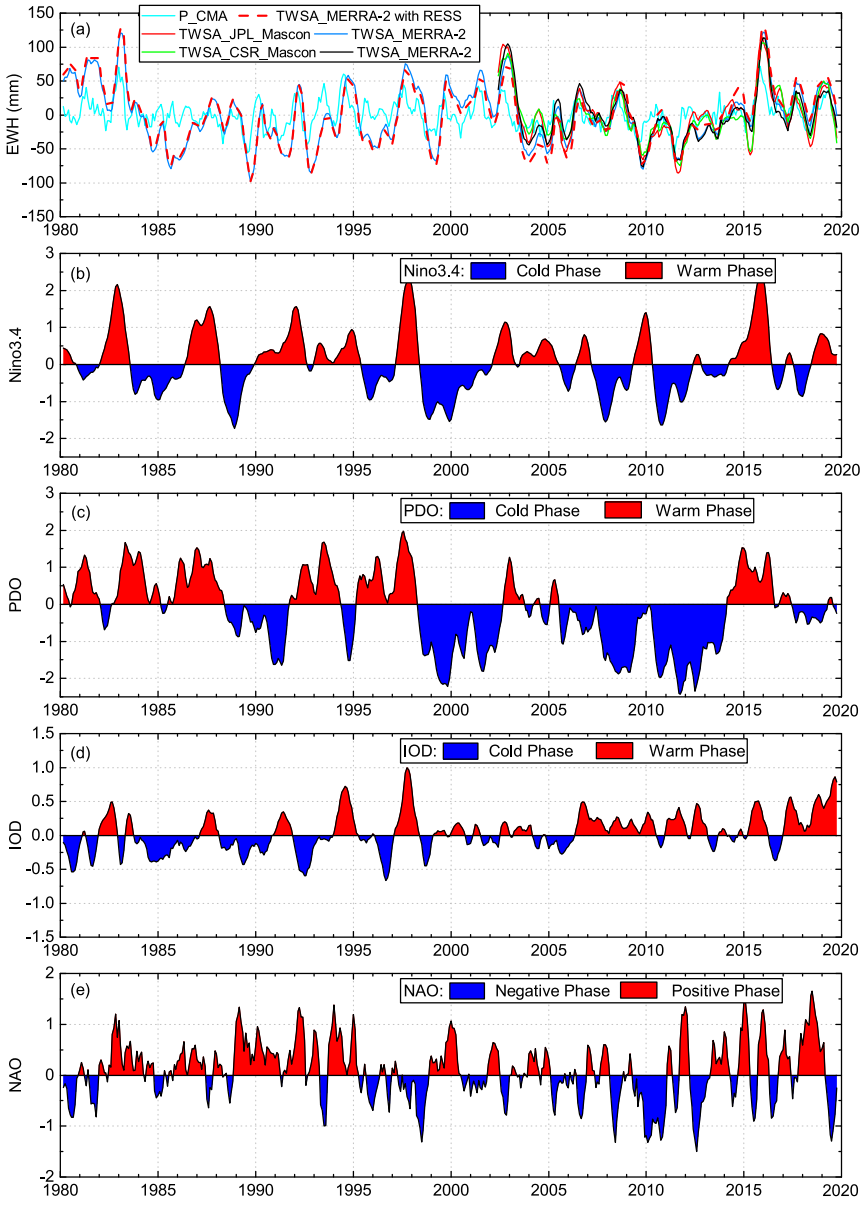

FIG. 9. (a) Deseasonalized and detrended time series of precipitation from CMA and TWSA from MERRA-2 and two GRACE mascon solutions. For comparison with GRACE results, the deseasonalized and detrended TWSA estimated using the MERRA-2 data only during 2002–19 is also plotted. (b) Anomaly time series for the Niño-3.4 index. (c) Anomaly time series for PDO index. (d) Anomaly time series for the IOD index. (e) Anomaly time series for the NAO index. All the time series in (a)–(e) are smoothed using a 5-month moving average. Please note that in (a), except for the notation "with RESS", all the other legends are the results without RESS. RESS means reservoir storage based on synthetic data (see section 2d).

between the IOD and precipitation (also TWSA) with a significant negative correlation of −0.34 (−0.25 to −0.31) (Figs. 10c,g, Table 5). Precipitation (TWSA) exhibits a lag of 9 (11–12) months after NAO during 2002–19 with a significant correlation above 0.30. The time lag of precipitation in response to NAO for the period 1980–2001 is −15 months (i.e., leading ahead) with a significant negative correlation of −0.29. TWSA represents a lag of 19 months after NAO during 1980–2001, but this lag seems to be indirectly linked to precipitation (Figs. 10d,h, Table 5). As seen, there is another trough for the cross-correlation between TWSA and NAO during 1980–2001, showing a significant negative correlation of −0.25 at a lag of −11 months (Fig. 10h, Table 5), which is more reasonable compared to the −15-month lag of precipitation in response to NAO.

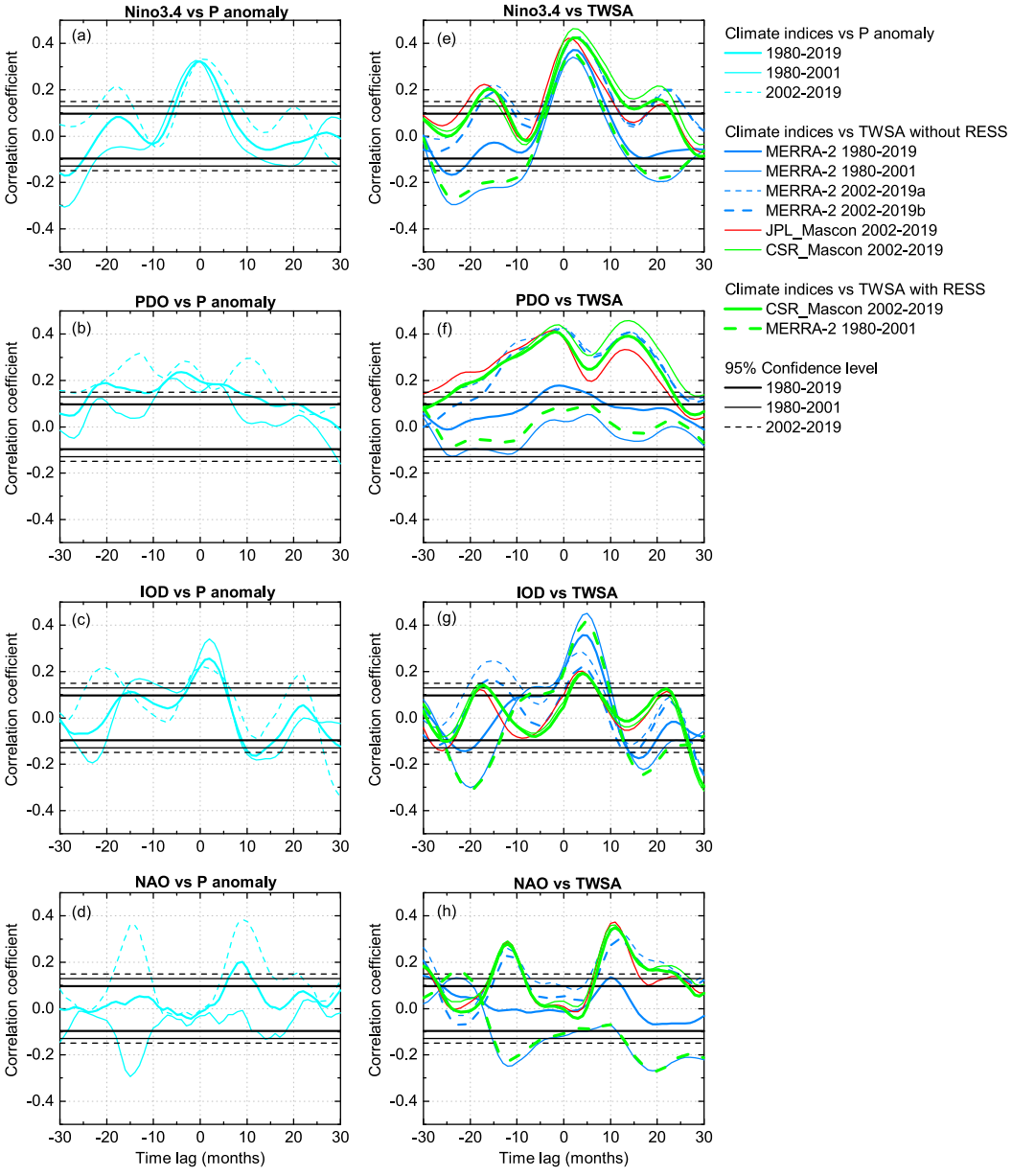

FIG. 10. (a)–(d) Cross-correlation coefficients between deseasonalized and detrended precipitation and the four climate indices for the period 1980–2019 and the two subperiods before and after 2002. (e)–(h) Cross-correlation coefficients between deseasonalized and detrended TWSA and the four climate indices for the period 1980–2019 and the two subperiods before and after 2002. Please note that the legend "MERRA-2 2002–2019a" means the results are derived using the data extracted from 1980 to 2019, while "MERRA-2 2002–2019b" means the results are derived using the data only during 2002–19. The notation "RESS" means reservoir storage based on synthetic data (see section 2d). The black lines show the upper and lower bound of the 95% confidence level determined based on a two-tailed t test. The time lag is determined at the lag of months with a maximum cross-correlation coefficient. The lag greater than zero means climate index leads to TWSA/precipitation. The time lag smaller than zero means TWSA/precipitation leads to climate index. The lag equal to zero means no time lag.

The above results indicate that the correlation between TWSA/precipitation and climate indices are different for the two periods before and after 2002, indicating the different role of climate variability during different periods. During 2002–19, precipitation and TWSA are positively correlated with ENSO, PDO, and NAO with a significant correlation above 0.30. During 1980–2001, ENSO and IOD (NAO) play a positive (negative) role in influencing precipitation and TWSA, while PDO represents an insignificant impact on TWSA. ENSO is considered to be the dominant climate index that affects precipitation and TWSA

TABLE 5. Matrix of time lag and the maximum cross-correlation coefficients between precipitation/TWSA and four climate indices during 1980–2019 and the two subperiods. The period "2002–19a" means the results are derived using the data extracted from 1980 to 2019, while "2002–19b" means the results are derived using the data only during 2002–19. The 95% confidence intervals for the periods 1980–2001, 2002–19, and 1980–2019 are ±0.13, ±0.15, and ±0.096, respectively.

		Niño-3.4		PDO		IOD		NAO	
Periods	Data	Lag (months)	*r*	Lag (months)	*r*	Lag (months)	*r*	Lag (months)	*r*
1980–2001	P_CMA	−1	0.33	−6	0.21	2	0.34	−15	−0.29
2002–19		1	0.33	−13	0.32	30	−0.34	9	0.38
1980–2019		0	0.32	−4	0.24	2	0.26	9	0.20
				TWSA without RESS					
1980–2001	MERRA-2	2	0.34	−24	−0.13	5	0.45	19	−0.27
2002–19a		3	0.43	−1	0.43	3	0.28	11	0.35
2002–19b		3	0.43	−1	0.42	30	−0.25	12	0.30
2002–19	CSR_Mascon	2	0.46	14	0.46	30	−0.31	11	0.36
2002–19	JPL_Mascon	1	0.42	−2	0.41	30	−0.29	11	0.37
1980–2019	MERRA-2	2	0.37	−1	0.18	4	0.36	−30	0.14
				TWSA with RESS					
1980–2001	MERRA-2	1	0.36	5	0.10	5	0.41	19	−0.27
2002–19	CSR_Mascon	2	0.42	−2	0.41	30	−0.30	11	0.35

considering the shorter time lag (≤3 months) and significant lagged correlation for the entire period 1980–2019. Although precipitation/TWSA represents a positive lagged correlation in response to ENSO, it does not mean that the occurrence of the warm (cool) phase of ENSO [i.e., El Niño (La Niña)] will result in a high (low) TWSA. As shown in Figs. 9a and 9b, only the droughts during 2003–05 and 2009 co-occurred with El Niño. The droughts in 1989, 1992–93, and 2006 were during the transition stage from La Niña to El Niño, and the droughts in 1999 and 2011 were under the impacts of La Niña. This phenomenon is associated with the asymmetric impacts of El Niño and La Niña on rainfall under the East Asian climate system (Chen et al. 2008; Z. Zhang et al. 2015). This asymmetric impact may be caused by the anomalous circulation over the western North Pacific in the lower troposphere that can be affected by the intraseasonal oscillation (R. Zhang et al. 2015). Further, the four climate indices are not independent of each other. Instead, they have a combined influence on the East Asia climate system and hence exert impacts on precipitation and TWSA in the PRB [see section 4e(3)].

2) IMPACT OF RESS CHANGE ON TELECONNECTION ANALYSIS

The above results do not consider RESS contribution. Figure 10 and Table 5 also show the results with RESS contribution. The interannual TWSA without RESS contribution represents a slightly higher cross-correlation with ENSO and PDO during 2002–19, and with PDO and IOD during 1980–2001 than that with RESS contribution. The slightly higher cross-correlation without RESS contribution is caused by excluding the impact of human regulation on RESS whose peak generally occurs later in October (see the peak of LTRS in Fig. 2b) than the peak in July of TWSA (see Fig. 7e). Overall, the impact of human regulation of RESS on the relationship between climate variability and interannual TWSA is minor. The inclusion of RESS contribution does not influence the determination of time lag based on the maximum correlation coefficient except for the relationship between PDO and TWSA. During 2002–19, the interannual TWSA using CSR mascon with (without) RESS contribution represents a lag of −2 (14) months in response to PDO with the lagged correlation of 0.41 (0.46). This may indicate that human regulation of RESS is more sensitive to PDO than other climate indices.

3) INFLUENCE OF CLIMATE VARIABILITY ON PRECIPITATION/TWSA

TWSA is directly regulated by precipitation, while precipitation is related to the atmospheric circulations that influence the monsoon system (Niu 2013; Niu et al. 2014). The East Asian monsoon (EAM) is the dominant climate system that affects the precipitation and TWSA in the PRB (Xiao et al. 2016; Zheng et al. 2017). The EAM generally begins in April and brings substantial (little) rainfall in summer (winter) through the strong southerly (northerly) winds (Chen et al. 2013). ENSO is the major factor that influences the EAM and finally controls the summer precipitation in the PRB. Considering the different atmospheric response to sea surface temperature, the influence of ENSO on the East Asian summer monsoon (EASM) is different during the developing and decaying phases of ENSO. In the developing phases of El Niño, southern China may experience dry conditions since the subtropical high moves southward triggered by the colder western tropical Pacific in summer, while opposite summer rainfall anomalies may occur after the peak of El Niño as the subtropical high shifts northward (Chen et al. 2013; Huang and Wu 1989). It should be noted that there is an unstable interannual relationship between ENSO and EASM, and they show significant multidecadal variations.

ENSO is the principal but not the only factor that affects EAM. Previous studies have demonstrated that EAM is jointly influenced by ENSO, PDO, IOD, and NAO (Xiao et al. 2016). As for PDO, it plays a critical role in modulating the relationship between the ENSO-related East Asian winter monsoon (EAWM) and EASM (Chen et al. 2013). During the cool phase of PDO, there are significantly low-level temperature

changes over eastern Asia due to the strong impact on the EAWM exerted by ENSO (Wang et al. 2008), whereas during the warm phase of PDO, a stronger EASM would follow the weak ENSO-related EAWM (Chen et al. 2013). Besides, Chan (2005) figured out that there tends to be below or above normal South China monsoon rainfall when ENSO and PDO are in phase. As for IOD, it usually co-occurs with ENSO (Du et al. 2009; Liu et al. 2020). IOD strongly influences the precipitation in the PRB (in particular for the central and eastern parts of the PRB; Niu 2013) through affecting the onset of the South China Sea summer monsoon and Asian summer monsoon in the following year (Yuan et al. 2008). During the positive (negative) phase of IOD, the Asian summer monsoon breaks out later (earlier) with higher (lower) intensity than normal. As a result, more precipitation occurs in the eastern and southern parts of China (Yan and Zhang 2004). It was also pointed out by Niu et al. (2014) that the dominant variabilities of runoff and soil moisture show a good correlation with IOD in the PRB. A connection between NAO and southern China precipitation was revealed by Zhou (2013), with increased (decreased) spring precipitation in southern China following a strong (weak) winter NAO. This connection was significantly weakened after the early 1980s due to the changing climatology of the East Asian jet. Moreover, spring NAO can also enhance the relationship between EASM and ENSO by inducing a tripole sea surface temperature anomaly in the North Atlantic that sustains from spring until summer, which develops the subpolar teleconnections downstream across northern Eurasia and finally results in a stronger or weaker EASM (Wu et al. 2009).

In addition to the four indices mentioned in this study, some other factors also influence the drought and flood in the PRB such as Tibetan Plateau heating (Tao and Ding 1981) and Arctic Oscillation (Barriopedro et al. 2012; Wang et al. 2015). For instance, the snow cover on the Tibetan Plateau can modulate the ENSO teleconnections and ENSO is significantly correlated with the EASM only during summer when snow cover on Tibetan Plateau is reduced (Wu et al. 2012).

5. Limitations and future directions

There are some limitations to this study. First, although long-term TWSA data are used for drought and flood characterization, no robust results are presented for the period before the GRACE mission due to the inaccurate long-term MERRA-2 reanalysis data. Nevertheless, this study provides caveats for future studies that focus on a hydrological study using MERRA-2 TWSA data from a long-term perspective. Unlike the previous papers (e.g., Awange et al. 2016; Ndehedehe et al. 2018) that directly use long-term MERRA-2 TWSA for spatiotemporal analysis, future studies should first examine the accuracy of long-term TWSA by validating the historical precipitation data (mainly for the period before 1998). There are also some other options for using long-term TWSA data such as 1) extending the time series by combining modeled TWSA and long-term precipitation and temperature data using relevant algorithms like neural networks (X. Chen et al. 2018; Long et al. 2014; Sun et al. 2020), random forest, and a spatially moving window structure (Jing et al. 2020); 2) reconstructing long-term TWSC time series using satellite/observation-based precipitation, evapotranspiration, and streamflow data based on a water balance approach (Pellet et al. 2020); and 3) using reliable model simulation like the LEAF-Hydro-Flood (LHF) model as used by Chaudhari et al. (2019). Besides, future MERRA-2 data assimilation system can include GRACE data to improve its outputs.

Second, this study only considers a single large reservoir with short monthly time series of observed water level data and constructs the synthetic long-term RESS time series based on simplified assumptions due to the lack of data. Considering the difficulty of obtaining observed reservoir water level/storage data, future studies could choose the water level data retrieved from altimetric satellites like *Envisat*, *ICESat*, *Jason*, *Cryosat-2*, and the planned Surface Water and Ocean Topography (SWOT) mission, etc. Combining the altimetric water level and the water area of reservoirs/lakes detected by optical satellites like Landsat (Pekel et al. 2016) or MODIS (Han and Niu 2020), the storage change can be estimated.

Third, this study only considers precipitation as the influencing factor for drought, flood characterization, and teleconnection analysis. Future studies could consider the impact of temperature on drought and flood under the warming climate from a long-term perspective. Fourth, the drought characterization and the teleconnection analysis only focus on the whole PRB on the basin-averaged basis. Previous studies have indicated that the drought, flood, and teleconnection in the PRB vary spatially as impacted by the weather system and topography (Y. Chen et al. 2018; Zheng et al. 2017). Future studies may take a deeper insight into the drought and flood characteristics at the subbasin scale. However, as the subbasin scale is below the GRACE footprint, the TWSA will contain large leakage errors, especially for those induced by large man-made reservoirs. Therefore, the recovery of the biased signal would be the major challenge for subbasin-scale studies using GRACE data. Fifth, considering the uncertainty in GRACE data, the large footprint, and limited temporal resolution, it is recommended to use GRACE as a tool for forecasting and warning future drought and flood events as indicated by Reager et al. (2014) rather than as a management tool, which is also suggested by Alley and Konikow (2015).

6. Summary

This study characterizes the drought and flood and analyzes their teleconnection to climate variability in the PRB using long-term GRACE (2002–19) and MERRA-2 reanalysis (1980–2019) data. GRACE TWSC is consistent with water balance–based TWSC. Reasonably good agreement is found between MERRA-2 and GRACE TWSA during 2002–19 except for some overestimation of amplitude in summer. The timing of MERRA-2-based WSDI is consistent with scPDSI during 1987–2001. However, the magnitude of MERRA-2 TWSA before 1998 is unreliable owing to the significant underestimation of precipitation.

Two major drought periods are identified over the last two decades: 2003–06 and 2009–12. The two GRACE mascon

solutions (JPL and CSR) show good agreement on identifying drought duration but have large discrepancies in quantifying total drought severity owing to the accumulation of monthly uncertainty. MERRA-2 does not well capture the 2009–12 drought duration and significantly overestimates the drought duration and severity during 1980–2000, in particular for the period 1985–87. Good agreement on timing is found between WSDI and scPDSI, both of which exhibit apparent time lags after precipitation anomalies, indicating the propagation from meteorological drought to hydrological drought. However, comparing to scPDSI, WSDI may underestimate the drought intensity by one category for over 40% of the identified drought events. Overall, the whole PRB is getting drier with more severe drought events after 2000 than before. This conclusion is consistent with the findings by Luo et al. (2017). Four major flood events are identified by GRACE mascon data: August 2002 (FPI = 1.0), followed by June 2008 (FPI > 0.80), July 2006 (FPI ≥ 0.79), and July 2019 (FPI = 0.76). MERRA-2 overestimates the flood potential in summer. Owing to the consideration of precipitation for calculating FPI, the spatial pattern of FPI based on CSR mascon data is consistent with precipitation (TRMM) at monthly and subbasin scales.

ENSO is the dominant climatic factor that affects precipitation and TWSA in the PRB, representing a lag from −1 to 3 months (the lag of TWSA may be longer than precipitation in response to ENSO) over the entire period 1980–2019. Different climate indices play different roles in influencing precipitation/TWSA before and after 2002. A significant correlation is found between PDO (or NAO) and precipitation/TWSA during 2002–19, and between IOD (or NAO) and precipitation/TWSA during 1980–2001, but with a longer time lag (from −13 to 30 months) than that between ENSO and precipitation/TWSA. EAM is the dominant climate system that affects the precipitation and TWSA in the PRB. EAM is mainly influenced by ENSO, with the combined impact from PDO, IOD, and NAO. The impacts of El Niño and La Niña on rainfall under the East Asian climate system are asymmetric. Generally, a drought or flood event is jointly influenced by different climate systems with different response times.

RESS change is of great significance for long-term drought and flood characterization. The inclusion of RESS contribution results in longer (shorter) drought duration and overestimated (underestimated) severity for the period 2003–06 (2009–10). RESS mainly influences the FPI estimation for the period after reservoir impoundment, leading to some overestimation. The overall impact of RESS change on cross-correlation analysis between TWSA and climate indices is minor. The inclusion of RESS contribution slightly enhances the correlation between TWSA and ENSO/PDO. Among the four indices, RESS change only modulates the lag for TWSA in response to PDO, which might indicate the sensitivity of human activity to PDO.

Acknowledgments. This study was funded by the National Natural Science Foundation of China (41877172) and Seed fund program granted to HKU-ZIRI. The authors declare no conflict of interest.

Data availability statement. The data used in this study can be freely accessed from the following links: GRACE CSR mascon data: http://www2.csr.utexas.edu/grace/; GRACE JPL mascon data: https://grace.jpl.nasa.gov/data/get-data/; GSFC mascon data: https://earth.gsfc.nasa.gov/geo/; MERRA-2 data: https://disc.gsfc.nasa.gov/; CMA Precipitation data: http://www.nmic.cn/; MODIS ET data: http://files.ntsg.umt.edu/data/MOD16_TP/; GLEAM ET data: https://www.gleam.eu/; Streamflow data (China River Sediment Bulletin): http://www.mwr.gov.cn/sj/tjgb/zghlnsgb/; ScPDSI and CRU precipitation data: https://crudata.uea.ac.uk/cru/data/drought/; Niño 3.4 and NAO index data: https://www.cpc.ncep.noaa.gov/; IOD and PDO index data: https://www.esrl.noaa.gov/.

APPENDIX

Acronyms and Definitions

13-MA	13-point moving average
CMA	China Meteorological Administration
CRU	Climatic Research Unit
CSR	Center for Space Research
EAM	East Asia monsoon
EASM	East Asia summer monsoon
EAWM	East Asia winter monsoon
ENSO	El Niño–Southern Oscillation
ET	Evapotranspiration
EWH	Equivalent water height
FPA	Flood potential amount
FPI	Flood potential index
GLDAS	Global Land Data Assimilation System
GLEAM	Global Land Evaporation Amsterdam Model
GRACE	Gravity Recovery and Climate Experiment
GSFC	Goddard Space Flight Center
IESA	Integrated Earth System Analysis
IOD	Indian Ocean dipole
JPL	Jet Propulsion Laboratory
LHF	LEAF-Hydro-Flood
LTR	Longtan Reservoir
LTRS	Longtan Reservoir storage
Mascon	Mass concentration
MODIS	Moderate Resolution Imaging Spectroradiometer
NAO	North Atlantic Oscillation
P	Precipitation
PAC	Percentage of precipitation or streamflow above the climatology
PDO	Pacific decadal oscillation
PRB	Pearl River basin
R	Streamflow or runoff
RB	River basin
RESS	Reservoir storage
RMSD	Root-mean-square difference
SWOT	Surface Water and Ocean Topography
TRMM	Tropical Rainfall Measuring Mission
TWSA	Total terrestrial water storage anomaly
WSD	Water storage deficit
WSDI	Water storage deficit index

REFERENCES

A, G., J. Wahr, and S. Zhong, 2012: Computations of the viscoelastic response of a 3-D compressible Earth to surface loading: An application to glacial isostatic adjustment in Antarctica and Canada. *Geophys. J. Int.*, **192**, 557–572, https://doi.org/10.1093/GJI/GGS030.

Abelen, S., F. Seitz, R. Abarca-del-Rio, and A. Güntner, 2015: Droughts and floods in the La Plata basin in soil moisture data and GRACE. *Remote Sens.*, **7**, 7324–7349, https://doi.org/10.3390/rs70607324.

AghaKouchak, A., A. Farahmand, F. S. Melton, J. Teixeira, M. C. Anderson, B. D. Wardlow, and C. R. Hain, 2015: Remote sensing of drought: Progress, challenges, and opportunities. *Rev. Geophys.*, **53**, 452–480, https://doi.org/10.1002/2014RG000456.

Ahi, G. O., and S. Jin, 2019: Hydrologic mass changes and their implications in Mediterranean-climate Turkey from GRACE measurements. *Remote Sens.*, **11**, 120, https://doi.org/10.3390/rs11020120.

Alley, W. M., and L. F. Konikow, 2015: Bringing GRACE down to earth. *Ground Water*, **53**, 826–829, https://doi.org/10.1111/gwat.12379.

Awange, J. L., E. Forootan, M. Kuhn, J. Kusche, and B. Heck, 2014: Water storage changes and climate variability within the Nile basin between 2002 and 2011. *Adv. Water Resour.*, **73**, 1–15, https://doi.org/10.1016/j.advwatres.2014.06.010.

——, Khandu, M. Schumacher, E. Forootan, and B. Heck, 2016: Exploring hydro-meteorological drought patterns over the Greater Horn of Africa (1979–2014) using remote sensing and reanalysis products. *Adv. Water Resour.*, **94**, 45–59, https://doi.org/10.1016/j.advwatres.2016.04.005.

Barriopedro, D., C. M. Gouveia, R. M. Trigo, and L. Wang, 2012: The 2009/10 drought in China: Possible causes and impacts on vegetation. *J. Hydrometeor.*, **13**, 1251–1267, https://doi.org/10.1175/JHM-D-11-074.1.

Bosilovich, M. G., and Coauthors, 2015: MERRA-2: Initial evaluation of the climate. NASA Tech. Memo. NASA/TM-2015-104606, Vol. 43, 145 pp., https://gmao.gsfc.nasa.gov/pubs/docs/Bosilovich803.pdf.

——, R. Lucchesi and M. Suarez, 2016: MERRA-2: File specification. GMAO Office Note 9 (version 1.1), 73 pp., https://gmao.gsfc.nasa.gov/pubs/docs/Bosilovich785.pdf.

Chan, J. C. L., 2005: PDO, ENSO, and the early summer monsoon rainfall over south China. *Geophys. Res. Lett.*, **32**, L08810, https://doi.org/10.1029/2004GL022015.

Chaudhari, S., Y. Pokhrel, E. Moran, and G. Miguez-Macho, 2019: Multi-decadal hydrologic change and variability in the Amazon River basin: Understanding terrestrial water storage variations and drought characteristics. *Hydrol. Earth Syst. Sci.*, **23**, 2841–2862, https://doi.org/10.5194/hess-23-2841-2019.

Chen, J. L., C. R. Wilson, B. D. Tapley, Z. L. Yang, and G. Y. Niu, 2009: 2005 drought event in the Amazon River basin as measured by GRACE and estimated by climate models. *J. Geophys. Res.*, **114**, B05404, https://doi.org/10.1029/2008JB006056.

——, ——, and ——, 2010: The 2009 exceptional Amazon flood and interannual terrestrial water storage change observed by GRACE. *Water Resour. Res.*, **46**, W12526, https://doi.org/10.1029/2010WR009383.

Chen, J.-M., T. Li, and C. Shih, 2008: Asymmetry of the El Niño–spring rainfall relationship in Taiwan. *J. Meteor. Soc. Japan*, **86**, 297–312, https://doi.org/10.2151/jmsj.86.297.

Chen, W., J. Feng, and R. Wu, 2013: Roles of ENSO and PDO in the link of the East Asian winter monsoon to the following summer monsoon. *J. Climate*, **26**, 622–635, https://doi.org/10.1175/JCLI-D-12-00021.1.

Chen, X., J. Jiang, and H. Li, 2018: Drought and flood monitoring of the Liao River basin in Northeast China using extended GRACE data. *Remote Sens.*, **10**, 1168, https://doi.org/10.3390/rs10081168.

Chen, Y., L. Guo, C. Ye, and L. Zhu, 2018: Spatial-temporal characteristics of the drought and flood in Pearl River basin (in Chinese). *China Rural Water Hydropower*, 113–120+125.

Cheng, M., J. C. Ries, and B. D. Tapley, 2011: Variations of the Earth's figure axis from satellite laser ranging and GRACE. *J. Geophys. Res.*, **116**, B01409, https://doi.org/10.1029/2010JB000850.

——, B. D. Tapley, and J. C. Ries, 2013: Deceleration in the Earth's oblateness. *J. Geophys. Res. Solid Earth*, **118**, 740–747, https://doi.org/10.1002/jgrb.50058.

Cui, W. Z., J. Chen, Y. P. Wu, and Y. D. Wu, 2007: An overview of water resources management of the Pearl River. *Water Sci. Technol. Water Supply*, **7**, 101–113, https://doi.org/10.2166/ws.2007.045.

Du, Y., S. Xie, G. Huang, and K. Hu, 2009: Role of air–sea interaction in the long persistence of El Niño–induced north Indian ocean warming. *J. Climate*, **22**, 2023–2038, https://doi.org/10.1175/2008JCLI2590.1.

Editorial Board of China Agriculture Yearbook, 1980–2017 (published yearly): *China Agriculture Yearbook* (in Chinese). China Agriculture Press, https://data.cnki.net/area/Yearbook.

Editorial Board of China Water Conservancy Yearbook, 2018–2019: *China Water Conservancy Yearbook* (in Chinese), China Water & Power Press, https://data.cnki.net/area/Yearbook.

Forootan, E., and Coauthors, 2019: Understanding the global hydrological droughts of 2003–2016 and their relationships with teleconnections. *Sci. Total Environ.*, **650**, 2587–2604, https://doi.org/10.1016/j.scitotenv.2018.09.231.

Gelaro, R., and Coauthors, 2017: The Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2). *J. Climate*, **30**, 5419–5454, https://doi.org/10.1175/JCLI-D-16-0758.1.

Han, Q., and Z. Niu, 2020: Construction of the long-term global surface water extent dataset based on water-NDVI spatiotemporal parameter set. *Remote Sens.*, **12**, 2675, https://doi.org/10.3390/rs12172675.

Han, Z., and Coauthors, 2019: Propagation dynamics from meteorological to groundwater drought and their possible influence factors. *J. Hydrol.*, **578**, 124102, https://doi.org/10.1016/j.jhydrol.2019.124102.

Heddinghaus, T. R., and P. Sabol, 1991: A review of the Palmer Drought Severity Index and where do we go from here. *Proc. Seventh Conf. on Applied Climatology*, Boston, MA, Amer. Meteor. Soc., 242–246.

Heumann, B. W., 2011: Satellite remote sensing of mangrove forests: Recent advances and future opportunities. *Prog. Phys. Geogr.*, **35**, 87–108, https://doi.org/10.1177/0309133310385371.

Hirabayashi, Y., R. Mahendran, S. Koirala, L. Konoshima, D. Yamazaki, S. Watanabe, H. Kim, and S. Kanae, 2013: Global flood risk under climate change. *Nat. Climate Change*, **3**, 816–821, https://doi.org/10.1038/nclimate1911.

Hosseini-Moghari, S., S. Araghinejad, K. Ebrahimi, and M. J. Tourian, 2019: Introducing modified total storage deficit index (MTSDI) for drought monitoring using GRACE observations. *Ecol. Indic.*, **101**, 465–475, https://doi.org/10.1016/j.ecolind.2019.01.002.

Huang, R., and Y. Wu, 1989: The influence of ENSO on the summer climate change in China and its mechanism. *Adv. Atmos. Sci.*, **6**, 21–32, https://doi.org/10.1007/BF02656915.

Huang, Z., and Coauthors, 2019a: Detection of large-scale groundwater storage variability over the karstic regions in southwest China. *J. Hydrol.*, **569**, 409–422, https://doi.org/10.1016/j.jhydrol.2018.11.071.

——, J. J. Jiao, X. Luo, Y. Pan, and C. Zhang, 2019b: Sensitivity analysis of leakage correction of GRACE data in southwest China using a-priori model simulations: Inter-comparison of spherical harmonics, mass concentration and in situ observations. *Sensors*, **19**, 3149, https://doi.org/10.3390/s19143149.

Huffman, G. J., and Coauthors, 2007: The TRMM Multi-satellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeor.*, **8**, 38–55, https://doi.org/10.1175/JHM560.1.

Hurrell, J. W., Y. Kushnir, G. Ottersen, and M. Visbeck, 2003: An overview of the North Atlantic Oscillation. The *North Atlantic Oscillation: Climatic Significance and Environmental Impact*, *Geophys. Monogr.*, Vol. 134, Amer. Geophys. Union, 1–35.

Idowu, D., and W. Zhou, 2019: Performance evaluation of a potential component of an early flood warning system—A case study of the 2012 flood, lower Niger River basin, Nigeria. *Remote Sens.*, **11**, 1970, https://doi.org/10.3390/rs11171970.

Jin, S., and G. Feng, 2013: Large-scale variations of global groundwater from satellite gravimetry and hydrological models, 2002–2012. *Global Planet. Change*, **106**, 20–30, https://doi.org/10.1016/j.gloplacha.2013.02.008.

Jing, W., P. Zhang, X. Zhao, Y. Yang, H. Jiang, J. Xu, J. Yang, and Y. Li, 2020: Extending GRACE terrestrial water storage anomalies by combining the random forest regression and a spatially moving window structure. *J. Hydrol.*, **590**, 125239, https://doi.org/10.1016/j.jhydrol.2020.125239.

Karl, T. R., 1986: The sensitivity of the Palmer Drought Severity Index and Palmer's Z-index to their calibration coefficients including potential evapotranspiration. *J. Climate Appl. Meteor.*, **25**, 77–86, https://doi.org/10.1175/1520-0450(1986)025<0077:TSOTPD>2.0.CO;2.

Liu, L. L., T. Jiang, J. G. Xu, J. Q. Zhai, and L. Yong, 2012: Responses of hydrological processes to climate change in the Zhujiang River basin in the 21st century. *Adv. Climate Change Res.*, **3**, 84–91, https://doi.org/10.3724/SP.J.1248.2012.00084.

Liu, X., X. Feng, P. Ciais, and B. Fu, 2020: Widespread decline in terrestrial water storage and its link to teleconnections across Asia and eastern Europe. *Hydrol. Earth Syst. Sci.*, **24**, 3663–3676, https://doi.org/10.5194/hess-24-3663-2020.

Long, D., Y. Shen, A. Sun, Y. Hong, L. Longuevergne, Y. Yang, B. Li, and L. Chen, 2014: Drought and flood monitoring for a large karst plateau in Southwest China using extended GRACE data. *Remote Sens. Environ.*, **155**, 145–160, https://doi.org/10.1016/j.rse.2014.08.006.

Luo, L., T. Jiang, H. Sun, C. Jing, and B. Su, 2017: Variation of droughts and floods and their connection with atmospheric circulation in the Pearl River basin (in Chinese with English abstract). *J. Arid Land Resour. Environ.*, **31**, 142–147.

Luo, Z., C. Yao, Q. Li, and Z. Huang, 2016: Terrestrial water storage changes over the Pearl River basin from GRACE and connections with Pacific climate variability. *Geod. Geodyn.*, **7**, 171–179, https://doi.org/10.1016/j.geog.2016.04.008.

Luthcke, S. B., T. J. Sabaka, B. D. Loomis, A. A. Arendt, J. J. McCarthy, and J. Camp, 2013: Antarctica, Greenland, and Gulf of Alaska land-ice evolution from an iterated GRACE global mascon solution. *J. Glaciol.*, **59**, 613–631, https://doi.org/10.3189/2013JoG12J147.

Mantua, N. J., and S. R. Hare, 2002: The Pacific decadal oscillation. *J. Oceanogr.*, **58**, 35–44, https://doi.org/10.1023/A:1015820616384.

Martens, B., and Coauthors, 2017: GLEAM v3: Satellite-based land evaporation and root-zone soil moisture. *Geosci. Model Dev.*, **10**, 1903–1925, https://doi.org/10.5194/gmd-10-1903-2017.

McKee, T. B., 1993: The relationship of drought frequency and duration to time scales. *Proc. Eighth Conference on Applied Climatology*, Anaheim, CA, Amer. Meteor. Soc., 179–184.

Molodtsova, T., S. Molodtsov, A. Kirilenko, X. Zhang, and J. VanLooy, 2016: Evaluating flood potential with GRACE in the United States. *Nat. Hazard. Earth Sys.*, **16**, 1011–1018, https://doi.org/10.5194/nhess-16-1011-2016.

Monteith, J. L., 1965: Evaporation and environment. *The State and Movement of Water in Living Organisms*, B. D. Fogg, Ed., Cambridge University Press, 205–234.

Morid, S., V. Smakhtin, and M. Moghaddasi, 2006: Comparison of seven meteorological indices for drought monitoring in Iran. *Int. J. Climatol.*, **26**, 971–985, https://doi.org/10.1002/joc.1264.

Mu, Q., F. A. Heinsch, M. Zhao, and S. W. Running, 2007: Development of a global evapotranspiration algorithm based on MODIS and global meteorology data. *Remote Sens. Environ.*, **111**, 519–536, https://doi.org/10.1016/j.rse.2007.04.015.

——, M. Zhao, and S. W. Running, 2011: Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Remote Sens. Environ.*, **115**, 1781–1800, https://doi.org/10.1016/j.rse.2011.02.019.

Narasimhan, B., and R. Srinivasan, 2005: Development and evaluation of Soil Moisture Deficit Index (SMDI) and Evapotranspiration Deficit Index (ETDI) for agricultural drought monitoring. *Agric. For. Meteor.*, **133**, 69–88, https://doi.org/10.1016/j.agrformet.2005.07.012.

Ndehedehe, C. E., J. L. Awange, M. Kuhn, N. O. Agutu, and Y. Fukuda, 2017: Climate teleconnections influence on West Africa's terrestrial water storage. *Hydrol. Processes*, **31**, 3206–3224, https://doi.org/10.1002/hyp.11237.

——, ——, N. O. Agutu, and O. Okwuashi, 2018: Changes in hydrometeorological conditions over tropical West Africa (1980–2015) and links to global climate. *Global Planet. Change*, **162**, 321–341, https://doi.org/10.1016/j.gloplacha.2018.01.020.

Ni, S., J. Chen, C. R. Wilson, J. Li, X. Hu, and R. Fu, 2018: Global terrestrial water storage changes and connections to ENSO events. *Surv. Geophys.*, **39**, 1–22, https://doi.org/10.1007/s10712-017-9421-7.

Niu, J., 2010: A comprehensive analysis of terrestrial hydrological processes over the Pearl River basin in South China. Ph.D. thesis, The University of Hong Kong, 202 pp.

——, 2013: Precipitation in the Pearl River basin, South China: Scaling, regional patterns, and influence of large-scale climate anomalies. *Stochastic Environ. Res. Risk Assess.*, **27**, 1253–1268, https://doi.org/10.1007/s00477-012-0661-2.

——, J. Chen, and B. Sivakumar, 2014: Teleconnection analysis of runoff and soil moisture over the Pearl River basin in southern China. *Hydrol. Earth Syst. Sci.*, **18**, 1475–1492, https://doi.org/10.5194/hess-18-1475-2014.

Palmer, W. C., 1965: Meteorological drought. Research Paper 45, U.S. Department of Commerce, Weather Bureau, 58 pp., https://www.ncdc.noaa.gov/temp-and-precip/drought/docs/palmer.pdf.

Pawlowicz, R., B. Beardsley, and S. Lentz, 2002: Classical tidal harmonic analysis including error estimates in MATLAB using T_TIDE. *Comput. Geosci.*, **28**, 929–937, https://doi.org/10.1016/S0098-3004(02)00013-4.

Pearl River Water Conservancy Commission, 2003–2018: *Water Resources Bulletin of Pearl River Basin* (in Chinese). Pearl River

Water Conservancy Commission, http://www.pearlwater.gov.cn/zwgkcs/lygb/szygb/.

——, 2005: *Pearl River Flood Prevention Handbook*. Pearl River Water Resource Commission, 150 pp.

Pekel, J., A. Cottam, N. Gorelick, and A. S. Belward, 2016: High-resolution mapping of global surface water and its long-term changes. *Nature*, **540**, 418–422, https://doi.org/10.1038/nature20584.

Pellet, V., F. Aires, F. Papa, S. Munier, and B. Decharme, 2020: Long-term total water storage change from a satellite water cycle reconstruction over large southern Asian basins. *Hydrol. Earth Syst. Sci.*, **24**, 3033–3055, https://doi.org/10.5194/hess-24-3033-2020.

Peltier, W. R., 2004: Global glacial isostatic and the surface of the ice-age Earth: The ICE-5G (VM2) model and GRACE. *Annu. Rev. Earth Planet. Sci.*, **32**, 111–149, https://doi.org/10.1146/annurev.earth.32.082503.144359.

——, D. F. Argus, and R. Drummond, 2018: Comment on ''An assessment of the ICE-6G_C (VM5a) glacial isostatic adjustment model'' by Purcell et al. *J. Geophys. Res. Solid Earth*, **123**, 2019–2028, https://doi.org/10.1002/2016JB013844.

Phillips, T., R. S. Nerem, B. Fox-Kemper, J. S. Famiglietti, and B. Rajagopalan, 2012: The influence of ENSO on global terrestrial water storage using GRACE. *Geophys. Res. Lett.*, **39**, L16705, https://doi.org/10.1029/2012GL052495.

Reager, J. T., and J. S. Famiglietti, 2009: Global terrestrial water storage capacity and flood potential using GRACE. *Geophys. Res. Lett.*, **36**, L23402, https://doi.org/10.1029/2009GL040826.

——, B. F. Thomas, and J. S. Famiglietti, 2014: River basin flood potential inferred using GRACE gravity observations at several months lead time. *Nat. Geosci.*, **7**, 588–592, https://doi.org/10.1038/ngeo2203.

Reichle, R. H., C. S. Draper, Q. Liu, M. Girotto, S. P. P. Mahanama, R. D. Koster, and G. J. M. De Lannoy, 2017: Assessment of MERRA-2 land surface hydrology estimates. *J. Climate*, **30**, 2937–2960, https://doi.org/10.1175/JCLI-D-16-0720.1.

Saji, N. H., B. N. Goswami, P. N. Vinayachandran, and T. Yamagata, 1999: A dipole mode in the tropical Indian Ocean. *Nature*, **401**, 360–363, https://doi.org/10.1038/43854.

Save, H., S. Bettadpur, and B. D. Tapley, 2016: High-resolution CSR GRACE RL05 mascons. *J. Geophys. Res. Solid Earth*, **121**, 7547–7569, https://doi.org/10.1002/2016JB013007.

Schneider, N., and B. D. Cornuelle, 2005: The forcing of the Pacific decadal oscillation. *J. Climate*, **18**, 4355–4373, https://doi.org/10.1175/JCLI3527.1.

Shukla, S., and A. W. Wood, 2008: Use of a standardized runoff index for characterizing hydrologic drought. *Geophys. Res. Lett.*, **35**, L02405, https://doi.org/10.1029/2007GL032487.

Sinha, D., T. H. Syed, J. S. Famiglietti, J. T. Reager, and R. C. Thomas, 2017: Characterizing drought in India using GRACE observations of terrestrial water storage deficit. *J. Hydrometeor.*, **18**, 381–396, https://doi.org/10.1175/JHM-D-16-0047.1.

——, ——, and J. T. Reager, 2019: Utilizing combined deviations of precipitation and GRACE-based terrestrial water storage as a metric for drought characterization: A case study over major Indian river basins. *J. Hydrol.*, **572**, 294–307, https://doi.org/10.1016/j.jhydrol.2019.02.053.

Sun, Z., X. Zhu, Y. Pan, and J. Zhang, 2017: Assessing terrestrial water storage and flood potential using GRACE data in the Yangtze River basin, China. *Remote Sens.*, **9**, 1011, https://doi.org/10.3390/rs9101011.

——, ——, ——, ——, and X. Liu, 2018: Drought evaluation using the GRACE terrestrial water storage deficit over the Yangtze River basin, China. *Sci. Total Environ.*, **634**, 727–738, https://doi.org/10.1016/j.scitotenv.2018.03.292.

——, D. Long, W. Yang, X. Li, and Y. Pan, 2020: Reconstruction of GRACE data on changes in total water storage over the global land surface and 60 basins. *Water Resour. Res.*, **56**, e2019WR026250, https://doi.org/10.1029/2019WR026250.

Swenson, S., D. Chambers, and J. Wahr, 2008: Estimating center variations from a combination of GRACE and ocean model output. *J. Geophys. Res.*, **113**, B08410, https://doi.org/10.1029/2007JB005338.

Tao, S., and Y. Ding, 1981: Observational evidence of the influence of the Qinghai-Xizang (Tibet) Plateau on the occurrence of heavy rain and severe convective storms in China. *Bull. Amer. Meteor. Soc.*, **62**, 23–30, https://doi.org/10.1175/1520-0477(1981)062<0023:OEOTIO>2.0.CO;2.

Tapley, B. D., S. Bettadpur, M. Watkins, and C. Reigber, 2004: The Gravity Recovery and Climate Experiment: Mission overview and early results. *Geophys. Res. Lett.*, **31**, L09607, https://doi.org/10.1029/2004GL019920.

Thomas, A. C., J. T. Reager, J. S. Famiglietti, and M. Rodell, 2014: A GRACE-based water storage deficit approach for hydrological drought characterization. *Geophys. Res. Lett.*, **41**, 1537–1545, https://doi.org/10.1002/2014GL059323.

van der Schrier, G., J. Barichivich, K. R. Briffa, and P. D. Jones, 2013: A scPDSI-based global data set of dry and wet spells for 1901–2009. *J. Geophys. Res. Atmos.*, **118**, 4025–4048, https://doi.org/10.1002/jgrd.50355.

Vishwakarma, B. D., B. Devaraju, and N. Sneeuw, 2016: Minimizing the effects of filtering on catchment-scale GRACE solutions. *Water Resour. Res.*, **52**, 5868–5890, https://doi.org/10.1002/2016WR018960.

Wang, L., W. Chen, and R. Huang, 2008: Interdecadal modulation of PDO on the impact of ENSO on the East Asian winter monsoon. *Geophys. Res. Lett.*, **35**, L20702, https://doi.org/10.1029/2008GL035287.

——, ——, W. Zhou, and G. Huang, 2015: Drought in southwest China: A review. *Atmos. Oceanic Sci. Lett.*, **8**, 339–344, https://doi.org/10.3878/AOSL20150043.

Watkins, M. M., D. N. Wiese, D. Yuan, C. Boening, and F. W. Landerer, 2015: Improved methods for observing Earth's time variable mass distribution with GRACE using spherical cap mascons. *J. Geophys. Res. Solid Earth*, **120**, 2648–2671, https://doi.org/10.1002/2014JB011547.

Wells, N., S. Goddard, and M. J. Hayes, 2004: A self-calibrating Palmer drought severity index. *J. Climate*, **17**, 2335–2351, https://doi.org/10.1175/1520-0442(2004)017<2335:ASPDSI>2.0.CO;2.

Winsemius, H. C., and Coauthors, 2016: Global drivers of future river flood risk. *Nat. Climate Change*, **6**, 381–385, https://doi.org/10.1038/nclimate2893.

Wu, Z., B. Wang, J. Li, and F. Jin, 2009: An empirical seasonal prediction model of the East Asian summer monsoon using ENSO and NAO. *J. Geophys. Res.*, **114**, D18120, https://doi.org/10.1029/2009JD011733.

——, J. Li, Z. Jiang, and T. Ma, 2012: Modulation of the Tibetan Plateau snow cover on the ENSO teleconnections: From the East Asian summer monsoon perspective. *J. Climate*, **25**, 2481–2489, https://doi.org/10.1175/JCLI-D-11-00135.1.

Xiao, M., Q. Zhang, V. P. Singh, and L. Liu, 2016: Transitional properties of droughts and related impacts of climate indices in the Pearl River basin, China. *J. Hydrol.*, **534**, 397–406, https://doi.org/10.1016/j.jhydrol.2016.01.012.

Yan, X., and M. Zhang, 2004: A study of the Indian Ocean dipole influence on climate variations over East Asian monsoon region. *Climate Environ. Res.*, **9**, 435–444.

Yirdaw, S. Z., K. R. Snelgrove, and C. O. Agboma, 2008: GRACE satellite observations of terrestrial moisture changes for drought characterization in the Canadian prairie. *J. Hydrol.*, **356**, 84–92, https://doi.org/10.1016/j.jhydrol.2008.04.004.

Yuan, Y., H. Yang, W. Zhou, and C. Li, 2008: Influences of the Indian Ocean dipole on the Asian summer monsoon in the following year. *Int. J. Climatol.*, **28**, 1849–1859, https://doi.org/10.1002/joc.1678.

Zargar, A., R. Sadiq, B. Naser, and F. I. Khan, 2011: A review of drought indices. *Environ. Rev.*, **19**, 333–349, https://doi.org/10.1139/a11-013.

Zhang, Q., C. Y. Xu, S. Becker, Z. X. Zhang, Y. D. Chen, and M. Coulibaly, 2009: Trends and abrupt changes of precipitation maxima in the Pearl River basin, China. *Atmos. Sci. Lett.*, **10**, 132–144, https://doi.org/10.1002/asl.221.

Zhang, R., T. Li, M. Wen, and L. Liu, 2015: Role of intraseasonal oscillation in asymmetric impacts of El Niño and La Niña on the rainfall over southern China in boreal winter. *Climate Dyn.*, **45**, 559–567, https://doi.org/10.1007/s00382-014-2207-4.

Zhang, Z., B. F. Chao, J. Chen, and C. R. Wilson, 2015: Terrestrial water storage anomalies of Yangtze River basin droughts observed by GRACE and connections with ENSO. *Global Planet. Change*, **126**, 35–45, https://doi.org/10.1016/j.gloplacha.2015.01.002.

Zhao, M., A. Geruo, I. Velicogna, and J. S. Kimball, 2017: Satellite observations of regional drought severity in the continental United States using GRACE-based terrestrial water storage changes. *J. Climate*, **30**, 6297–6308, https://doi.org/10.1175/JCLI-D-16-0458.1.

Zheng, Y., Y. He, and X. Chen, 2017: Spatiotemporal pattern of precipitation concentration and its possible causes in the Pearl River basin, China. *J. Clean. Prod.*, **161**, 1020–1031, https://doi.org/10.1016/j.jclepro.2017.06.156.

Zhou, B., 2013: Weakening of winter North Atlantic oscillation signal in spring precipitation over southern China. *Atmos. Oceanic Sci. Lett.*, **6**, 248–252, https://doi.org/10.1080/16742834.2013.11447089.

Zhou, H., Z. Luo, N. Tangdamrongsub, L. Wang, L. He, C. Xu, and Q. Li, 2017: Characterizing drought and flood events over the Yangtze River basin using the HUST-Grace2016 solution and ancillary data. *Remote Sens.*, **9**, 1100, https://doi.org/10.3390/rs9111100.