Attributing Terrestrial Water Storage Variations across China to Changes in Groundwater and Human Water Use

MEIXIA LV,^a ZHUGUO MA,^{a,b} AND NAIMING YUAN^a

^a *Key Laboratory of Regional Climate–Environment for Temperate East Asia, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China*

^b *University of Chinese Academy of Sciences, Beijing, China*

(Manuscript received 11 April 2020, in final form 18 October 2020)

ABSTRACT: This study investigated the attribution of terrestrial water storage (TWS) variations across China to changes in groundwater and human water use. As one vital storage component, the groundwater storage (GWS) derived from the Jet Propulsion Laboratory's GRACE (Gravity Recovery and Climate Experiment) mass concentration solution compared reasonably well with the in situ groundwater table depth, with the correlation coefficients ranging from -0.83 to -0.18 , all of which were statistically significant at the 95% confidence level. About 71% of the trends in derived GWS had the same sign as those of observations, without systematic deviation, across China. The GWS variation contributed a large portion of the TWS trend in most regions of China, and the majority of contribution values reached 50%–150% in the Hai River basin, the Loess Plateau, and the middle portion of the Yangtze River basin. The dominant role of GWS is closely related to the detected long-term “memories” in both TWS and GWS. The increase of irrigation consumption accelerated the TWS depletion trend by 13.4% in the Huai River basin, while the decrease of consumptive agricultural water use alleviated the TWS decline rate by 4.1% in the Hai River basin. Importantly, the correlation coefficients reached 0.74–0.95 between the TWS change and the residual of precipitation, evapotranspiration, flow into the sea, and irrigation consumption in the four river basins of particular interest. The findings of this study are helpful for understanding regional water cycles in China.

KEYWORDS: Precipitation; Water masses/storage; Anthropogenic effects; Evapotranspiration; Hydrometeorology; Remote sensing

1. Introduction

Under the influences of climate change and intensive anthropogenic activities, the terrestrial water cycle has changed markedly, especially in dry regions with large populations (Ahmed et al. 2014; Feng et al. 2018; Rodell et al. 2018; Scanlon et al. 2018). Fortunately, the terrestrial water storage (TWS) variation products from the Gravity Recovery and Climate Experiment (GRACE) satellites (Tapley et al. 2004), which represent changes related to both climate variability and human impacts, have provided an unprecedented opportunity for studies on the terrestrial water cycle and climate change (Long et al. 2017; Tapley et al. 2019; Thomas and Famiglietti 2019). GRACE-TWS includes all forms of water stored above and beneath the ground. Changes in water storage under the ground, especially in soil moisture, are closely related to flood generation (Blöschl et al. 2019; Yang et al. 2019) and agricultural drought. Groundwater is crucial for water supply in arid and drought-impacted regions where other types of freshwater are deficient.

Comprehensive understanding of changes in TWS is critical for our understanding of the water cycle, as well as for water resource management.

Attribution of the change in TWS to its influencing factors in local areas of China has been addressed in some studies (Lv et al. 2019b; Sun et al. 2020), and it is important to investigate how the changes of different water storage compartments can potentially impact the availability and utilization of water resources (Felfelani et al. 2017). However, spatial, quantitative contributions of water storage components to TWS variations across China, as well as in terms of the influence of human water use, have been less common (Huang et al. 2013; Luo et al. 2016; Lv et al. 2019b; Mo et al. 2016). One likely reason is that the reliability of groundwater storage (GWS) derived from GRACE has not been evaluated well from a comprehensive perspective across China, which is largely due to the not readily available long-term in situ measurements of groundwater and associated specific yield for unconfined aquifers and the storage coefficient for confined aquifers (Feng et al. 2013, 2018; Gong et al. 2018). The directly observed groundwater level changes need to be converted to water storage changes by multiplying specific yields for unconfined aquifers and storage coefficients for confined aquifers, which are generally obtained through pumping tests, and hence their spatial values are very difficult to obtain. For instance, Zhang et al. (2020) evaluated the correlation of annual trends between GRACE-based GWS in China and observations without converting the groundwater level changes to storage changes, and they did not distinguish unconfined and confined groundwater. However, the contribution of groundwater depletion from

Denotes content that is immediately available upon publication as open access.

Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JHM-D-20-0095.s1>.

Corresponding author: Zhuguo Ma, mazg@tea.ac.cn

DOI: 10.1175/JHM-D-20-0095.1

© 2020 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#) (www.ametsoc.org/PUBSReuseLicenses).

deep aquifers has been emphasized by Feng et al. (2013) in north China.

Zhang et al. (2019) calculated the relative contributions of precipitation, evapotranspiration (ET), and runoff for 168 global river basins, but the anthropogenic influences were ignored. Although some studies have discussed human influences on GRACE-TWS (Ahmed et al. 2014; Hosseini-Moghari et al. 2019), little attention has been paid to spatial, quantitative contributions of direct human water-use impacts (Deng and Chen 2017; Huang et al. 2015; Panda and Wahr 2016; Sun et al. 2020). Felfelani et al. (2017) isolated human-induced TWS change by deducting the natural TWS variation simulated by hydrological models from the GRACE-TWS change (i.e., the trend in GRACE-TWS minus that in natural TWS). However, there was no analysis of the spatial pattern of human-induced TWS in their study, and different kinds of human activities were not distinguished. Moreover, Yi et al. (2016) came up with a linear relationship between precipitation and TWS, i.e., $m = a(P - P_0) + b$, where m is the annual TWS change, P_0 is the average precipitation over a long-term period, and the fitted parameter b is used to represent the anthropogenic contribution, which was used to investigate the anthropogenic contribution as a whole and indirectly. Although some previous GRACE-based studies have agreed that GWS depletion associated with intensive irrigation was responsible for the TWS decline in the North China Plain (NCP) and surroundings (Ebead et al. 2017; Feng et al. 2013; Moiwio et al. 2013; Rodell et al. 2018), the contribution of irrigation to TWS has been less well quantified. Therefore, the direct effect of human water use on water storage needs to be further investigated.

Given the above background, the primary objective of this study was to investigate the attribution of GRACE-TWS variations across China to changes in groundwater and human water use. First, the spatiotemporal variations in the monthly TWS anomaly (TWSA) were examined across China from both spherical-harmonic and mass concentration (mascon) solutions. Second, GWS was obtained from GRACE-TWSA in combination with the Global Land Data Assimilation System (GLDAS). The GWS results were comprehensively evaluated with respect to observed groundwater from both unconfined and confined aquifers. Third, the relative contributions of water storage components to the TWSA trend were estimated with consideration of irrigation consumption. Finally, the variations of TWS change (hereafter referred to as TWSC, which is the change in TWS between two study time steps, and was estimated using GRACE-TWSA data) were investigated according to the water balance equation.

2. Data and methods

a. Study area

China, a land area of approximately 9.6 million km², has multiple climate types, including humid and hot climate (cold and dry climate) in the southeast (northwest). As shown in Fig. 1a, there are 10 water resource regions: the Yellow River (795 000 km²), Yangtze River (1 800 000 km²), Hai River (318 000 km²), Huai River (330 000 km²), Songhua River (935 000 km²), Liao River (314 000 km²), Northwest rivers

FIG. 1. Study area of China, including (a) 10 water resource regions, and (b) the gross domestic product reported in Huang et al. (2014).

(3 387 000 km²), Southwest rivers (847 000 km²), Pearl River (578 000 km²), and Southeast rivers (241 000 km²). The Huai River region (numbered 4 in the figure) is composed of the Huai River basin and Shandong Peninsula. The sediment deposits from the Yellow, Huai, and Hai Rivers formed the heavily populated NCP which provides fertile soils for agriculture (Kang and Eltahir 2018). There is also a narrow definition of the NCP, which refers to the plain area between the Taihang Mountains in the northwest and the Yellow River in the southeast, i.e., the area of the NCP within the Hai River basin (Feng et al. 2018). Hereinafter, the narrow definition of the NCP is referred to as “the narrow NCP.” Importantly, the NCP contributes a large portion of the gross domestic product (GDP) of China (Fig. 1b), followed by the Yangtze River and Yellow River basins. Consequently, this study paid particular attention to the following four river basins because of their important roles in the socioeconomic development and availability of census water-use data: the Yellow River, Yangtze River, Hai River, and Huai River basins.

TABLE 1. Overview of the datasets used in this study.

Variable	Source	Scale	Time span
Precipitation	CMA, http://data.cma.cn/en	Monthly, China, 0.5°	1961–present
Evapotranspiration (ET)	GLDAS v2.1 (Noah 3.3)	Monthly, global, 0.25°	2000–present
	GLDAS v2.1 (Noah 3.3)	Monthly, global, 0.25°	2000–present
	GLEAM v3.3a	Monthly, global, 0.25°	1980–2018
	GLEAM v3.3b	Monthly, global, 0.25°	2003–18
Flow into the sea; census human water-use data (irrigation, manufacturing, and domestic)	Water Resource Bulletin	Annual, basin integrated	1998–present
Terrestrial water storage anomaly (TWSA) from GRACE	Spherical-harmonic solutions from CSR, JPL, and CFZ	Monthly, global, 1°	April 2002–March 2020
	Mass concentration (mascon) solution from JPL	Monthly, global, 0.5°	April 2002–March 2020
	Mass concentration (mascon) solution from CSR	Monthly, global, 0.25°	April 2002–May 2020
Aggregated total land water storage anomaly from GLDAS	https://podaac.jpl.nasa.gov/dataset/TELLUS_GLDAS-NOAH-3.3_TWS-ANOMALY_MONTHLY	Monthly, global, 1°	April 2002–January 2020
Observed groundwater table depth	Geological Environmental Monitoring of Groundwater Level Yearbooks of China	Monthly, China	2005–16
Reconstructed sectoral human water-use data	Huang et al. (2018)	Monthly, global, 0.5°	1971–2010

b. Precipitation, ET, and flow into the sea

The precipitation data employed in this study were from the 0.5° gridded monthly dataset provided by the China Meteorological Administration (CMA; available at <http://data.cma.cn/en>; Table 1). The data were interpolated from 2472 gauge stations across China with records that have been through strict quality control procedures to ensure data accuracy. Furthermore, another precipitation dataset was collected from GLDAS, version 2.1 (GLDAS v2.1). Three ET datasets were collected from the Global Land Evaporation Amsterdam Model (GLEAM) and GLDAS v2.1 (Table 1), in which the two GLEAM v3.3 datasets differed in their forcing and temporal coverage. The Noah ET of GLDAS v2.1 was found to be close to the basin-averaged water balance ET estimate over the Yellow River basin, which was computed as the residual of observation-based precipitation, gauged streamflow, and GRACE-TWSC (Lv et al. 2019a). Moreover, the linear trend of GLDAS v2.1 Noah ET (4.1 mm yr^{-1}) was close to that of the ET dataset of Jung et al. (2010) (3.9 mm yr^{-1}), which was based on observations from a global network of eddy covariance towers and a machine-learning algorithm, over the Yellow River basin (Lv et al. 2019b). In addition, the annual flow data for the flow into the sea of each basin were collected from the Water Resource Bulletins provided by each River Conservancy Commission (available at <http://www.yrcc.gov.cn/other/hhgb/>, <http://www.cjw.gov.cn/zwzc/bmgb/>, <http://www.hwcc.gov.cn/hwcc/wwgj/xxgb/szygb/>, and <http://www.hrc.gov.cn/main/szygb/index.jhtml>).

c. TWSA and observed groundwater table depth

Five TWSA datasets derived from GRACE and GRACE Follow On (GRACE-FO) were collected, which included two mascon solutions and three spherical-harmonic solutions

(available at <http://grace.jpl.nasa.gov/>; Table 1). The mascon solutions were from the Jet Propulsion Laboratory (JPL) and the Center for Space Research (CSR), which are denoted as JPL-M and CSR-M, respectively. The spherical-harmonic solutions were from JPL, CSR, and the GeoForschungsZentrum (GFZ), which are denoted as JPL-SH, CSR-SH, and GFZ-SH, respectively. Furthermore, aggregated total land water storage anomalies (sum of soil moisture, snow, and canopy water) derived from the GLDAS Noah model were also collected from https://podaac.jpl.nasa.gov/dataset/TELLUS_GLDAS-NOAH-3.3_TWS-ANOMALY_MONTHLY (Rodell et al. 2020). The monthly GLDAS water storage anomalies were computed over the same days during each month as the GRACE and GRACE-FO data. Additionally, in situ measurements of groundwater table depth were from the Geological Environmental Monitoring of Groundwater Level Yearbooks of China, compiled by the China Institute of Geological Environmental Monitoring. We selected 442 monitoring wells across China according to the integrity of data throughout the entire period of January 2005–December 2016 in order to obtain credible linear trends, but only 409 wells with a statistically significant change in trend were used in the evaluation of the GRACE-based GWS (Fig. 2). As the measurement is water table depth, a positive trend value represents a decreasing trend of groundwater. Among the 409 wells, there were 213 wells for unconfined aquifer and 196 for confined aquifer. There are many observation wells for confined aquifers in the NCP because the groundwater use was also from confined aquifers in the region (Cao et al. 2013; Feng et al. 2018).

d. Reconstructed consumptive human water use by sector

The dataset of sectoral human water use reconstructed by Huang et al. (2018) was employed in this study. It is a global

FIG. 2. Locations of monitoring wells and trends of observed groundwater table depth from both (a),(b) unconfined and (c),(d) confined aquifers at the monthly time scale after removal of the seasonal cycle in 2005–16.

0.5° gridded monthly dataset for the period 1971–2010 and was based on historical records reported by many agencies or organizations, outputs of models, and spatial/temporal statistical downscaling algorithms. The spatial downscaling was from

country (or state) level to grid scale, and the temporal downscaling was from a 5-yr interval to a monthly scale. For the irrigation sector, the simulated irrigation results by global hydrological models (GHMs) were corrected by correction

factors according to historical records. The dataset distinguishes the following six water-use sectors: irrigation, manufacturing, domestic, electricity generation, livestock, and mining. More details on these water-use categories can be found in [Huang et al. \(2018\)](#). Both withdrawal and consumption values are provided in this dataset, but we were only concerned with the consumptive results.

Before using them in our study, the reconstructed consumptive water-use data were first evaluated against census data also collected from the Water Resource Bulletins of China provided by each River Conservancy Commission. These census data are annual and basin-integrated values (i.e., each river basin is treated as one unit), which include withdrawn/consumed water-use data for agriculture, manufacturing, domestic activities, and ecological environments. Each river conservancy commission is responsible for the census data of the basin where the river is located. The evaluations were conducted for the Yellow River, Yangtze River, Hai River, and Huai River basins, due to the availability of the census data. As mentioned earlier, the reconstructed monthly irrigation water data of [Huang et al. \(2018\)](#) were obtained from reported historical records and simulated gridded irrigation water by four state-of-the-art GHMs: WaterGAP ([Döll and Siebert 2002](#)), LPJmL ([Rost et al. 2008](#)), H08 ([Hanasaki et al. 2008](#)), and PCR-GLOBWB ([van Beek et al. 2011](#)). Therefore, there were four sets of reconstructed usage data for irrigation. Since the focus of our study was on the linear trend in water use, the best reconstructed irrigation results derived from the above four GHMs were selected separately for the above four river basins in terms of the trend simulation performance. We focused on the trend of monthly water use to be in line with the trend analysis of monthly TWSA, and hence estimated the contribution of water-use consumption to the change in TWSA based on trend values. On the other hand, magnitudes of human water use generally cannot be accurately captured in numerical models, but trends can usually be reproduced. Nevertheless, the interannual variability of irrigation consumption is analyzed when assessing the quality of the water-use data of [Huang et al. \(2018\)](#) in this paper.

e. Methods

1) ESTIMATION AND EVALUATION OF GWS

In this study, gridded GWS was estimated as the difference between the GRACE-TWSA and aforementioned land water storage anomaly from GLDAS, which is a commonly used method to isolate GWS from GRACE-TWSA ([Feng et al. 2013](#); [Panda and Wahr 2016](#); [Zhang et al. 2020](#)). Five GWS results were obtained from two mascon and three spherical-harmonic solutions of GRACE-TWSA, and then their accuracies were evaluated with respect to in situ measurements of groundwater. The GWS assessments were performed in two ways. First, the time series of GRACE-based GWS was evaluated using directly monitored groundwater table depth data at the basin scale. The usage of directly observed groundwater level data does not involve the influence of extra uncertainty that may be induced by the converting process from groundwater level changes to storage changes. To be comparable, the estimated gridded GWS data were interpolated to the observation sites, i.e., 70, 79, 77, and 67

monitoring wells for the Hai River, Huai River, Yellow River, and Yangtze River basins, respectively. Then, correlation coefficients were calculated between the GRACE-based GWS and observed groundwater table depth at the basin scale, for which the correlation coefficients would be negative.

Second, the estimated GWS was evaluated further in terms of describing the trend at the pixel scale after converting the trends in observed groundwater level to water storage and interpolating the observations from monitoring wells to the GWS grids. Specifically, the trends in groundwater level from observation wells were multiplied by specific yields for unconfined aquifers and storage coefficients for confined aquifers, to be comparable to the trends of estimated GWS. The values of specific yield and storage coefficient were estimated for each monitoring well in this study. The specific yield values were obtained according to the trilinear graph of textural classification from [Johnson \(1967\)](#) together with soil texture data of China (percentage values of silt, clay, and sand) with 1-km spatial resolution, which was provided by the Data Center for Resources and Environmental Sciences, Chinese Academy of Sciences (<http://www.resdc.cn>). The storage coefficient for a certain confined aquifer was calculated through multiplying the aquifer thickness by its specific storage value, which was gained from the lookup table of [Zhang and Shen \(1998\)](#) and the information from [Yu \(2008\)](#) according to the aquifer characteristics reported in the Geological Environmental Monitoring of Groundwater Level Yearbooks of China. The uncertainty in the estimated storage coefficient for confined aquifers is generally higher than that of specific yield for unconfined aquifers. Therefore, the sensitivity of the GWS assessment to the storage coefficient value was further examined. In addition, whether a long-term “memory” exists in GWS was detected by detrended fluctuation analysis (DFA; [Yuan et al. 2010](#)). A long-term memory exists when the line's slope is larger than 0.5 in the log-log plot of the DFA curve; otherwise, only white noise exists ([Yuan et al. 2010](#)).

2) EOF DECOMPOSITION AND CONTRIBUTION ESTIMATION

Empirical orthogonal function (EOF) decomposition is usually used to decompose time series of variable fields into two components: the spatial function component that does not change over time and the temporal function component representing only temporal variations. Given the successive missing months in the transition period from GRACE to GRACE-FO (i.e., July 2017–May 2018), the trend in derived GWS of the period April 2002–January 2020 was estimated in terms of EOF decomposition. As TWS is composed of surface water and subsurface water, TWS was split into the aforementioned water storage from GLDAS (without groundwater component) and groundwater (represented by the estimated GWS) in this study, due to data availability. Furthermore, the spatial, quantitative contributions of surface water (represented by the water storage from GLDAS), groundwater, and irrigation consumption to the change rate of TWSA were estimated based on linear trend values. For instance, the contribution of GWS to the TWSA change rate was calculated as the ratio of the GWS trend to the TWSA trend. The uncertainties in the trends of TWSA and GWS were estimated as the standard deviation among multiple

FIG. 3. (a)–(e) Trends of terrestrial water storage anomaly after removal of the seasonal signal, and (f) uncertainty in the five trends in April 2002–June 2017, wherein the shaded areas represent statistical significance at the 95% confidence level of the Student's t test. Panels (a) and (b) show results of mascon solutions from CSR and JPL, and (c)–(e) show results of spherical-harmonic solutions from CSR, JPL, and GFZ.

trend results, which is a frequently used method for uncertainty estimation (Lv et al. 2019a).

3. Variations of TWS and GWS across China

Rodell et al. (2018) deemed that gap filling was necessary when computing the linear trend of monthly TWSA after

removal of the seasonal signal. The successive missing months of July 2017–May 2018 cannot be filled by linear interpolation, and the gap filling work is beyond the scope of this study. Therefore, the linear trend of monthly TWSA was computed from April 2002 to June 2017 after removal of the seasonal cycle in our study. Due to the data availability, the evaluation of GWS derived from GRACE was performed during 2005–16

at the monthly time scale, on the basis of the in situ measurements of both unconfined and confined groundwater. The trend analysis of gridded GWS was conducted from April 2002 to January 2020 using EOF decomposition.

a. Variations of TWSA from both mascon and spherical-harmonic solutions

As shown in Fig. 3, the TWSA decreased in the Hai River basin, Huai River basin, Yellow River basin (except the headwaters), Liao River basin, and the north part of the Northwest rivers region. The TWSA increased in the majority of the Yangtze River basin, the Pearl River basin, Southeast rivers region, and northern Tibetan Plateau. Importantly, almost all the changes were statistically significant at the 95% confidence level of the Student's *t* test in April 2002–June 2017. The spatial patterns of the trend were similar among the five TWSA datasets, but the trend magnitudes of the three spherical-harmonic solutions were generally smaller than those of JPL-M and CSR-M. According to the estimated uncertainty (Fig. 3f), large discrepancies in the trend were mainly concentrated in the southern Hai River basin, the downstream area of the Yellow River basin, and western China. However, as stated by Scanlon et al. (2018), the signal-to-noise ratio has been increased and uncertainties decreased owing to advances in GRACE processing from spherical-harmonics to more recent mascon solutions (Watkins et al. 2015). As a result, the trends from JPL-M and CSR-M are supposed to be more reliable.

The above TWSA variation is detrimental for the water supply in the narrow NCP, which already suffers from a severe water shortage (Cao et al. 2013; Long et al. 2020). As the NCP contributes a large portion of China's GDP, TWS depletion would likely bring about adverse effects on socioeconomic development. It should be noted that the increase in TWSA also likely plays a negative role in flood control in the region located in the southeast of the Yangtze River. That is because saturation excess (Dunne) runoff, which occurs after the soil water content reaches field capacity, dominates in humid regions. It has been reported that water storage changes under the ground, especially in soil moisture, are closely related to the generation of floods (Blöschl et al. 2019; Yang et al. 2019). Generally, saturation excess runoff is more likely to generate along with the increment of water storage under the ground, namely, a wetting underlying surface.

b. Evaluation of GWS derived from GRACE in 2005–16

1) GWS EVALUATION USING OBSERVED

GROUNDWATER TABLE DEPTH AT THE BASIN SCALE

Regarding the five derived gridded GWS datasets, the time series of GWS were first evaluated by using the observed groundwater table depth data at the basin scale. Given the well-known high spatial heterogeneity in groundwater, the gridded GWS data were interpolated to the observation stations to be comparable. The results showed that the correlation coefficients were all statistically significant at the 95% level between the observation and the GWS derived from JPL-M over the four river basins of particular interest (Fig. 4). Specifically, the correlation coefficient was -0.834 over the Hai River basin. For the GWS estimated from CSR-M, CSR-SH, JPL-SH, and GFZ-SH, the correlations were not as good as that of JPL-M (Figs. S1–S4

FIG. 4. Correlation coefficients (*r*) between the estimated groundwater storage from JPL-M and the observed groundwater table depth after standardized processing over river basins from 2005 to 2016 ($p < 0.05$ indicates significance at the 95% confidence level of the Student's *t* test).

in the online supplemental material). For instance, the correlation coefficients over the Huai River basin ranged from -0.250 to -0.065 , the absolute values of which were much smaller than that of -0.614 for the GWS derived from JPL-M. However, the GWS performance from CSR-M, for which the correlations were all significant at the 90% level, was overall better than those from spherical-harmonic solutions. Furthermore, the groundwater variation of the Hai River basin was generally well captured by all the five GWS results, which is likely due to the evident depletion signal of groundwater that can be relatively easily detected by GRACE. In summary, the GWS derived from JPL-M performs the best, followed by the CSR-M result, over the four river basins, with respect to observed groundwater table depth.

2) GWS EVALUATION USING THE OBSERVED GROUNDWATER TREND AT THE PIXEL SCALE

The GWS trends derived from the mascon solutions are shown in Figs. 5a and 5b, the absolute values of which were also

FIG. 5. (a),(b) Trends in GRACE-based groundwater storage and (c)–(f) observations from both unconfined and confined aquifers in January 2005–December 2016, with the seasonal cycle removed.

generally larger than those from the spherical-harmonic solutions (Fig. S5). The spatial pattern of uncertainty in the estimated GWS trends was consistent with that in GRACE-TWSA trends because the only difference among the derived GWS results came from the different TWSA datasets employed. The trends in the GRACE-based GWS were then evaluated from the perspective of trend description across China at the pixel scale, after converting the observed groundwater level trend to the trend in water storage based

on specific yield and storage coefficient values. The average of the estimated specific yield over the narrow NCP was 0.087 in our study, which is close to the value of 0.075 reported by [Cao et al. \(2013\)](#) but a little larger than the value of 0.06 stated in [Feng et al. \(2018\)](#). Furthermore, the average of the estimated storage coefficients for confined aquifers was 0.00138 over the narrow NCP, which is close to the value of 0.00125 reported by [Cao et al. \(2013\)](#). Therefore, the estimated specific yield and storage coefficient values of

our study were reasonable. Given the relatively large uncertainty in the estimated storage coefficient, the sensitivity of GWS assessment to the value of the storage coefficient was further investigated in this study, the findings of which are presented later.

The results showed that, although there were both increasing and decreasing trends of groundwater in observations, the numbers of wells with decreased groundwater (140 and 105) were more than those of wells with increased groundwater (73 and 91) for both unconfined and confined aquifers (Figs. 5c–f). Regarding unconfined groundwater, the ranges of the declining trend overall agreed with each other between the GWS derived from JPL-M and observations in northern China, but with relatively large trends at some monitoring wells. This is reasonable because 1) the scales are different in Fig. 5, 2) high spatial heterogeneity exists in groundwater, and 3) there were also increasing trends in unconfined groundwater which would offset some of the depletion trends in the observation. Furthermore, the trends from confined aquifers were considerable at some observation wells (Figs. 5e,f). Feng et al. (2013) also mentioned that the decline in groundwater in deep aquifers was important in the NCP. In terms of the trend, the GWS derived from JPL-M was closer to the observations than that from CSR-M, especially for the Huai River basin. The GWS from JPL-M exhibited significant declining trends throughout almost the entire basin, while the depletion trends in the GWS from CSR-M were not all significant and there were even increasing trends in some pixels. The conclusion is consistent with that from the assessment above using directly observed water table depth. As a result, the GWS derived from JPL-M performed the best in our study. Zhang et al. (2020) also found, on the basis of in situ measurements, that the GWS derived from JPL-M, combined with GLDAS v2.1, outperforms the WaterGAP Global Hydrological Model and earth2Observe in its reflection of the groundwater changes in China. However, they only evaluated the correlation of annual trends without converting the observed groundwater level changes to storage changes, and did not distinguish unconfined and confined groundwater.

Given the problem of spatial mismatch between the gridded GRACE-based GWS and point-scale measurements mentioned above, the observed trends in water storage were interpolated to the GWS grids. As shown in Fig. 6, the observed groundwater mainly decreased in the NCP and increased in the lower reaches of the Yangtze River. Then, the trends in GRACE-based GWS and observations were further compared in terms of sign and magnitude. Regarding the GWS derived from JPL-M, 71% of pixels had the same sign in trend, among which 58% were overestimated and 42% were underestimated in terms of the absolute trend value. For the GWS derived from CSR-M, 63% of pixels had the same sign in trend, among which 52% and 48% were overestimated and underestimated, respectively. That is to say, no obvious systematic bias in the GRACE-based GWS trend was found in the evaluation. To test the sensitivity of the above conclusion to the value of the storage coefficient for confined aquifers, the storage coefficient values were reduced to be

FIG. 6. Trends in observed groundwater storage interpolated based on both unconfined and confined groundwater in January 2005–December 2016. The dots represent the locations of monitoring wells.

10% and increased up to 1000%, from which a similar conclusion of GWS performance was obtained. With the reduced (increased) storage coefficient, 72% (72%) of pixels had the same sign in trend, among which 57% (45%) were overestimated and 43% (55%) were underestimated for the GWS derived from JPL-M. For the GWS based on CSR-M, the three percentage values became 62% (62%), 55% (38%), and 45% (62%), respectively. Consequently, the GWS derived from JPL-M can be further used to investigate the variation in groundwater across China over a longer period than the evaluation. The trend in the GWS from CSR-M was also examined as a side analysis.

c. Variations of GWS across China during April 2002–June 2017 and April 2002–January 2020

Given the successive missing months in the transition period of July 2017–May 2018 from GRACE to GRACE-FO, the spatial pattern of the GWS trend was first examined in April 2002–June 2017 (Figs. 7a,d), and the trends were then used to estimate the contributions of GWS to the TWSA change rates, as reported in the next section. Compared with 2005–16, the GWS still reduced markedly in the southern Hai River basin and the downstream region of the Yellow River basin, but the GWS depletions from both JPL-M and CSR-M were alleviated in the Huai River basin. To investigate the GWS variation over the entire period of GRACE and GRACE-FO, EOF decomposition was carried out for the period April 2002–January 2020 across China. As shown in Figs. 7b, c, e and f, the first EOF mode, together with its associated principal component (PC) 1, represented the apparent long-term trend in GWS. The first mode explained 49.9% and 44.6% of the total variance in the GWS derived from JPL-M and CSR-M, respectively. The shift points in PC 1 occurred in October 2011 for the GWS derived from JPL-M and in May 2012 for the GWS from CSR-M. The results showed that the groundwater depletion continues until now in the Hai River basin.

FIG. 7. (a),(d) Trends of GRACE-based groundwater storage in April 2002–June 2017 and (b),(c),(e),(f) EOF analyses of groundwater storage in April 2002–January 2020, wherein (a)–(c) are results from JPL-M and (d)–(f) are from CSR-M. “Mode 1” represents the EOF first mode, and “PC 1” is its associated principal component.

4. Attribution analyses of variations in TWSA and TWS

a. Attribution of the TWSA trend to changes in water storage components

The contributions of surface-water and groundwater components are shown in Fig. 8, which demonstrates that the GWS variation contributed a large portion of the TWSA trend in most regions of China, from the results of both JPL-M and

CSR-M. The results from JPL-M are paid particular attention in the text below, as its GWS was of higher quality than that from CSR-M. The contribution values of GWS mainly ranged from 50% to 150% in the Hai River basin and the Loess Plateau, as well as the middle part of the Yangtze River basin (Fig. 8a). The contribution value more than 100% means that the trend in GWS exceeded that in TWSA, which generally occurred when the surface water had an opposite trend relative to that of TWSA. The large impact of GWS on TWSA in the

FIG. 8. Contributions of surface water and groundwater to the trend of terrestrial water storage anomaly during the period of April 2002–June 2017 (the yellow star indicates the location of the Three Gorges Dam).

Hai River basin was in line with the conjecture of Rodell et al. (2018) that the TWSA decline mainly resulted from groundwater depletion in the vast agricultural region surrounding Beijing. Felfelani et al. (2017) also found that the sum of soil moisture and groundwater contributed more than 50% of the total variation in TWS in the Yellow River and Yangtze River basins at the basin scale, but their study was based on results of a land surface model. Note that the high contribution value surrounding the Three Gorges Dam in the Yangtze River basin was partially related to the reservoir impoundment. The reservoir started to impound water in June 2003 and was first filled to its near-full capacity in November 2008 (Wang et al. 2011). On the other hand, the water storage from GLDAS is not affected by human activity, and hence the impact of human water management was probably passed to the GRACE-based GWS through the subtraction operation of GRACE-GLDAS.

Furthermore, the difference in contribution was not so wide between the surface water and GWS in the Pearl River basin with a wet climate condition, with most contribution values both falling within the range of 20%–80%. The contributions of GWS were large in the north of Xinjiang Province, with values even higher than 150%, which may be due to the dry climate condition with limited surface water resources. In other words, the increase in surface water could not easily reverse the

decreasing TWSA trend in the region. The situation is a little more complicated in the Tibetan Plateau, largely because of the complex terrain and melting of snow and glaciers. The increase in TWSA over the northern Tibetan Plateau has been reported as mainly due to lake expansion from increased precipitation and glacial melt (Long et al. 2017; Rodell et al. 2018; Yang et al. 2014; Zhang et al. 2011). For the Songhua River region and the border between the Huai River basin and the Yangtze River basin in Fig. 8a, we noted that the negative contributions of GWS basically corresponded to the locations of the Daxing'an Mountain range, Changbai Mountain, and the Dabie Mountain. The cold alpine areas, that have complex hydrologic processes including glaciers, snowpack, and permafrost, should be investigated further in the future.

Importantly, the crucial role played by GWS in explaining the TWSA trend is closely related to long-term memory. As shown in Fig. 9, both the derived GWS and TWS from JPL-M exhibited an apparent long-term memory, as determined by the DFA method. Moreover, the memory signals of GWS (0.831–0.959) were more obvious than those of TWS (0.729–0.786) in the four river basins of particular concern. The conclusion from CSR-M was similar to that from JPL-M (Fig. S6). It should be noted that the long-term memory is a vital reason why we generally cannot obtain a satisfactory result from an

attribution analysis of the TWSA trend from the perspectives of precipitation, ET, and runoff in the same time span. For instance, the trend values of groundwater were much greater than those of precipitation in the study of Li et al. (2019). Besides, the counterpart of precipitation, ET and runoff is TWSC rather than TWSA in terms of the water balance equation, which has been mentioned in Lv et al. (2019b). However, unlike TWSA, there is no significant trend in TWSC in most regions of China, which makes the attribution analysis of TWSC change rate difficult. Instead, the interannual variability of TWSC is investigated in section 4c.

b. Impact of irrigation consumption

As mentioned earlier, the impact of irrigation on TWS has been less well quantified in the NCP. Irrigation water is the largest sector among human water uses, which can be abstracted from both surface-water and groundwater bodies. As reported in this section, the contribution of consumed irrigation water to the TWSA change rate was estimated. The collected reconstructed results of consumptive human water use were first evaluated using the census values. As defined in the Water Resource Bulletin, the census of consumptive water use represents the water volume that cannot return to the surface and subsurface water bodies during the processes of use.

First, all four sets of the reconstructed consumptive irrigation water, i.e., based on H08, WaterGAP, LPJmL, and PCR-GLOBWB GHMs, were evaluated for the four river basins of particular interest (Fig. S7). For the irrigation water simulations among these four GHMs, Huang et al. (2018) reported that the uncertainty arising from model structure (mainly the parameterization and assumptions of the irrigation scheme) was larger than that induced by forcing data. The magnitudes of reconstructed irrigation consumption were generally underestimated, among which the irrigation amount from PCR-GLOBWB was the smallest over the four river basins, likely as a result of the different crop calendars employed. PCR-GLOBWB used the crop calendar data of Portmann et al. (2010), while the other three models employed climate conditions to simulate crop calendars (Huang et al. 2018). Furthermore, no model could reproduce all the irrigation trends well in the four river basins, which is partially related to the fact that irrigation management decision is flexible and also influenced by individual actions of farmers (Lawston et al. 2017; Lei et al. 2014). Overall, it remains difficult at present for actual irrigation practices to be accurately parameterized in numerical models. In this study, we focused on the linear trends rather than magnitudes of the irrigation consumption. In terms of linear trend, the results from WaterGAP, LPJmL, and H08 were chosen for the Yellow River, Yangtze River, and Huai River basins, respectively (Fig. 10). Note that the available census data of the Hai River basin was consumptive agricultural water use, which includes not only agricultural irrigation but also water use for forestry, animal husbandry, and fisheries. The agricultural irrigation consumption accounted for ~95% of the consumptive agricultural water use in the year 1998 (i.e., the only year with census irrigation data). Therefore, the evaluation was conducted over a relatively long time period for the Hai River basin to reduce the possible influence of water

use in forestry, animal husbandry, and fisheries. Combined with the irrigation performances in 2003–10, the average of WaterGAP, LPJmL, and PCR-GLOBWB was used to investigate the irrigation consumption change in the Hai River basin. The trends in the selected optimal irrigation consumptions compared fairly well with that of the census data for the Yellow River, Yangtze River, and Hai River basins. The trend in the H08 result was about half that in the census data in the Huai River basin, which was partially due to the existence of transferred water from the Yellow River and Yangtze River to the Huai River basin. The total amount of transferred water ranged from 13.8 to 30.3 mm yr⁻¹ in 2003–10, with an obvious increasing trend of 1.8 mm yr⁻¹, which cannot be easily taken into account in GHMs (Wada et al. 2014). Furthermore, the interannual variabilities in the census data were largely captured by the selected optimal irrigation results.

The spatial pattern of the trend in irrigation consumption is shown in Fig. 11a, for which the trends were derived from individual optimized reconstructed results from the four GHMs (as in Fig. 10) at the monthly scale for the period 2003–10. The consumptive irrigation exhibited a significant upward trend in most parts of the Yellow River and Huai River basins, as well as the western Hai River basin. There were both positive and negative trends in the Yangtze River basin, but almost only the negative trends concentrated in two areas were significant. Considering that increased irrigation consumption induces decreased water storage, the trend of irrigation consumption was multiplied by -1.0 in the contribution estimation (Figs. 11b–f). Note that the spatial patterns of contribution were basically similar among the five spherical-harmonic and mascon solutions, and the different signs of contribution values in the Huai River basin were due to the different signs of the TWSA trend (Fig. S8). The results showed that the contributions of irrigation change to the TWSA trend were small in the four basins except the Huai River basin. The increase of irrigation consumption accelerated the TWSA decline trend by less than 2% in most irrigated areas of the Yellow River basin, while the increase of irrigation consumption accelerated the TWSA decrease trend by 2%–10% in the northern Huai River basin. The large absolute values of contribution were ignored in the south part of the Huai River basin because of the nonsignificant trends of TWSA (Fig. S8). As noted earlier, the trend in the reconstructed irrigation consumption was underestimated in the Huai River basin, and thus the contribution should be higher than that shown in Fig. 11. Hence, the impacts of irrigation consumption based on the annual census data were also investigated over a longer period (2003–16) at the basin scale (Fig. 12). It was found that the increase of irrigation consumption accelerated the TWSA depletion trend by 4.9% and 13.4% in the Yellow River and Huai River basins, respectively. The decrease of consumptive agricultural water use alleviated the TWSA decline rate by 4.1% in the Hai River basin. The heavily managed Hai River basin has experienced severe water shortages over the past two decades, and policies (e.g., reduction in irrigation pumpage) have been made to restrict groundwater extraction (Long et al. 2020). In the Yangtze River basin, the reduction of consumed irrigation water accelerated the TWSA increase trend by 3.4%. In summary, the

FIG. 9. Long-term memories of (left) terrestrial water storage and (right) groundwater storage derived from JPL-M in April 2002–June 2017. A slope > 0.5 represents a long-term memory.

FIG. 10. Evaluations of reconstructed irrigation consumptions over the (a) Yellow River basin, (b) Yangtze River basin, (c) Hai River basin, and (d) Huai River basin.

contribution of irrigation consumption change to the TWSA trend was generally less than 5%, but it was considerable in the Huai River basin.

As detailed earlier, we found that the TWSA decline in the narrow NCP could largely be attributed to GWS depletion, which was believed to closely link to the intensive groundwater-based irrigation (Feng et al. 2013; Rodell et al. 2018). However, given the common problem of water budget nonclosure, the direct contribution of irrigation consumption to TWSC remains difficult to quantify with large uncertainty (not shown). Besides irrigation, the consumptive water use in the manufacturing, domestic, livestock, mining, and electricity generation sectors were also examined. Regarding manufacturing and domestic water uses, seven of eight trends in the reconstructed results had the same sign as those of the census data over the four basins, but with generally underestimated trend values (not shown). Although manufacturing and mining water uses mainly increased in China with statistically significant trends, the trends were fairly small relative to that of irrigation (Fig. S9). The contributions of the above five water-use sectors were much smaller than that of irrigation to the TWSA trend.

c. Attribution of TWSC variation according to the water balance equation with consideration of irrigation

Given there were almost no significant trends in TWSC across China, the interannual variability of TWSC was investigated at the basin scale in terms of the water balance equation, which was conducted based on observed precipitation and flow into the sea as well as census irrigation consumption.

The correlations between TWSC and the residual of precipitation, ET, flow into the sea, and irrigation consumption were calculated, as shown in Fig. 13. The observed annual data of flow into the sea already included the influences of human water use (e.g., reservoir regulation and water withdrawal), and thus correspond well to GRACE-TWSC compared with model-simulated streamflow. The consideration of irrigation consumption was used to account for the evaporated water from irrigation that cannot be well reflected by currently available ET products. As stated by Huang et al. (2015), “the irrigated water is partly consumed by ET and outflow to the surface water bodies, whereas the other part recharges groundwater.” Although irrigation information is believed to have been incorporated in satellite sensor products (Anderson et al. 2007; Senay et al. 2013), remotely sensed ET products generally tend to underestimate ET (Lei et al. 2014; Lv et al. 2019a). The machine-learning algorithm-based ET of Jung et al. (2009) was found by Lv et al. (2019a) to underestimate the ET over the Yellow River basin. Regarding the ET product of Jung et al. (2011), as well as two remote sensing-based ET datasets from the Moderate Resolution Imaging Spectroradiometer and Zhang et al. (2010), Lei et al. (2014) documented that they cannot correctly reproduce the trend of simulated ET with observed irrigation in the Hai River basin. In short, irrigation impacts have not been well represented in currently available ET products.

Figure 13 shows that the correlation coefficients between TWSC and the residual value ranged from 0.742 to 0.952 and were all statistically significant at the 99% level in the four river basins of particular concern. Although this conclusion is based

FIG. 11. (a) Trends in irrigation consumption and (b)–(f) contributions of irrigation change to the trend of terrestrial water storage anomaly for the four river basins during January 2003–December 2010. A positive contribution value indicates that the trend value of irrigation multiplied by -1.0 has the same sign as that of terrestrial water storage anomaly, considering that increased irrigation consumption induces decreased water storage.

on precipitation from the CMA and ET from GLEAM v3.3a, the interannual variations of different combinations of precipitation minus ET (including precipitation from GLDAS v2.1 and ET from GLEAM v3.3b and GLDAS v2.1) were close to each other (Fig. S10). We noted that the consideration of consumed irrigation water did not remarkably improve the correlations (Fig. S11), which is partially due to the large uncertainty

in ET. The standard deviation among different ET datasets reached $28.02\text{--}79.83\text{ mm yr}^{-1}$ in the four basins, and the standard deviation among different combinations of precipitation minus ET was $24.76\text{--}70.78\text{ mm yr}^{-1}$. Taking the Huai River basin as an example, the standard deviation of 79.83 mm yr^{-1} in ET was comparable to the mean-annual consumed irrigation water of 92.48 mm yr^{-1} . Regarding the uncertainty of GRACE-TWSC,

FIG. 12. Time series of terrestrial water storage anomaly and consumed irrigation water over the four river basins of particular interest in 2003–16. Numbers in parentheses are absolute values of the ratios of the trend in irrigation consumption to the trend in terrestrial water storage anomaly.

the standard deviations were 7.82, 6.73, 12.20, and 14.52 mm yr⁻¹ in the Yellow River, Yangtze River, Hai River, and Huai River basins, respectively. This phenomenon involves the problem of water budget nonclosure among GRACE-TWSC and currently available water budget terms, which has been discussed in [Lv et al. \(2017\)](#). For instance, the attribution of water budget nonclosure errors to ET has been reported to be the greatest among precipitation, ET, runoff, and TWSC, over the globe (45.4%) and in the Yellow River basin (51%) by [Zhang et al. \(2018\)](#). It can be concluded that there are two obstacles for quantitative attributions of TWSC at present. One is that there is almost no significant trend in TWSC, and the other is that water budget closure is hardly obtainable among GRACE-TWSC and currently available water budget terms.

5. Summary and conclusions

In this study, the spatiotemporal variation of TWS and its attribution were investigated across China from the perspectives of water storage components (surface water and GWS) and water

FIG. 13. Correlations (r) between terrestrial water storage change and the residual of precipitation, evapotranspiration, flow into the sea, and irrigation consumption over the four river basins at the annual time scale from 2003 to 2016, wherein precipitation is from the CMA, and evapotranspiration is from GLEAM v3.3a.

fluxes (precipitation, ET, and flow into the sea), with consideration of irrigation consumption. It was found that the dry area became drier in the NCP and the wet area became wetter in the southeast of the Yangtze River, which may bring disadvantages to both the water supply in the NCP (a water-stressed and densely populated region) and flood control in southeastern China.

The GWS data derived from five GRACE-TWSA products were evaluated comprehensively in order to estimate the contributions of surface water and groundwater to the TWSA trend. The GWS derived from JPL-M performs the best, followed by the result from CSR-M, with respect to 409 monitoring wells across China. The time series of JPL-M-based GWS compared reasonably well with the observed groundwater table depth from both unconfined and confined aquifers at the basin scale, as indicated by the correlations ranging from -0.83 to -0.18 that were all statistically significant at the 95% confidence level. A pixel-by-pixel analysis showed that $\sim 71\%$ of the trends in the GWS from JPL-M had the same sign as that of the observations without systematic deviation.

The results showed that the GWS variation was responsible for a large portion of the TWSA trend in most regions of China, and most contribution values reached 50%–150% in the Hai River basin and the Loess Plateau, as well as the middle portion of the Yangtze River basin. The primary role of GWS is closely linked to the detected apparent long-term memory in both TWS and GWS. Furthermore, the TWSA decline in the Hai River basin can be largely attributed to GWS depletion, and meanwhile the decrease of consumptive agricultural water use alleviated the TWSA reduction rate by 4.1% in the basin. The increase of irrigation consumption accelerated the TWSA depletion trend by 4.9% and 13.4% in the Yellow River and Huai River basins, respectively. The direct contribution of irrigation to TWSC remains difficult to estimate at present due to the problem of water budget nonclosure. Importantly, the correlation coefficients were 0.742–0.952 between TWSC and the residual of precipitation, ET, flow into the sea, and irrigation consumption over the four river basins of particular interest. The consideration of irrigation consumption did not remarkably improve the above correlations, which is partially due to the large uncertainty in ET.

The findings of this study are helpful for our understanding of regional water cycles and water resource management in China. The missing TWSA data during the transition period from GRACE to GRACE-FO need to be filled in the future to investigate the TWSA variations over as long a period as possible. In addition, inelastic water storage changes in deep confined aquifers accompanying land subsidence need to be further studied in the future.

Acknowledgments. This study is jointly sponsored by the National Key R&D Program of China (2018YFA0606002; 2016YFA0600404), and National Natural Science Foundation of China (41530532; 41705072). GRACE land data are available at <http://grace.jpl.nasa.gov>, supported by the NASA MEaSUREs Program. The dataset of reconstructed monthly water withdrawals by Huang et al. (2018) are available at <https://doi.org/10.5281/zenodo.1209296>. The 0.5° gridded monthly precipitation dataset provided by the China Meteorological Administration are available at <http://data.cma.cn/en>. The GLEAM data are available at <https://www.gleam.eu/>. The gross domestic product is available at <http://geodoi.ac.cn/WebCn/doi.aspx?Id=125>. We thank the editor and the anonymous reviewers for their instructive comments and suggestions.

REFERENCES

- Ahmed, M., M. Sultan, J. Wahr, and E. Yan, 2014: The use of GRACE data to monitor natural and anthropogenic induced variations in water availability across Africa. *Earth-Sci. Rev.*, **136**, 289–300, <https://doi.org/10.1016/j.earscirev.2014.05.009>.
- Anderson, M. C., J. M. Norman, J. R. Mecikalski, J. A. Otkin, and W. P. Kustas, 2007: A climatological study of evapotranspiration and moisture stress across the continental United States based on thermal remote sensing: 2. Surface moisture climatology. *J. Geophys. Res.*, **112**, D11112, <https://doi.org/10.1029/2006JD007507>.
- Blöschl, G., and Coauthors, 2019: Changing climate both increases and decreases European river floods. *Nature*, **573**, 108–111, <https://doi.org/10.1038/s41586-019-1495-6>.
- Cao, G., C. Zheng, B. R. Scanlon, J. Liu, and W. Li, 2013: Use of flow modeling to assess sustainability of groundwater resources in the North China Plain. *Water Resour. Res.*, **49**, 159–175, <https://doi.org/10.1029/2012WR011899>.
- Deng, H., and Y. Chen, 2017: Influences of recent climate change and human activities on water storage variations in Central Asia. *J. Hydrol.*, **544**, 46–57, <https://doi.org/10.1016/j.jhydrol.2016.11.006>.
- Döll, P., and S. Siebert, 2002: Global modeling of irrigation water requirements. *Water Resour. Res.*, **38**, 1037, <https://doi.org/10.1029/2001WR000355>.
- Ebead, B., M. Ahmed, Z. Niu, and N. Huang, 2017: Quantifying the anthropogenic impact on groundwater resources of North China using Gravity Recovery and Climate Experiment data and land surface models. *J. Appl. Remote Sens.*, **11**, 026029, <https://doi.org/10.1117/1.JRS.11.026029>.
- Felfelani, F., Y. Wada, L. Longuevergne, and Y. N. Pokhrel, 2017: Natural and human-induced terrestrial water storage change: A global analysis using hydrological models and GRACE. *J. Hydrol.*, **553**, 105–118, <https://doi.org/10.1016/j.jhydrol.2017.07.048>.
- Feng, W., M. Zhong, J.-M. Lemoine, R. Biancale, H.-T. Hsu, and J. Xia, 2013: Evaluation of groundwater depletion in North China using the Gravity Recovery and Climate Experiment (GRACE) data and ground-based measurements. *Water Resour. Res.*, **49**, 2110–2118, <https://doi.org/10.1002/wrcr.20192>.
- , C. Shum, M. Zhong, and Y. Pan, 2018: Groundwater storage changes in China from satellite gravity: An overview. *Remote Sens.*, **10**, 674, <https://doi.org/10.3390/rs10050674>.
- Gong, H., and Coauthors, 2018: Long-term groundwater storage changes and land subsidence development in the North China Plain (1971–2015). *Hydrogeol. J.*, **26**, 1417–1427, <https://doi.org/10.1007/s10040-018-1768-4>.
- Hanasaki, N., S. Kanae, T. Oki, K. Masuda, K. Motoya, N. Shirakawa, Y. Shen, and K. Tanaka, 2008: An integrated model for the assessment of global water resources—Part 1: Model description and input meteorological forcing. *Hydrol. Earth Syst. Sci.*, **12**, 1007–1025, <https://doi.org/10.5194/hess-12-1007-2008>.
- Hosseini-Moghari, S. M., 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, Y., M. S. Salama, M. S. Krol, R. van der Velde, A. Y. Hoekstra, Y. Zhou, and Z. Su, 2013: Analysis of long-term terrestrial water storage variations in the Yangtze River basin. *Hydrol. Earth Syst. Sci.*, **17**, 1985–2000, <https://doi.org/10.5194/hess-17-1985-2013>.
- , D. Jiang, and J. Fu, 2014: GDP distribution data set of China on 1 km grids (in Chinese). *Acta Geogr. Sin.*, **69**, 45–48, <https://doi.org/10.3974/geodh.2014.01.07.V1>.
- , M. S. Salama, M. S. Krol, Z. Su, A. Y. Hoekstra, Y. Zeng, and Y. Zhou, 2015: Estimation of human-induced changes in terrestrial water storage through integration of GRACE satellite detection and hydrological modeling: A case study of the Yangtze River basin. *Water Resour. Res.*, **51**, 8494–8516, <https://doi.org/10.1002/2015WR016923>.
- Huang, Z., and Coauthors, 2018: Reconstruction of global gridded monthly sectoral water withdrawals for 1971–2010 and analysis of their spatiotemporal patterns. *Hydrol. Earth Syst. Sci.*, **22**, 2117–2133, <https://doi.org/10.5194/hess-22-2117-2018>.
- Johnson, A. I., 1967: Specific yield—compilation of specific yields for various materials. U.S. Geological Survey Water Supply Paper 1662-D, 80 pp., <https://pubs.usgs.gov/wsp/1662d/report.pdf>.

- Jung, M., M. Reichstein, and A. Bondeau, 2009: Towards global empirical upscaling of FLUXNET eddy covariance observations: Validation of a model tree ensemble approach using a biosphere model. *Biogeosciences*, **6**, 2001–2013, <https://doi.org/10.5194/bg-6-2001-2009>.
- , and Coauthors, 2010: Recent decline in the global land evapotranspiration trend due to limited moisture supply. *Nature*, **467**, 951–954, <https://doi.org/10.1038/nature09396>.
- , and Coauthors, 2011: Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. *J. Geophys. Res.*, **116**, G00J07, <https://doi.org/10.1029/2010JG001566>.
- Kang, S., and E. A. B. Eltahir, 2018: North China Plain threatened by deadly heatwaves due to climate change and irrigation. *Nat. Commun.*, **9**, 2894, <https://doi.org/10.1038/s41467-018-05252-y>.
- Lawston, P. M., J. A. Santanello Jr., T. E. Franz, and M. Rodell, 2017: Assessment of irrigation physics in a land surface modeling framework using non-traditional and human-practice datasets. *Hydrol. Earth Syst. Sci.*, **21**, 2953–2966, <https://doi.org/10.5194/hess-21-2953-2017>.
- Lei, H., D. Yang, H. Yang, Z. Yuan, and H. Lv, 2014: Simulated impacts of irrigation on evapotranspiration in a strongly exploited region: A case study of the Haihe River basin, China. *Hydrol. Processes*, **29**, 2704–2719, <https://doi.org/10.1002/hyp.10402>.
- Li, B., M. Rodell, J. Sheffield, E. Wood, and E. Sutanudjaja, 2019: Long-term, non-anthropogenic groundwater storage changes simulated by three global-scale hydrological models. *Sci. Rep.*, **9**, 10746, <https://doi.org/10.1038/s41598-019-47219-z>.
- Long, D., Y. Pan, J. Zhou, Y. Chen, X. Hou, Y. Hong, B. R. Scanlon, and L. Longuevergne, 2017: Global analysis of spatiotemporal variability in merged total water storage changes using multiple GRACE products and global hydrological models. *Remote Sens. Environ.*, **192**, 198–216, <https://doi.org/10.1016/j.rse.2017.02.011>.
- , and Coauthors, 2020: South-to-North Water Diversion stabilizing Beijing's groundwater levels. *Nat. Commun.*, **11**, 3665, <https://doi.org/10.1038/s41467-020-17428-6>.
- 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>.
- Lv, M., Z. Ma, X. Yuan, M. Lv, M. Li, and Z. Zheng, 2017: Water budget closure based on GRACE measurements and reconstructed evapotranspiration using GLDAS and water use data for two large densely-populated mid-latitude basins. *J. Hydrol.*, **547**, 585–599, <https://doi.org/10.1016/j.jhydrol.2017.02.027>.
- , —, L. Chen, and S. Peng, 2019a: Evapotranspiration reconstruction based on land surface models and observed water budget components while considering irrigation. *J. Hydrometeor.*, **20**, 2163–2183, <https://doi.org/10.1175/JHM-D-19-0090.1>.
- , —, M. Li, and Z. Zheng, 2019b: Quantitative analysis of terrestrial water storage changes under the Grain for Green Program in the Yellow River basin. *J. Geophys. Res. Atmos.*, **124**, 1336–1351, <https://doi.org/10.1029/2018JD029113>.
- Mo, X., J. J. Wu, Q. Wang, and H. Zhou, 2016: Variations in water storage in China over recent decades from GRACE observations and GLDAS. *Nat. Hazards Earth Syst. Sci.*, **16**, 469–482, <https://doi.org/10.5194/nhess-16-469-2016>.
- Moiwo, J. P., F. Tao, and W. Lu, 2013: Analysis of satellite-based and in situ hydro-climatic data depicts water storage depletion in North China Region. *Hydrol. Processes*, **27**, 1011–1020, <https://doi.org/10.1002/hyp.9276>.
- Panda, D. K., and J. Wahr, 2016: Spatiotemporal evolution of water storage changes in India from the updated GRACE-derived gravity records. *Water Resour. Res.*, **52**, 135–149, <https://doi.org/10.1002/2015WR017797>.
- Portmann, F. T., S. Siebert, and P. Döll, 2010: MIRCA2000—Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling. *Global Biogeochem. Cycles*, **24**, GB1011, <https://doi.org/10.1029/2008GB003435>.
- Rodell, M., J. S. Famiglietti, D. N. Wiese, J. T. Reager, H. K. Beaudoin, F. W. Landerer, and M. H. Lo, 2018: Emerging trends in global freshwater availability. *Nature*, **557**, 651–659, <https://doi.org/10.1038/s41586-018-0123-1>.
- , and Coauthors, 2020: Monthly gridded Global Land Data Assimilation System (GLDAS) from NOAA-v3.3 land hydrology model for GRACE and GRACE-FO over nominal months, version 3.3. PO.DAAC, accessed 9 June 2020, <https://doi.org/10.5067/GGDAS-3NH33>.
- Rost, S., D. Gerten, A. Bondeau, W. Lucht, J. Rohwer, and S. Schaphoff, 2008: Agricultural green and blue water consumption and its influence on the global water system. *Water Resour. Res.*, **44**, W09405, <https://doi.org/10.1029/2007WR006331>.
- Scanlon, B. R., and Coauthors, 2018: Global models underestimate large decadal declining and rising water storage trends relative to GRACE satellite data. *Proc. Natl. Acad. Sci. USA*, **115**, E1080–E1089, <https://doi.org/10.1073/pnas.1704665115>.
- Senay, G. B., S. Bohms, R. K. Singh, P. H. Gowda, N. M. Velpuri, H. Alemu, and J. P. Verdin, 2013: Operational evapotranspiration mapping using remote sensing and weather datasets: A new parameterization for the SSEB approach. *J. Amer. Water Resour. Assoc.*, **49**, 577–591, <https://doi.org/10.1111/jawr.12057>.
- Sun, W., Y. Jin, J. Yu, G. Wang, B. Xue, Y. Zhao, Y. Fu, and S. Shrestha, 2020: Integrating satellite observations and human water use data to estimate changes in key components of terrestrial water storage in a semi-arid region of North China. *Sci. Total Environ.*, **698**, 134171, <https://doi.org/10.1016/j.scitotenv.2019.134171>.
- 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>.
- , and Coauthors, 2019: Contributions of GRACE to understanding climate change. *Nat. Climate Change*, **9**, 358–369, <https://doi.org/10.1038/s41558-019-0456-2>.
- Thomas, B. F., and J. S. Famiglietti, 2019: Identifying climate-induced groundwater depletion in GRACE observations. *Sci. Rep.*, **9**, 4124, <https://doi.org/10.1038/s41598-019-40155-y>.
- van Beek, L. P. H., Y. Wada, and M. F. P. Bierkens, 2011: Global monthly water stress: 1. Water balance and water availability. *Water Resour. Res.*, **47**, W07517, <https://doi.org/10.1029/2010WR009791>.
- Wada, Y., D. Wisser, and M. F. P. Bierkens, 2014: Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources. *Earth Syst. Dyn.*, **5**, 15–40, <https://doi.org/10.5194/esd-5-15-2014>.
- Wang, X., C. de Linage, J. Famiglietti, and C. S. Zender, 2011: Gravity Recovery and Climate Experiment (GRACE) detection of water storage changes in the Three Gorges Reservoir of China and comparison with in situ measurements. *Water Resour. Res.*, **47**, W12502, <https://doi.org/10.1029/2011WR010534>.

- Watkins, M. M., D. N. Wiese, D. N. 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>.
- Yang, K., H. Wu, J. Qin, C. Lin, W. Tang, and Y. Chen, 2014: Recent climate changes over the Tibetan Plateau and their impacts on energy and water cycle: A review. *Global Planet. Change*, **112**, 79–91, <https://doi.org/10.1016/j.gloplacha.2013.12.001>.
- Yang, L., L. Wang, X. Li, and J. Gao, 2019: On the flood peak distributions over China. *Hydrol. Earth Syst. Sci.*, **23**, 5133–5149, <https://doi.org/10.5194/hess-23-5133-2019>.
- Yi, S., W. Sun, W. Feng, and J. Chen, 2016: Anthropogenic and climate-driven water depletion in Asia. *Geophys. Res. Lett.*, **43**, 9061–9069, <https://doi.org/10.1002/2016GL069985>.
- Yu, Z., 2008: *Principle and Application of Distributed Hydrology in Watershed* (in Chinese). Science Press, 224 pp.
- Yuan, N., Z. Fu, and J. Mao, 2010: Different scaling behaviors in daily temperature records over China. *Physica A*, **389**, 4087–4095, <https://doi.org/10.1016/j.physa.2010.05.026>.
- Zhang, G., H. Xie, S. Kang, D. Yi, and S. F. Ackley, 2011: Monitoring lake level changes on the Tibetan Plateau using ICESat altimetry data (2003–2009). *Remote Sens. Environ.*, **115**, 1733–1742, <https://doi.org/10.1016/j.rse.2011.03.005>.
- Zhang, J., K. Liu, and M. Wang, 2020: Seasonal and interannual variations in China's groundwater based on GRACE data and multisource hydrological models. *Remote Sens.*, **12**, 845, <https://doi.org/10.3390/rs12050845>.
- Zhang, K., J. S. Kimball, R. R. Nemani, and S. W. Running, 2010: A continuous satellite-derived global record of land surface evapotranspiration from 1983 to 2006. *Water Resour. Res.*, **46**, W09522, <https://doi.org/10.1029/2009WR008800>.
- Zhang, W., and R. Shen, 1998: *Groundwater Hydrology and Groundwater Regulation* (in Chinese). China Water Conservancy and Hydropower Press, 226 pp.
- Zhang, Y., and Coauthors, 2018: A Climate Data Record (CDR) for the global terrestrial water budget: 1984–2010. *Hydrol. Earth Syst. Sci.*, **22**, 241–263, <https://doi.org/10.5194/hess-22-241-2018>.
- Zhang, Z., K. Liu, H. Zhou, H. Lin, D. Li, and X. Peng, 2019: The relative contributions of precipitation, evapotranspiration, and runoff to terrestrial water storage changes across 168 river basins. *J. Hydrol.*, **571**, 1–10, <https://doi.org/10.1016/j.jhydrol.2019.01.041>.