





## Research Article

# Reconstructing Terrestrial Water Storage Anomalies Using Satellite Data to Evaluate Water Resource Shortages from 1980 to 2016 in the Inland Yongding River Basin, China

Kangning Sun <sup>1,2,3</sup> Litang Hu <sup>1,2,3</sup> Xin Liu <sup>1,2,3</sup> and Wenjie Yin <sup>4</sup>

<sup>1</sup>College of Water Sciences, Beijing Normal University, Beijing 100875, China

<sup>2</sup>Beijing Key Laboratory of Urban Hydrological Cycle and Sponge City Technology, Beijing 100875, China

<sup>3</sup>Engineering Research Center of Groundwater Pollution Control and Remediation of Ministry of Education, Beijing 100875, China

<sup>4</sup>Qian Xuesen Laboratory of Space Technology, China Academy of Space Technology, 104 Youyi Road Beijing, 100094, China

Correspondence should be addressed to Litang Hu; [litanghu@bnu.edu.cn](mailto:litanghu@bnu.edu.cn)

Received 27 May 2021; Revised 21 July 2021; Accepted 6 August 2021; Published 17 August 2021

Academic Editor: Andrew H. Manning

Copyright © 2021 Kangning Sun et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Water resources in the Yongding River basin (YRB) are one of the important fundamental conditions in supporting regional water conservation and ecological development. However, the historical changes in water resources under recent human activities remain unknown due to very limited observation data. In this study, terrestrial water storage anomalies (TWSA) as well as multiple precipitation and actual evapotranspiration products from satellites were collected, and the accuracy of the data was verified by observed data or pairwise comparisons. The TWSA during 1980–2016 was reconstructed by using the water balance method, and the reconstructed TWSA was verified using GRACE-observed TWSA, the average depth to groundwater in the Beijing Plain from historical document records and the observed runoff from Guanting Reservoir. The reconstructed TWSA data indicated that the significant decrease occurred during 2000–2016 and the average rate of decreasing trend was  $-11$  mm/year, which may have been caused by a decrease in groundwater storage due to agricultural development. However, the reconstructed TWSA decreased slightly during 1980–1999. The establishment of the water storage deficit index (WSDI) showed that there was no drought or mild drought during 1980–1999; however, the water resource shortage during 2000–2016 was more serious due to groundwater storage decreases caused by agricultural development. The WSDI was verified by using the commonly used self-calibrated Palmer drought severity index. The findings are valuable for sustainable water resource management in the YRB.

## 1. Introduction

Water resources play an important role in human life and ecosystem maintenance worldwide. Terrestrial water storage (TWS) is a critical element of the global and continental water resource cycles, including groundwater storage (GWS), surface water storage (SWS, including lakes, wetlands, reservoirs, and canopy interception), soil moisture storage (SMS), and snow water equivalent (SWE). Therefore, the accurate estimation of TWS is an important issue for understanding the behavior of the hydrological cycle under the influence of human activities. The quantitative study of TWS has been mostly based on hydrological models for the past few decades. Hydrological models often require a large

amount of field observation data to construct. However, field observations are often inadequate or uneven for most regions [1, 2] and are constrained by difficulties related to access, cost, and logistics [3]. Moreover, the construction of the model is time- and money-consuming [4].

Satellite remote sensing methods with wide spatial resolutions and consistent data records provide new ways to measure TWS or hydrological flux and are widely used in areas where observed data are scarce [5, 6]. The Gravity Recovery and Climate Experiment (GRACE), launched in March 2002, can measure the change in the gravity field caused by large variations in a water mass. To date, GRACE provides the first and most unique method to detect TWS anomalies (TWSA) [7], and many studies have demonstrated

the accuracy and effectiveness of TWSA from GRACE [8–11]. GWS anomalies (GWSA) can be derived by assimilating other data products to GRACE-observed TWSA and have also been proven to have high accuracies in evaluating groundwater reserves in the North China Plain [12]. Although GRACE provides an available solution for measuring TWSA, the short lifetime of GRACE means that it cannot provide long-term TWSA variability information. Hence, to obtain the long-term TWS variations, the water balance method is often used [4, 13]. Remote sensing products provide a more convenient method to obtain precipitation, actual evapotranspiration (AET), and runoff when compared with in situ observations. For example, the Tropical Rainfall Measuring Mission product (TRMM) and Moderate Resolution Imaging Spectroradiometer (MODIS) provide datasets independent from ground observations.

The Yongding River basin (YRB) is an important water conservation area and ecological barrier in the Beijing-Tianjin-Hebei region of China. However, water resources are gradually decreasing due to excessive utilization [14]. The average total amount of water resources during 2001–2014 decreased by 21% compared with that during 1956–2010, and the river channel began to dry up in the lower reaches of the YRB after 1996 according to historical document records [15]. The Chinese government issued the *General scheme of comprehensive treatment and ecological restoration of the Yongding River* in 2016. The scheme required that measurements be taken to restore the ecological function of the river as well as water resource utilization [15]. It is very important to evaluate the long-term change in TWS in the YRB and identify the main characteristics to improve water resource management and scheme implementation. However, due to the lack of long-term monitoring data such as groundwater level, the knowledge of change patterns of the water resources in the basin and in the historical period since 1980 is limited to documentary records.

The objective of this paper was to estimate the long-term change in TWSA in the YRB. First, the accuracy of multiple remote sensing products, such as TWSA, precipitation, and AET data, was validated using observed data and pairwise comparisons. Then, the TWSA over the period from 1980 to 2016 were reconstructed by the water balance method using selected precipitation and AET product. The TWSA from GRACE observations, the average depth to groundwater in the Beijing Plain from historical document records, and the runoff from the Guanting Reservoir were used to verify the accuracy of the reconstructed TWSA. Finally, the changes and characterization in the reconstructed TWSA and the water resource shortage characteristics were discussed. The results will improve our understanding of why the water resources in the YRB have changed in recent years.

## 2. Material and Methods

**2.1. Study Area.** The YRB is located in North China Plain. The longitude is  $112^{\circ}$ – $117^{\circ}45'E$ , and the latitude is  $39^{\circ}$ – $41^{\circ}20'N$ . The area is 47 thousand  $\text{km}^2$ . Sanjiadian station is the boundary between the mountain area and the plain area, of which the mountain area is approximately 4.51 thousand

$\text{km}^2$  and the plain area is 1953  $\text{km}^2$ . The altitude of the mountain area is above 1500 m. The YRB (Figure 1) includes two major branches: the northern branch (Yang River) and the southern branch (Sanggan River). The two branches converge and then flow into the Beijing Plain (called the Yongding River). The study area has a temperate continental monsoon climate. The annual average temperature and precipitation are  $6.9^{\circ}\text{C}$  and 360–650 mm, respectively. Over seventy percent of the precipitation is concentrated in the period from June to September.

**2.2. Data Sources.** The data used in this study are listed in Table 1, mainly including precipitation, AET, TWSA, and other hydrological flux data from remote sensing and observed data. The observation wells are shown in Figure 1 and are mainly distributed on the plain. Observed runoff data for the Yongding River at Guanting Reservoir and the average depth to groundwater in the Beijing Plain were collected from historical document records.

**2.2.1. Precipitation Products.** Monthly precipitation datasets ( $0.5^{\circ} \times 0.5^{\circ}$ ) from the China Meteorological Administration (CMA) and TRMM 3B43 ( $0.25^{\circ} \times 0.25^{\circ}$ ) were used to verify their consistency. Then, a suitable set of precipitation products were selected to calculate the reconstructed TWSA. The gridded CMA precipitation dataset was spatially interpolated based on the 2472 gauge stations in China [16], with a time span from 1961 to the present. TRMM 3B43 is a standard monthly precipitation product that has high resolution, high credibility, and good consistency and is widely used in climatological applications [17]. The time span of TRMM 3B43 is from 1998 to the present.

**2.2.2. Evapotranspiration Products.** Four kinds of AET monthly dataset products were selected in this study, including Noah AET ( $1^{\circ} \times 1^{\circ}$ ) data (Noah-AET), MODIS16 global AET data (MODIS-AET, 1 km), Global Land Evaporation Amsterdam Model (GLEAM) AET data (GLEAM-AET,  $0.25^{\circ} \times 0.25^{\circ}$ ), and ERA5-AET data (ERA-AET,  $0.25^{\circ} \times 0.25^{\circ}$ ). The MODIS-AET data were estimated using the Penman-Monteith equation [18–20] with the time span from 2003 to the present. GLEAM is a set of algorithms driven by satellite-based observations that separately estimate daily global AET changes [21, 22] over the time span from 1980 to the present. ERA5 is the 5th generation of global climate reanalysis data released by the European Centre for Medium-Range Weather Forecast with a time span from 1979 to the present.

**2.2.3. GRACE Data.** The monthly  $0.5^{\circ} \times 0.5^{\circ}$  gridded GRACE level-3 mascon (mass concentration) datasets are available from the RL06 time-variable gravity field model, which is provided by the Jet Propulsion Laboratory (JPL). In this study, the monthly TWSA from January 2003 to December 2016 covering 168 months were used. Data of 17 months are not available during the study period. The missing values were filled by the linear difference method that is averaging the values of the two months before and after the month [23]. For the convenience of analysis, the  $0.5^{\circ}$  of GRACE data was averaged to  $1^{\circ}$ .

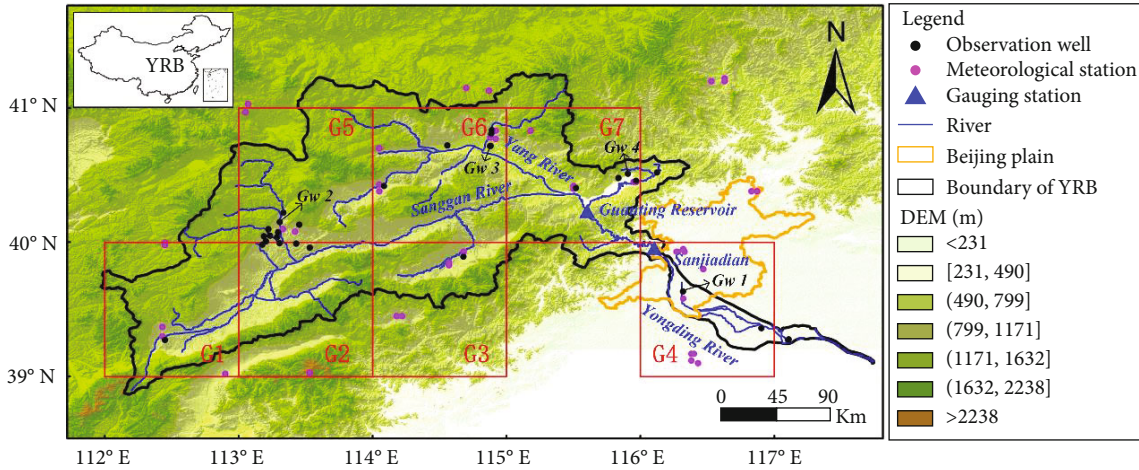


FIGURE 1: Topography of the study area and locations of rivers, gauging stations, meteorological stations, and observation wells.

TABLE 1: List of data sources used in this study.

Category	Data	Spatial and temporal resolution	Time span	Source
Depth to groundwater	Observation wells	Monthly	2005-2014	—
Runoff	Gauging reservoir	Yearly	1980-2016	—
Precipitation	CMA	0.5°, monthly	1961-present	<a href="http://www.nmic.cn/">http://www.nmic.cn/</a>
	TRMM 3B43	0.25°, monthly	1998-present	<a href="http://www.hwcc.gov.cn/">http://www.hwcc.gov.cn/</a>
AET	Noah model v2.1	1°, monthly	2000-present	<a href="https://disc.sci.gsfc.nasa.gov/">https://disc.sci.gsfc.nasa.gov/</a>
	MODIS 16	1 km, monthly	2003-present	<a href="http://files.nts.gov.umt.edu/data/MOD16_TP/">http://files.nts.gov.umt.edu/data/MOD16_TP/</a>
	GLEAM v3.1a	0.25, monthly	1980-present	<a href="https://www.gleam.eu/">https://www.gleam.eu/</a>
SMS, SWS, and SWE	ERA5	0.25, monthly	1979-present	<a href="https://www.ecmwf.int/en/forecasts/datasets">https://www.ecmwf.int/en/forecasts/datasets</a>
	Noah model v2.1	1°, monthly	2000-present	<a href="https://disc.sci.gsfc.nasa.gov/">https://disc.sci.gsfc.nasa.gov/</a>
TWSA from GRACE	JPL-mascon	0.5°, monthly	2002-2017	<a href="https://podaac-opendap.jpl.nasa.gov/opendap/allData/tellus/L3/mascon/RL06/JPL/v02/CRI/netcdf/">https://podaac-opendap.jpl.nasa.gov/opendap/allData/tellus/L3/mascon/RL06/JPL/v02/CRI/netcdf/</a>
Land cover from MODIS	MCD12C1	500 m, yearly	2001-present	<a href="https://lpdaac.usgs.gov/products/mcd12q1v006/">https://lpdaac.usgs.gov/products/mcd12q1v006/</a>
scPDSI	CRU TS 4.04	0.5°, monthly	1901-2019	<a href="https://crudata.uea.ac.uk/cru/data/drought/#global">https://crudata.uea.ac.uk/cru/data/drought/#global</a>

2.2.4. *GLDAS Products.* As a global, high-resolution, offline terrestrial modeling system, the global land data assimilation system (GLDAS) incorporates satellite and ground-based observations to produce optimal fields of land surface states and fluxes in near-real-time. It includes a series of land surface states (e.g., soil moisture) and fluxes (e.g., AET) [24, 25]. Currently, GLDAS drives four land surface models: Mosaic, Noah, CLM, and VIC, and is widely used [26, 27]. In this study, the monthly AET (Noah-AET), SWS (including surface runoff and canopy interception), SWE, and SMS (within 2 m from the surface) from the Noah model in

GLDAS-2.1 with a spatial resolution of  $1^\circ \times 1^\circ$  and time span from 2000 to present were selected.

2.2.5. *Land Cover Products.* The MODIS land cover type product (MCD12Q1) provides global maps of land cover at a yearly scale with a spatial resolution of 500 m for 2001-present. This product extracts the type of land cover based on the Terra and Aqua data after processing. MCD12Q1 contains 13 scientific datasets, and LC\_Type1 (annual IGBP classification) was selected for use in this study. Cropland and urban and built-up land were abstracted from LC\_

Type1. Cropland means that at least 60% of the area is cultivated cropland. Urban and built-up land means there is at least 30% impervious surface area, including building materials, asphalt, and vehicles.

**2.2.6. scPDSI.** The self-calibrated Palmer drought severity index (scPDSI) is compared with the water storage deficit index (WSDI). The scPDSI was developed based on the original Palmer drought severity index [28], which is a drought index based on the relationship between water supply and demand. When the local water supply is less than the demand, it represents a drought; otherwise, it is humid. The scPDSI not only considers the current water condition but also considers the water condition and duration at an earlier time, which can effectively characterize the severity and duration of water resource storage [29, 30]. The global scPDSI grid data were directly used in this paper with a spatial resolution of  $0.5^\circ \times 0.5^\circ$ . The time span of scPDSI is from 1901 to 2019. A lower value indicates a more serious drought. The scPDSI categories are listed in Table 2.

### 2.3. Methods

**2.3.1. Water Balance Method.** The change in water resource storage can be written as Equation (1), which reflects that the water balance results in accumulated precipitation, evapotranspiration, surface runoff, and subsurface runoff within a given area [31]:

$$P - ET - R = \frac{dTWSA}{dt} = \frac{TWSA(t+1) - TWSA(t)}{t}, \quad (1)$$

where  $dTWSA/dt$  represents the change rate of TWSA,  $P$  is the monthly precipitation,  $ET$  is the evapotranspiration, and  $R$  represents the surface runoff and subsurface runoff and is always assumed to be negligible due to their small values [32].

According to the principle of water balance, GWSA can be isolated from TWSA by Equation (2). SMS, SWS, and SWE can be estimated from GLDAS Noah 2.1. TWSA time-series data obtained from GRACE are the monthly difference between the actual TWS and the average TWS of January 2004 to December 2009. To be consistent with TWSA data, the SMS, SWS, and SWE data were processed based on similar operations to obtain SMS anomalies (SMSA), SWS anomalies (SWSA), and SWE anomalies (SWEA). The GWSA calculated by Equation (2) is actually the difference between the actual values and the average value of January 2004 to December 2009.

$$GWSA = TWSA - (SMSA + SWSA + SWEA). \quad (2)$$

**2.3.2. Water Storage Deficit Estimation.** We used WSDI to estimate the situation of water resource shortages. The water storage deficit (WSD) is represented as the residuals by subtracting the climatology from the reconstructed TWSA time series [33]. The climatology was calculated by averaging the TWSA of each month of the reconstructed record (e.g., averaging the values of each January in the whole data record), which is shown in

$$WSD_{i,j} = TWSA_{i,j} - TWSA_j, \quad (3)$$

where  $TWSA_{i,j}$  is the reconstructed time series of TWSA for month  $j$  in year  $i$ .  $TWSA_j$  is the climatology of reconstructed TWSA. WSD is normalized to the WSDI using the zero-mean normalization method as follows:

$$WSDI = \frac{WSD - \mu}{\sigma}, \quad (4)$$

where  $\mu$  is the mean WSD and  $\sigma$  is the standard deviation of the WSD. The lower the value is, the more serious the water resource shortage is [34]. The WSDI categories are listed in Table 3.

**2.3.3. Mann-Kendall Trend Analysis.** The change trend was analyzed by the Mann-Kendall (MK) trend test [36, 37]. The MK trend test is applied to effectively distinguish whether the change of time-series data is in natural fluctuation or has a certain obvious trend. If there is a certain obvious trend, the change rate of this trend can be calculated by the MK method. For a set of time series  $X = (x_1, x_2, \dots, x_n)$ , the  $S$  statistic is calculated as

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{sgn}(x_j - x_i) \text{sgn}(x_j - x_i) = \begin{cases} 1(x_j - x_i) > 0 \\ 0(x_j - x_i) = 0 \\ -1(x_j - x_i) < 0 \end{cases}, \quad (5)$$

where  $n$  is the number of data and  $x_j$  and  $x_i$  are the corresponding values of times  $j$  and  $i$  ( $j > i$ ). Variance  $V(S)$  is defined as

$$\text{VAR}(S) = \frac{1}{18} \left[ n(n-1)(2n+5) - \sum_{p=1}^q t_p(t_p-1)(2t_p+5) \right]. \quad (6)$$

When  $n > 10$ , the standardized calculation of test statistic  $Z_s$  is shown in

$$Z_s = \begin{cases} \frac{S-1}{\sqrt{\text{VAR}(S)}} & S > 0 \\ 0 & S = 0 \\ \frac{S+1}{\sqrt{\text{VAR}(S)}} & S < 0 \end{cases}. \quad (7)$$

$Z_s > 0$  means that the time-series data show an increasing trend; conversely, the time series data show a decreasing trend. If there is a change trend in the time-series data, the change rate ( $\beta$ ) can be estimated by Equation (8) [38]:

$$\beta = \text{median} \left( \frac{x_i - x_j}{i - j} \right). \quad (8)$$

TABLE 2: scPDSI categories [29].

Index	No drought	Mild drought	Moderate drought	Severe drought	Extreme drought
scPDSI	scPDSI > -1	-1 ≥ scPDSI > -2	-2 ≥ scPDSI > -3	-3 ≥ scPDSI > -4	scPDSI ≤ -4

TABLE 3: WSDI categories [35].

Index	No drought	Mild drought	Moderate drought	Severe drought	Extreme drought
WSDI	WSDI > 0	0 ≥ WSDI > -1	-1 ≥ WSDI > -2	-2 ≥ WSDI > -3	WSDI ≤ -3

**2.3.4. Evaluation Index.** The root mean square error (RMSE) and the correlation coefficient ( $r$ ), which are expressed in Equation (9) and Equation (10), respectively, were used to quantitatively assess the differences in the datasets. The RMSE describes the global discrepancy between the observed and simulated time-series data and is more sensitive to erroneous data. A smaller RMSE denotes a better model performance. The  $r$  value measures the degree of correlation among different datasets, and a higher absolute value of  $r$  indicates a higher degree of correlation [17]. In addition, Nash coefficient (NSE, dimensionless) which is expressed in Equation (11) was used to assess the performance of reconstructed TWSA. The closer the NSE value is to the maximum value of 1, the higher the simulation accuracy is.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (x_i - y_i)^2}{n}}, \quad (9)$$

$$r = \frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2} \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}}, \quad (10)$$

$$\text{NSE} = 1 - \frac{\sum_{i=1}^n (x_i - y_i)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad (11)$$

where  $x_i$  or  $y_i$  is the observed (simulated) value,  $i$  is the serial number,  $n$  is the total number of samples, and  $\bar{x}(\bar{y})$  is the average value over  $n$  for  $x$  ( $y$ ).

The uncertain of the reconstructed TWSA results is mainly caused by propagating of the errors from water budget components and can be calculated according to Gaussian error propagation [39] and the expressed is shown in

$$\sigma_{\text{Reconstructed-TWSA}} = \sqrt{\sigma_P^2 + \sigma_{\text{AET}}^2}, \quad (12)$$

where  $\sigma$  is the standard deviation of each component.

### 3. Results and Discussion

**3.1. Validation and Analysis of Groundwater Storage Data.** The observed depth to groundwater can be compared with the GWSA estimated from the GRACE-observed TWSA by subtracting the SWSA, SWEA, and SMSA. Four observation wells distributed in the plain area of four grids (shown in Figure 1) with better time continuity and good consistency with other adjacent wells were used to represent the depth to groundwater in the grid. Generally, the smaller the GWSA

value is, the larger the value of depth to groundwater is, indicating that the depth to groundwater has a negative correlation with the GWSA. As shown in Figure 2, the negative correlation coefficients between the GWSA and the depth to groundwater were all above 0.6 ( $p < 0.01$ ). The results showed a good negative correlation between the GWSA and the depth to groundwater and provided some confidence in the accuracy of the TWSA from GRACE. Figure 2 also shows that the GWSA decreased from 2003 to 2016. Compared with 2016, the average GWSA decreased by 193 mm, with a decreasing trend of -14 mm/year passing the MK test in seven grid blocks (shown in Figure 1). The value of GWSA is the multiply of changes in groundwater levels with specific yield, which usually ranges from 0.001 to 0.02. Thus, the changing amplitude of GWSA is smaller than those of depth of groundwater.

#### 3.2. Comparison of Different Precipitation and Evapotranspiration Data

**3.2.1. Comparison of Different Precipitation Data.** The time-series data of CMA and TRMM during 2003-2016 were compared. The monthly time-series data (Figure 3(a)) indicated that the CMA and TRMM data were basically consistent in the trend of periodic change, with a high correlation coefficient ( $r = 0.99$ ) and an RMSE of 6.5 mm. The correlation coefficient between the TRMM and CMA data reached 0.93, and the RMSE was 37.3 mm on a yearly scale (Figure 3(b)). The yearly averaged precipitation estimated by TRMM and CMA during 2003-2016 is 481.2 and 453.3 mm, respectively. Both the two datasets show the drought condition in 2005-2007, 2009, 2011, and 2014 and no drought condition in other years. The largest difference between the TRMM and CMA data appeared in 2013 (74.6 mm) which may be mainly caused by overestimating of TRMM in rainy season on the monthly scale. On a seasonal scale (Figure 3(c)), the correlation coefficient and the RMSE between the TRMM and CMA data were 0.998 and 3.2 mm, respectively. Both datasets of precipitation reached a maximum in July, with 108.6 mm in CMA and 113.6 mm in TRMM. Precipitation mainly occurs from June to September, accounting for approximately 72% of the annual precipitation, which is consistent with the CMA observations (73%). These results provide confidence in the CMA data over the study area. Hence, the CMA precipitation data are used for later discussion.

**3.2.2. Comparison of Different Evapotranspiration Products.** The time-series data of MODIS-AET, GLEAM-AET, ERA-

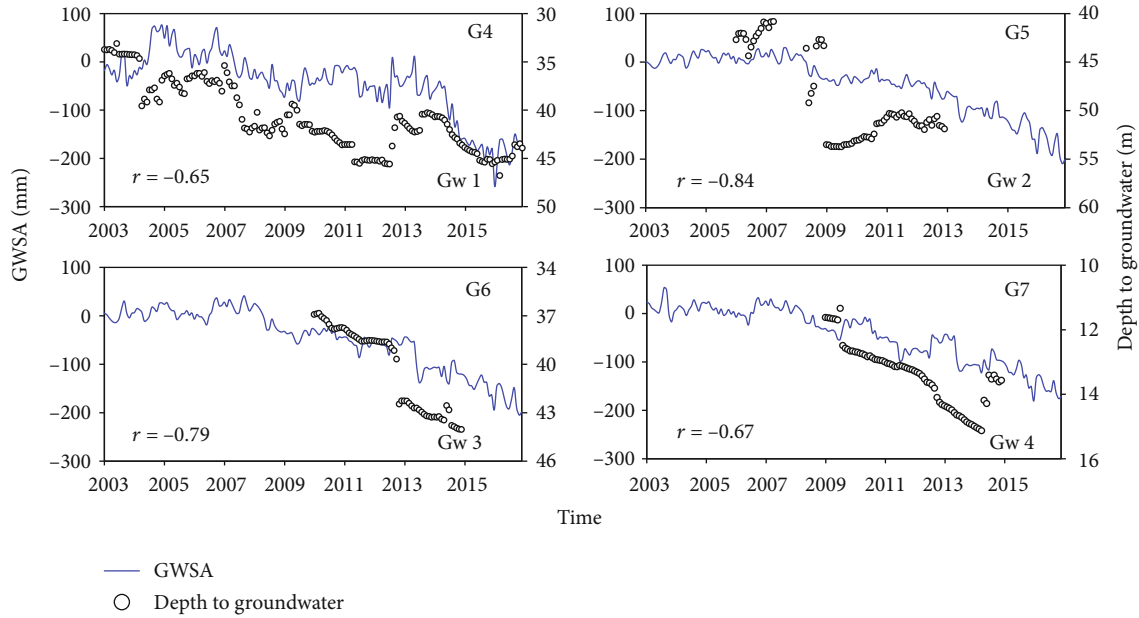


FIGURE 2: Comparison between the GWSA and the depth to groundwater in four grid blocks.

AET, and Noah-AET during 2003–2016 were compared. The correlation coefficients of MODIS-AET, GLEAM-AET, and ERA-AET with Noah-AET were 0.93, 0.94, and 0.96, respectively. The RMSEs of MODIS-AET, GLEAM-AET, and ERA-AET with Noah-AET were 12.4 mm, 14.6 mm, and 10.2 mm, respectively (Figure 4(a)). MODIS-AET, GLEAM-AET, and ERA-AET showed trends and periodic changes similar to those of Noah-AET. On a yearly scale (Figure 4(b)), the average MODIS-AET, GLEAM-AET, ERA-AET, and Noah-AET were 452.8, 389.3, 493.8, and 456.5 mm, respectively. The GLEAM-AET data were much lower, and the reason may be that GLEAM underestimated the peak value of AET (Figure 4(a)). The correlation coefficients of MODIS-AET, GLEAM-AET, and ERA-AET with Noah-AET were 0.46, 0.85, and 0.38, respectively. The GLEAM-AET data had the highest correlation coefficient with the Noah-AET data on the yearly scale. On the seasonal scale (Figure 4(c)), the correlation coefficients of MODIS-AET, GLEAM-AET, and ERA-AET with Noah-AET were higher than 0.96, and the peak value was in July. More than 75% of the seasonal AET occurs from March to September. The peak value of Noah-AET data was higher than that of MODIS-AET, GLEAM-AET, and ERA-AET. In summary, the phase and amplitude from the four datasets were consistent, and the GLEAM-AET data showed the best correlation coefficient ( $>0.8$ ) with the Noah-AET data at the monthly, yearly, or seasonal scales.

**3.3. Reconstruction of Long-Term Terrestrial Water Storage Anomalies.** The TWSA was reconstructed by using the monthly CMA precipitation and MODIS-AET, GLEAM-AET, ERA-AET, or Noah-AET data over the YRB based on the water balance method (Equation (1)) from 2003 to 2016. Among the four reconstructed TWSA from MODIS-AET, GLEAM-AET, ERA-AET, and Noah-AET data, the reconstructed TWSA based on the Noah-AET and CMA

precipitation had the best agreement with the GRACE-observed TWSA, with a correlation coefficient of 0.82 and an RMSE of 58.2 mm. However, the Noah-AET data only start in 2000. Considering that the GLEAM-AET data have a high correlation coefficient with the Noah-AET data from the former discussion, a revised scaling factor was used for the GLEAM-AET data by multiplying a coefficient to match the Noah-AET data for the period from 2003 to 2016, and this value was 1.205 based on the least-squares regression method. The reconstructed TWSA calculated by the CMA and scaled GLEAM-AET data on a yearly scale during 1980–2016 is shown in Figure 5(a), with a correlation coefficient of 0.86, an RMSE of 39 mm, and a NSE of 0.9 when compared with the GRACE-observed TWSA during 2003–2016. The uncertainties in precipitation and scaled GLEAM-AET are 63.4 mm and 48.4 mm, respectively. Then, the uncertainty in reconstructed TWSA is 79.8 mm which is estimated according to Equation (12).

The reconstructed TWSA during 1980–2016 showed a significant period of decreasing trends after 2000 with a decreasing trend of  $-11$  mm/year passing the MK test (Figure 5(a)). The average depth to groundwater in the Beijing Plain from historical document records and the runoff from the Guanting Reservoir approximately represent the changes in groundwater and surface water, respectively. The average depth to groundwater and runoff showed a similar pattern ( $r = -0.76$ ), and both decreased significantly after 2000 (Figure 5(b)). The annual average precipitation during 2000–2016 was 434 mm and was lower than the annual average precipitation during 1980–2016 (440 mm), which contributed to the decrease in TWSA in terms of meteorology. However, another important reason is the decrease in the GWSA. As shown in Figure 6, the GWSA also significantly decreased during this period. The correlation coefficient between the GWSA and the reconstructed TWSA

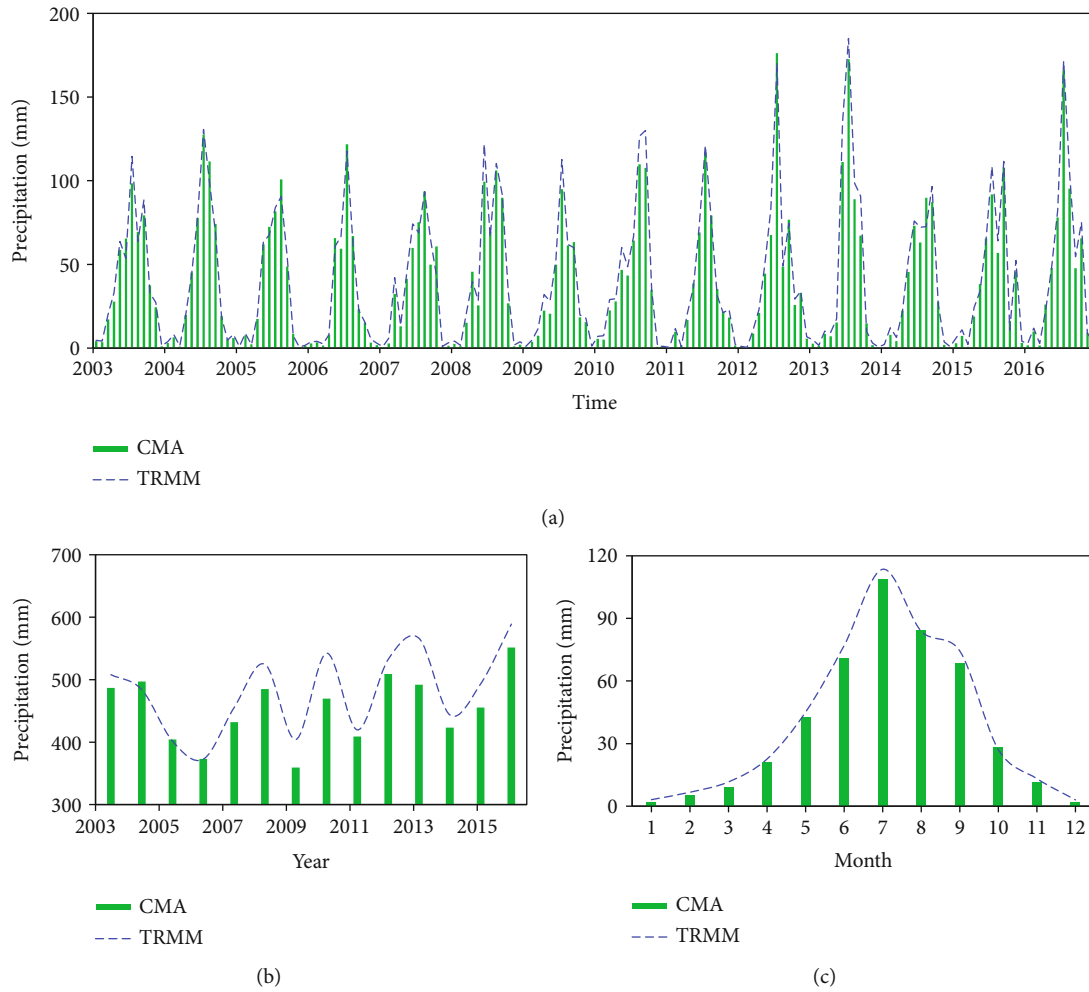


FIGURE 3: Comparison between data from CMA and TRMM precipitation data: (a) monthly time-series data; (b) yearly time-series data; (c) mean seasonal time-series data.

during 2003-2016 was 0.93, and the high relevance demonstrated that the decrease in the TWSA in the study area was mainly caused by the decrease in the GWSA. Rapid economic development occurred in the YRB during 2001-2016. As shown in Figure 6, cropland increased from 6174 km<sup>2</sup> in 2001 to 8476 km<sup>2</sup> in 2016 (an increment of 2301 km<sup>2</sup>), and the agricultural area increased by nearly one-third. Groundwater in this area sustains approximately 70% of irrigation, mainly for wheat production in the dry seasons of winter and spring [40]. The correlation coefficients of cropland area with the GWSA and reconstructed TWSA were -0.87 and -0.78, respectively, which proved that the rapid development of agriculture led to an increase in groundwater consumption and a decrease in GWSA. In addition, urban and built-up land increased from 1364 km<sup>2</sup> in 2001 to 1448 km<sup>2</sup> in 2016 (an increase of 84 km<sup>2</sup>). The development of cities has also led to an increase in water consumption.

It should also be noted that the reconstructed TWSA decreased slightly during the period from 1980 to 1999, with a decreasing trend of -1 mm/year passing the MK test (Figure 5(a)). As shown in Figure 5(b), both the depth to groundwater in the Beijing Plain from historical document records and the runoff from the Guanting Reservoir

decreased slightly during 1980-1999 probably because of the construction of dam and reservoirs in the upper reaches of the basin. The annual average annual precipitation during 1980-1999 was 445 mm, which was slightly higher than that during 1980-2016 (440 mm), and precipitation was sufficient. Both the depth to groundwater and the runoff demonstrated the results from the reconstructed TWSA that although water resources decreased in this period, the decreasing trend was very small.

Figure 5(a) also shows that the change in reconstructed TWSA had a time-lag effect on the change in precipitation. For example, the precipitation reached a peak in 1995 (601 mm); however, the reconstructed TWSA peaked in the next year, and the runoff and depth to groundwater also showed the same time-lag effect. The time-lag effect of TWSA changes in response to precipitation was also clearly observed in 1993. The time-lag effect may be caused by the fact that it usually takes much time for the infiltration of precipitation and rivers to reach the groundwater through the relatively thick vadose zone.

3.4. *Water Shortage of Water Resources.* The changes in the WSDI calculated by the reconstructed TWSA are shown in

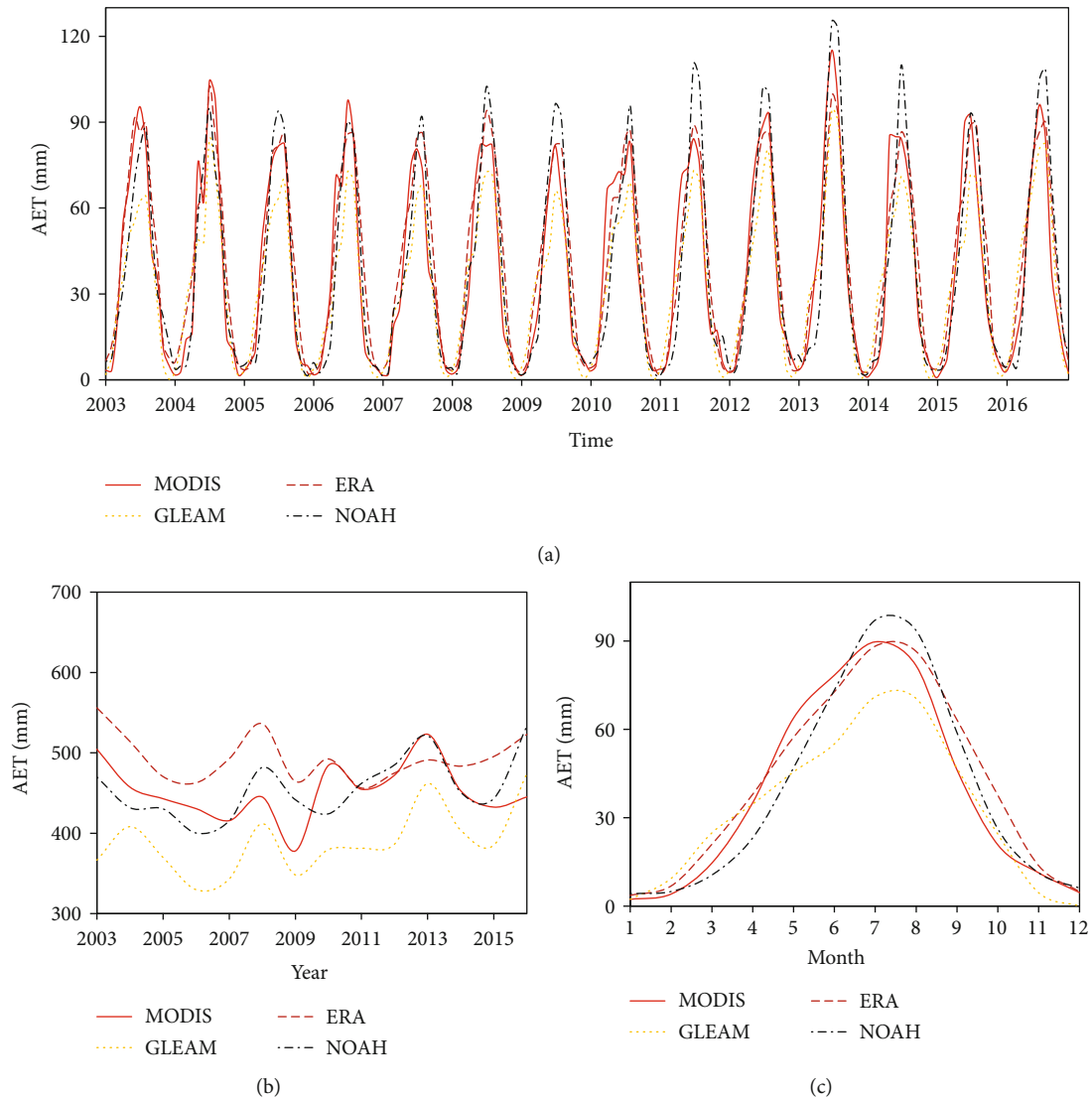


FIGURE 4: Comparison among four AET datasets from MODIS, GLEAM, ERA, and Noah: (a) monthly time-series data; (b) yearly time-series data; (c) mean seasonal time-series data.

Figure 7, and the scPDSI was used to verify the accuracy of the WSDI over the YRB during 1980-2016. From 1980-1999, the water resource shortage estimated from the WSDI was almost no drought or mild drought. The average WSDI was 0.34, which indicated a no drought situation in terms of the water resources during this period. The drought in 1994 was the most serious, which may have been caused by the scarce precipitation in the previous year (Figure 5(a)). The water resource shortage from the scPDSI also indicated no drought or mild drought and occurred at low values in 1994 during 1980~1999. The average scPDSI was -1.08, which indicated a mild drought situation in water storage and was one level different from the WSDI. Although moderate drought years were observed from the scPDSI (1981, 1984, and 1993), the value of the scPDSI was small, and the drought situation was not serious. The scPDSI proves that the WSDI estimation was very accurate during this period. Water resource shortages were not serious, and pre-

cipitation was sufficient (analyzed in Section 3.3) during this period. The supply and demand of water resources were relatively balanced.

From 2000 to 2016, the water resource shortage estimated by the WSDI changed from no drought to severe drought. Differences of approximately one or two levels were present between the WSDI and the scPDSI in this period. In particular, in 2002 and 2007, the difference between the scPDSI and the WSDI was three levels, where the WSDI result was no drought but the scPDSI result was severe drought. The reason may be that the scPDSI was heavily affected by current and previous precipitation. The annual precipitation in 1999-2002 and 2005-2007 was continuously less than the annual average precipitation (Figure 5(a)), which may lead to the estimation results of serious drought conditions. However, both the WSDI and the scPDSI showed more serious water resource storage than that in 1980-1999. As discussed in Section 3.3, the TWS (especially GWS)



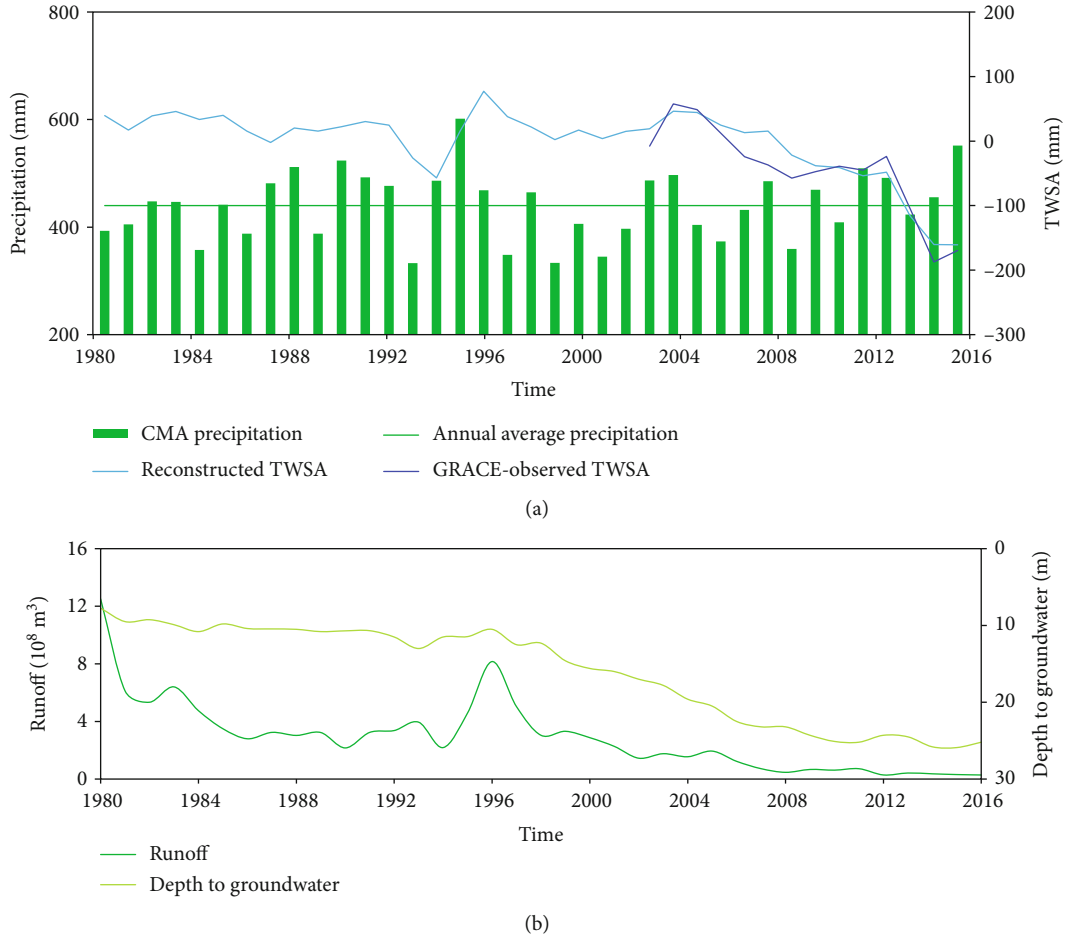


FIGURE 5: (a) The yearly time-series data of CMA precipitation, annual average precipitation, reconstructed TWSA, and GRACE-observed TWSA; (b) the yearly time-series data of average depth to groundwater in the Beijing Plain from historical document records and runoff from the Guanting Reservoir.

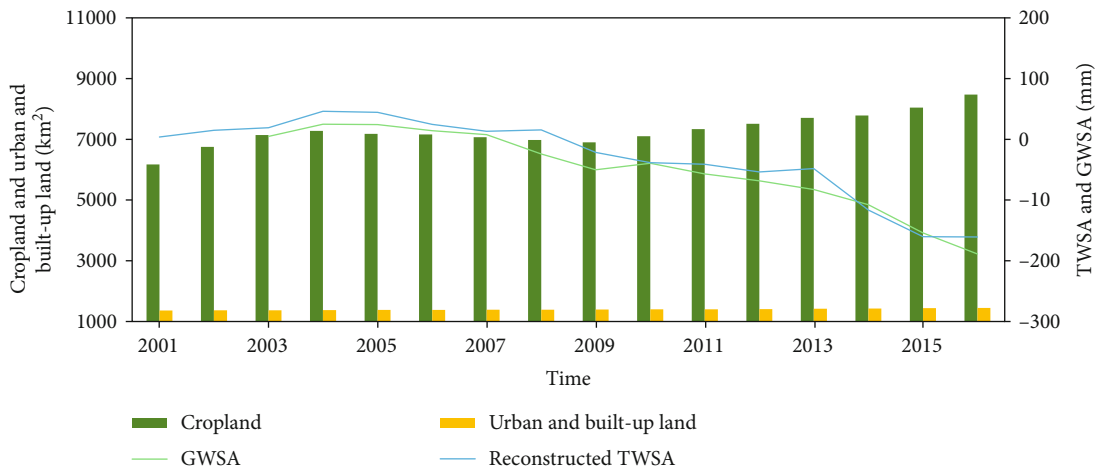


FIGURE 6: The yearly time-series data of cropland, urban and built-up land, GWSA, and reconstructed TWSA.

significantly decreased during this period, which contributed to the drought conditions. That is, human activity (agricultural development) played an important role in the serious water resource shortages during this period. The reason for

the differences between the WSDI and the scPDSI may be that the WSDI (calculated by reconstructed TWSA) and scPDSI are sensitive to evapotranspiration, and the components considered by the WSDI and scPDSI are not exactly

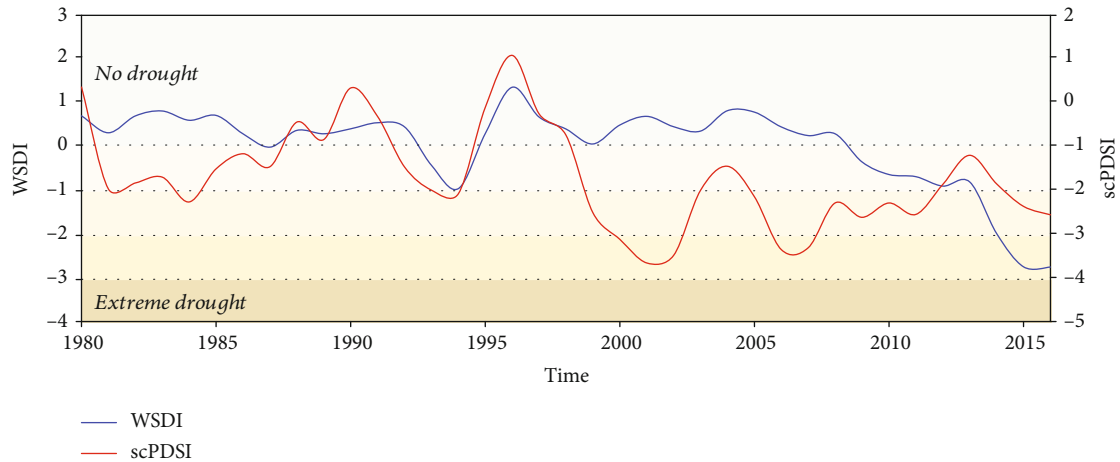


FIGURE 7: Comparisons of monthly time-series data of the WSDI and scPDSI.

the same. For example, the components of groundwater and surface water (e.g., lakes) are included in the WSDI but not in the scPDSI.

#### 4. Conclusions

Due to the excessive use of water resources for the purpose of economic development, the groundwater level has dropped significantly in recent years, and available water resources are decreasing in the YRB. A better understanding of past water resource utilization can aid in generating a rational utilization strategy for the sustainable development of water resources. The TWSA of the YRB from 1980 to 2016 were first reconstructed by using different remote sensing data. Then, the varieties of reconstructed TWSA and water resource shortage characteristics were analyzed. The main conclusions are as follows:

- (1) The change in the GRACE-observed GWSA over the period from 2003 to 2016 matched well with the observed depth to groundwater data ( $r < -0.6$ ). The correlation coefficient and RMSE of precipitation from TRMM and CMA sources were 0.99 and 6.5 mm, respectively, which demonstrated that there was little difference between the two datasets at both monthly and yearly time scales. The AET from MODIS, GLEAM, and ERA had a high correlation with Noah on the monthly scale ( $r > 0.9$ ) or seasonal scale ( $r > 0.8$ ). The GLEAM-AET data showed the best correlation coefficient ( $r > 0.8$ ) with the Noah-AET data at monthly, yearly, or seasonal scales
- (2) The water balance method using precipitation from CMA and scaled AET data from GLEAM was used to obtain the accurate reconstructed TWSA for the period from 1980 to 2016. The reconstructed TWSA matched well with the GRACE-observed TWSA, with correlation coefficient and RMSE of 0.86 and 39 mm, respectively, during 2003-2016. The trend of the reconstructed TWSA during 1980-2016 also agreed with the average depth to groundwater in

the Beijing Plain from historical document records and with the runoff from the Guanting Reservoir

- (3) An obvious decreasing pattern of the reconstructed TWSA was found during 2000-2016, and the average rate of decreasing trend was  $-11$  mm/year. The decreasing trend was mainly caused by a decrease in the GWSA due to the rapid development of agriculture. Before 2000, the reconstructed TWSA decreased slightly with an MK decreasing trend of  $-1$  mm/year
- (4) The water resource shortage was between no drought and mild drought during 1980-1999. Compared with 1980-1999, the water resource shortage was more serious during 2000-2016, which was mainly caused by human activities, especially agricultural development. The WSDI showed an increasingly serious trend from no drought to severe drought

A convenient method was developed to reconstruct long-term changes in TWSA in a region. Due to the difficulties in verifying the accuracy of AET data, the GRACE-observed TWSA was used to check and find the relationship between the reconstructed TWSA and precipitation-AET from the CMA and GLEAM datasets. Combinations of different precipitation and AET datasets have certain uncertainties and may affect the reconstructed results as well as the accuracy of GRACE data. The accuracy of the reconstructed TWSA is difficult to validate because of limited point-scale data and detailed water use data in this study. However, the reconstructed TWSA are at the regional scale and can provide a temporal picture of water resource changes.

#### Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

#### Conflicts of Interest

All authors declare that no conflicts of interest exist.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant Nos. 41877173 and 41831283) and the Beijing Municipal Science and Technology Project (Grant No: Z191100006919001). The authors are thankful for all of the support from for this basic research data.

## References

- [1] J. Liu, A. J. B. Zehnder, and H. Yang, "Global consumptive water use for crop production: the importance of green water and virtual water," *Water Resources Research*, vol. 45, no. 5, 2009.
- [2] J. Liu and H. Yang, "Spatially explicit assessment of global consumptive water uses in cropland: Green and blue water," *Journal of Hydrology*, vol. 384, no. 3-4, pp. 187–197, 2010.
- [3] N. Nie, W. Zhang, Z. Zhang, H. Guo, and N. Ishwaran, "Reconstructed terrestrial water storage change ( $\Delta TWS$ ) from 1948 to 2012 over the Amazon Basin with the latest GRACE and GLDAS products," *Water Resources Management*, vol. 30, no. 1, pp. 279–294, 2016.
- [4] W. Yin, L. Hu, S. C. Han, M. Zhang, and Y. Teng, "Reconstructing terrestrial water Storage variations from 1980 to 2015 in the Beishan area of China," *Geofluids*, vol. 2019, Article ID 3874742, 2019.
- [5] A. K. Sahoo, M. Pan, T. J. Troy, R. K. Vinukollu, J. Sheffield, and E. F. Wood, "Reconciling the global terrestrial water budget using satellite remote sensing," *Remote Sensing of Environment*, vol. 115, no. 8, pp. 1850–1865, 2011.
- [6] E. Forootan, J. Kusche, I. Loth et al., "Multivariate prediction of total water storage changes over West Africa from multi-satellite data," *Surveys in Geophysics*, vol. 35, no. 4, pp. 913–940, 2014.
- [7] M. Rodell, J. S. Famiglietti, J. Chen, S. I. Seneviratne, and P. Viterbo, "Basin scale estimates of evapotranspiration using GRACE and other observations," *Geophysical Research Letters*, vol. 31, no. 20, 2004.
- [8] R. Schmidt, S. Petrovic, A. Güntner, F. Barthelmes, J. Wunsch, and J. Kusche, "Periodic components of water storage changes from GRACE and global hydrology models," *Journal of Geophysical Research: Solid Earth*, vol. 113, no. 8, 2008.
- [9] B. R. Scanlon, Z. Zhang, R. C. Reedy et al., "Hydrologic implications of GRACE satellite data in the Colorado River basin," *Water Resources Research*, vol. 51, no. 12, pp. 9891–9903, 2015.
- [10] T. Yang, C. Wang, Y. Chen, X. Chen, and Z. Yu, "Climate change and water storage variability over an arid endorheic region," *Journal of Hydrology*, vol. 529, no. 1, pp. 330–339, 2015.
- [11] P. Yang, J. Xia, C. Zhan, Y. Qiao, and Y. Wang, "Monitoring the spatio-temporal changes of terrestrial water storage using GRACE data in the Tarim River basin between 2002 and 2015," *Science of the Total Environment*, vol. 595, pp. 218–228, 2017.
- [12] Z. Huang, Y. Pan, H. Gong et al., "Subregional-scale groundwater depletion detected by grace for both shallow and deep aquifers in North China plain," *Geophysical Research Letters*, vol. 42, no. 6, pp. 1791–1799, 2015.
- [13] N. Zeng, J. Yoon, A. Mariotti, and S. Swenson, "Variability of basin-scale terrestrial water storage from a PER water budget method: the Amazon and the Mississippi," *Journal of Climate*, vol. 21, no. 2, pp. 248–265, 2008.
- [14] L. Wang, Z. Wang, T. Koike, Y. Hang, D. Yang, and H. Shan, "The assessment of surface water resources for the semi-arid Yongding River Basin from 1956 to 2000 and the impact of land use change," *Hydrological Processes*, vol. 24, no. 9, pp. 1123–1132, 2010.
- [15] K. Sun, L. Hu, J. Guo, Z. Yang, and S. Zhang, "Enhancing the understanding of hydrological responses induced by ecological water replenishment using improved machine learning models: a case study in Yongding River," *Science of The Total Environment*, vol. 768, no. 7, p. 145489, 2021.
- [16] P. Bai, X. Liu, T. Yang, K. Liang, and C. Liu, "Evaluation of streamflow simulation results of land surface models in GLDAS on the Tibetan plateau," *Journal of Geophysical Research: Atmospheres*, vol. 121, no. 20, pp. 12,180–12,197, 2016.
- [17] Y. Yang and Y. Luo, "Evaluating the performance of remote sensing precipitation products CMORPH, PERSIANN, and TMPA, in the arid region of Northwest China," *Theoretical and Applied Climatology*, vol. 118, no. 3, pp. 429–445, 2014.
- [18] J. L. Monteith, "Evaporation and environment," *Symposia of the Society for Experimental Biology*, vol. 19, pp. 205–234, 1965.
- [19] Q. Mu, F. A. Heinsch, M. Zhao, and S. W. Running, "Development of a global evapotranspiration algorithm based on MODIS and global meteorology data," *Remote Sensing of Environment*, vol. 111, no. 4, pp. 519–536, 2007.
- [20] Q. Mu, M. Zhao, and S. W. Running, "Improvements to a MODIS global terrestrial evapotranspiration algorithm," *Remote Sensing of Environment*, vol. 115, no. 8, pp. 1781–1800, 2011.
- [21] B. Martens, D. G. Miralles, H. Lievens, R. V. D. Schalie, and N. Verhoest, "GLEAM v3: satellite-based land evaporation and root-zone soil moisture," *Geoscientific Model Development Discussions*, vol. 10, no. 5, 2016.
- [22] D. G. Miralles, R. A. M. de Jeu, J. H. Gash, T. R. H. Holmes, and A. J. Dolman, "Magnitude and variability of land evaporation and its components at the global scale," *Hydrology and Earth System Sciences*, vol. 15, no. 3, pp. 967–981, 2011.
- [23] Y. Y. Di Long, Y. Wada, Y. Hong et al., "Deriving scaling factors using a global hydrological model to restore GRACE total water storage changes for China's Yangtze River Basin," *Remote Sensing of Environment*, vol. 168, pp. 177–193, 2015.
- [24] J. Sheffield, G. Goteti, and E. F. Wood, "Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling," *Journal of Climate*, vol. 19, no. 13, pp. 3088–3111, 2006.
- [25] Q. Li, B. Zhong, Z. Luo, and C. Yao, "GRACE-based estimates of water discharge over the Yellow River basin," *Geodesy and Geodynamics*, vol. 7, no. 3, pp. 187–193, 2016.
- [26] V. Sridhar and D. A. Wedin, "Hydrological behaviour of grasslands of the Sandhills of Nebraska: water and energy-balance assessment from measurements, treatments, and modelling," *Ecohydrology*, vol. 2, no. 2, pp. 195–212, 2009.
- [27] V. Sridhar, M. M. Billah, and J. Hildreth, "Coupled Surface and Groundwater Hydrological Modeling in a Changing Climate," *Groundwater*, vol. 56, no. 4, pp. 618–635, 2018.
- [28] W. C. Palmer, *Meteorological Drought*, US Department of Commerce, Weather Bureau, 1965.
- [29] N. Wells, S. Goddard, and M. J. Hayes, "A self-calibrating Palmer drought severity index," *Journal of Climate*, vol. 17, no. 12, pp. 2335–2351, 2004.

- [30] G. van der Schrier, J. Barichivich, K. R. Briffa, and P. D. Jones, "A scPDSI-based global data set of dry and wet spells for 1901–2009," *Journal of Geophysical Research Atmospheres*, vol. 118, no. 10, pp. 4025–4048, 2013.
- [31] O. Lenk, "Satellite based estimates of terrestrial water storage variations in Turkey," *Journal of Geodynamics*, vol. 67, pp. 106–110, 2013.
- [32] S. Swenson and J. Wahr, "Estimating large-scale precipitation minus evapotranspiration from GRACE satellite gravity measurements," *Journal of Hydrometeorology*, vol. 7, no. 2, pp. 252–270, 2006.
- [33] A. C. Thomas, J. T. Reager, J. S. Famiglietti, and M. Rodell, "A GRACE-based water storage deficit approach for hydrological drought characterization," *Geophysical Research Letters*, vol. 41, no. 5, pp. 1537–1545, 2014.
- [34] D. Sinha, T. H. Syed, J. S. Famiglietti, J. T. Reager, and R. C. Thomas, "Characterizing drought in India using GRACE observations of terrestrial water storage deficit," *Journal of Hydrometeorology*, vol. 18, no. 2, pp. 381–396, 2017.
- [35] Z. Sun, X. Zhu, Y. Pan, J. Zhang, and X. Liu, "Drought evaluation using the GRACE terrestrial water storage deficit over the Yangtze River Basin, China," *Science of the Total Environment*, vol. 634, pp. 727–738, 2018.
- [36] H. B. Mann, "Nonparametric tests against trend," *Econometrica*, vol. 13, no. 3, pp. 245–259, 1945.
- [37] M. G. Kendall, *Rank Correlation Measures*, Charles Griffin, 1975.
- [38] R. M. Hirsch, J. R. Slack, and R. A. Smith, "Techniques of trend analysis for monthly water quality data," *Water Resources Research*, vol. 18, no. 1, pp. 107–121, 1982.
- [39] Y. P. Di Long, J. Zhou, Y. Chen et al., "Global analysis of spatiotemporal variability in merged total water storage changes using multiple GRACE products and global hydrological models," *Remote Sensing of Environment*, vol. 192, pp. 198–216, 2017.
- [40] C. Sun and L. Ren, "Assessment of surface water resources and evapotranspiration in the Haihe River basin of China using SWAT model," *Hydrological Processes*, vol. 27, no. 8, pp. 1200–1222, 2013.