remote sensing

MDPI

Article

Downscaling Simulation of Groundwater Storage in the Beijing, Tianjin, and Hebei Regions of China Based on GRACE Data

Jianchong Sun [1,2], Litang Hu [1,2,*], Fei Chen [3], Kangning Sun [1,2], Lili Yu [3] and Xin Liu [4]

[1] College of Water Sciences, Beijing Normal University, Beijing 100875, China
[2] Engineering Research Center of Groundwater Pollution Control and Remediation of Ministry of Education, Beijing Normal University, Beijing 100875, China
[3] General Institute of Water Resources and Hydropower Planning and Design, Ministry of Water Resources, Beijing 100120, China
[4] Powerchina Huadong Engineering Corporation Limited, Hangzhou 311122, China
* Correspondence: litanghu@bnu.edu.cn

Abstract: Gravity Recovery and Climate Experiment (GRACE)-derived groundwater storage anomalies (GWSA) have been used to highlight groundwater depletion in regional aquifer systems worldwide. However, the use of GRACE products in smaller areas is limited owing to the coarse spatial resolution of the data product. This study utilized a dynamic downscaling method to improve the GWSA resolution from 1° to 0.05° by constructing a groundwater storage numerical model in the Beijing, Tianjin, and Hebei regions of China (BTH). The results indicate that: (1) the GRACE-derived and calculated GWSA had a good match with an average root mean squared error (RMSE) of 2.61 cm equivalent water height (EWH) and an average Nash–Sutcliffe efficiency coefficient (NSE) of 0.84 for the calibration period. (2) The hydraulic gradient coefficient and specific yield had the highest sensitivity, and transmissivity had the lowest sensitivity; however, different forcing data had no obvious influence on the GWSA. (3) The downscaled results not only exhibited time series variations that were consistent with those of the GRACE-derived solutions but also revealed a finer spatial heterogeneity of the GWSA along with increasing correlation coefficients between the GRACE-derived GWSA and the in situ measurements of groundwater levels by 0.06 and reducing the RMSE by 8.85%. (4) The downscaled results reflected the spatiotemporal change characteristics of groundwater storage in different hydrogeological units and administrative regions well. This study demonstrates the potential applications of the proposed downscaling method for both regional and local groundwater resource management.

Keywords: groundwater model; GRACE; dynamic downscaling; model calibration; groundwater storage changes

check for updates

Citation: Sun, J.; Hu, L.; Chen, F.; Sun, K.; Yu, L.; Liu, X. Downscaling Simulation of Groundwater Storage in the Beijing, Tianjin, and Hebei Regions of China Based on GRACE Data. *Remote Sens.* **2023**, *15*, 1490. https://doi.org/10.3390/rs15061490

Academic Editors: Robert C. Walter and Timothy D. Bechtel

Received: 5 January 2023
Revised: 1 March 2023
Accepted: 6 March 2023
Published: 8 March 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Groundwater resources are critical for global food and water security [1,2]. In recent decades, excessive groundwater depletion has become a global problem, affecting economic and social development [3,4]. Due to considerable groundwater depletion [5,6], groundwater storage (GWS) changes under the influence of climate change and human activity have garnered increasing attention. Utilizing field and remote sensing observations for the quantitative estimation of the spatiotemporal variations in groundwater storage is critical for the sustainable management of water resources [7].

The Gravity Recovery and Climate Experiment (GRACE) mission and Its successor, GRACE Follow-On (GRACE-FO) have provided a new opportunity to monitor large-scale GWS changes with unprecedented accuracy by vertically disaggregating terrestrial water storage (TWS) change signals since 2002 [8–10]. GRACE gravity satellites effectively compensate for the shortcomings of ground-based measurements, including limited monitoring data coverage, uneven spatial distribution, difficulty in data acquisition, and point-specific

estimates [11,12]. GRACE data have been used worldwide to estimate groundwater storage change, such as in Northwest India [13], the U.S. High Plains and California's Central Valley [14,15], the Colorado River Basin [16], Australia [17,18], the Middle East [19], and North China [11,20–24]. In addition, some studies have proved that the GRACE data and hydrological model are mutually independent but complementary tools, that is, the GRACE-derived data can not only be used to evaluate the groundwater storage change at regional scale but can also be used to calibrate the regional groundwater model and reduce the uncertainty of the model [25]. GRACE data [26] are a significant guide for groundwater modeling, and the combined method composed of GRACE observation data, physical modeling and in situ monitoring data can more effectively manage groundwater. Most researchers applied the hydrological model or groundwater flow model [27–29] to verify the reliability of GRACE-derived groundwater storage changes. In order to reduce the uncertainty of regional groundwater storage changes estimation and improve the model reliability, GRACE data are also used for parameter calibration and data assimilation [30]. In recent years, with the development of remote sensing technologies, the integration of GRACE data and other multisource datasets has become a trend in water resource analysis [31–34]. Although remarkable achievements have been made for large areas, the application of GRACE observations is extremely limited in local areas because of the coarse spatial resolution [35]. Therefore, it is necessary to downscale the GRACE data to a higher resolution that could be practical for refined groundwater resource management at local scales.

Statistical and dynamical downscaling are the two most commonly used methods in downscaling applications [36]. Statistical approaches, such as multivariate regression, artificial neural networks, and machine learning, rely on various statistical techniques to approximate the stochastic between large-scale data and data from small-scale observations, which are comparatively cheap and computationally efficient [37]. Ning et al. [38] constructed an empirical regression model based on the water balance equation to improve the spatial resolution of the TWS. Yin et al. [39] used evapotranspiration data with a relatively high spatial resolution to downscale GRACE-derived groundwater storage anomalies (GWSA). Sun [40] downscaled GRACE-derived TWS by developing an artificial neural network model in the United States. Although statistical downscaling has performed well in improving GRACE-derived data resolution, some disadvantages still need to be considered: the rationale behind statistical downscaling models is not yet well understood; these models do not fully account for some fundamental hydrological processes (e.g., antecedent water storage) and the principle of water balance [41]; and they are not applicable to areas where the correlation between large-scale fields and local elements is not obvious [39].

In contrast to statistical downscaling, dynamic downscaling is used to integrate target datasets into numerical models to produce responses based on physically consistent processes using data assimilation methods. Based on specific physical relationships, dynamic downscaling uses global models to drive regional models, which can be applied at smaller scales to produce finer-resolution information without being limited by observations or the degree of correlation between the large-scale and local elements [36]. Simultaneously, the relatively complete physical process and parameterization scheme of the regional model also guarantees the accuracy of the simulation results. For example, Eicker et al. [42] assimilated GRACE-derived TWS into the WaterGAP Global Hydrology Model using a Kalman filter method in the Mississippi River Basin with reasonable results. Zhong et al. [43] downscaled the GRACE-derived TWS using an iterative adjustment method based on the self-calibration variance-component model and improved the TWS uncertainties through a combination of GRACE observations and the TWS simulated using the Land Surface Model. In addition, many studies [44–46] have demonstrated that dynamic downscaling methods can downscale the resolution of GRACE products to a model scale. These methods were based mostly on hydrological or land surface models, which regarded the groundwater in a grid as a linear reservoir and thus did not consider the influences of the aquifer parameters.

The North China Plain (NCP) is one of the regions where the relationship between water resources, food, energy, and ecology is the most significant and the joint risk is the most prominent [21]. Many scholars have studied the spatiotemporal changes of GWS in the NCP using such techniques as groundwater flow modelling [47], data analysis [11,20–22], combined hydrological models [24,48], and evaluations using monitoring wells [49,50]. Although these studies have explained the spatiotemporal changes of GWS over the region, a detailed evaluation using the GRACE and ground-observation data from municipalities such as Cangzhou, Xingtai, Handan, and Beijing is still not sufficient for the refined and cross-scale management of groundwater resources. Therefore, downscaling research must be conducted to provide high-resolution water storage estimates in the subregions. To the best of our knowledge, there has been a limited focus on groundwater flow models that apply freely available GRACE-derived GWSA data to conduct dynamic downscaling research.

The objective of this study is to construct a groundwater storage model with the freely available GRACE data to downscale the GWSA from 1° to 0.05°. The only difference between the proposed groundwater storage model and the traditional groundwater flow model is that the groundwater storage changes, instead of the hydraulic head, are used as the key variables in the equation, providing a clear physical meaning. In the next section, the study area information and data sources are prepared. In Section 3, the model downscaling method as well as model evaluation methods are described. In Section 4, the model calibration and validation, presentation, and verification of downscaled results are discussed. In Section 5, the downscaling results are used to evaluate the spatiotemporal changes of GWS and compared to the groundwater level (GWL) changes in different subregions; in addition, model parameters and uncertainty are also discussed. The results can be used as guidelines for enhanced groundwater resource management.

2. Materials

2.1. Study Area

The regions of concern in this study are the Beijing, Tianjin, and Hebei (BTH) Regions in China, which includes most of the NCP (Figure 1). The BTH region is the political, cultural, and economic center, extending from 36°03′N to 42°40′N and 113°27′E to 119°50′E, including the capital city Beijing, Tianjin municipality, and 11 city-level administrative entities in Hebei Province, with an area of approximately 218,000 km^2 [51]. The BTH region is within a littoral and semi-arid climatic zone; nearly 70% of the rainfall is concentrated from June to September. The average annual precipitation of the entire area is approximately 400–600 mm but less than 400 mm in the northwest plateau and hilly areas. The mean annual evapotranspiration of the region ranges from 1100 to 2000 mm.

2.2. Geology and Hydrogeology

The BTH contains a variety of geographical elements, from the Yanshan-Taihang Mountains in the northwest to the southeast, and gradually transitions into the NCP and the coastal plain, with basins and hills interspersed in the middle, exhibiting a terrain decrease from the southeast to the northwest (Figure 1). The plateau is mainly distributed in the northwest region. The BTH Plain, which is backed by mountains and faces the sea, is part of the NCP. According to the relative positions and causes, from west to east, the region can be divided into three principal zones: the piedmont plains (PP), central plains, and coastal plains. The piedmont and central plains have slow water circulation. Some studies have merged the central and coastal plains into the East Central Plain (ECP) [22,50]. The mountainous areas in the western and northwestern areas of the study area are identified as WM in this study.

Quaternary loose sediments with a thickness ranging from 150 to 600 m, are widely distributed throughout the plain. The aquifer structure transits from a single aquifer in the upper part of piedmont in the western fan to a multi-layer aquifer in the east, and the lithology transits from gravel and pebble to sand, silt, and clay. The PP is the

primary infiltration and replenishment area for the regional groundwater which has strong permeability. The central plain has a poor ability to receive infiltration from atmospheric precipitation due to contain a thick continuous impermeable layer. The coastal plain has poor aquifer permeability, water storage, and water supply capabilities. According to the burial and hydraulic properties, the groundwater can be divided into four aquifers from top to bottom. The first aquifer is unconfined groundwater, which is mainly represented in the PP with a thickness of approximately 15 to 55 m; the second to fourth layers are confined aquifers, which predominantly comprise the ECP [22]. Groundwater recharge mainly originates from precipitation infiltration, lateral flow in the mountains, and return flow from irrigation, whereas groundwater discharge primarily occurs through groundwater exploitation and phreatic evaporation [52].

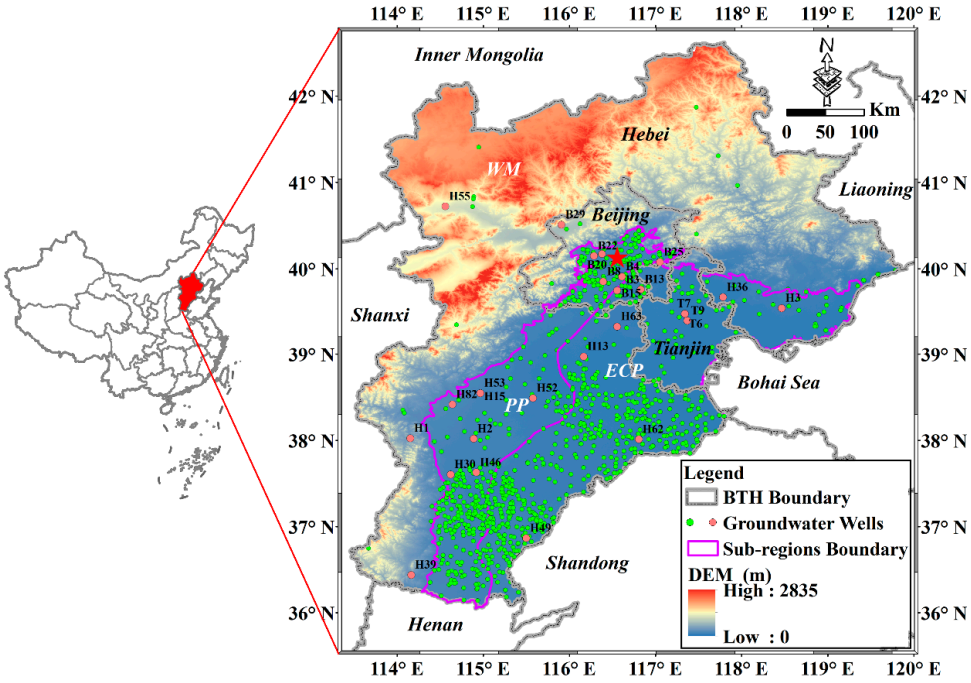


Figure 1. Location map of the BTH. The green and orange dots represent the locations of groundwater observation wells.

The discrepancy between water supply and demand in the BTH region is prominent, and the per capita volume of water resources is less than one ninth of the national average [51]. The extensive shortage of water resources has led to a serious problem of groundwater overexploitation, with the agricultural sector as the main driver in the BTH region. Overexploitation of groundwater is concentrated in the BTH plains, which has resulted in serious ecological deterioration. The South-to-North Water Diversion project was initiated to alleviate this water stress and has been supplying water to the BTH since 2014 [53]. The Action Program for Comprehensive Treatment of Groundwater Overexploitation in the NCP was issued by the China State Council in 2019 and has been adopted in the northern part of the NCP for the comprehensive treatment of groundwater overexploitation.

2.3. Data Sources

We have summarized the data used in this study (Table 1), and detailed descriptions are provided below.

(1) In situ GWL data

Monthly in situ GWL measurements from 986 monitoring wells were collected from the China Geological Environment Monitoring Institute and Ministry of Water Resources, including shallow (~10 to 55 m below the surface) and deep aquifers (~60 m to several hundred meters below the surface). Different observations have different starting times, and

the periods of available data varied from station to station. All the raw well observations were tested for quality during the preprocessing stage. Data with consecutive gaps of more than 2 years were not used. Following this strict quality control, 143 monitoring well stations were selected from the China Geological Environment Monitoring Institute between 2005 to 2014, and 843 monitoring well stations were selected from the Ministry of Water Resources in 2014, 2016, and 2018 to 2020. Note that the GWL measurements were without conversion to groundwater storage due to the uncertainty of the specific yield, which was mainly used for the qualitative validation.

Table 1. Summary of data used in the study.

Type	Product	Spatial (Temporal) Resolution	Period	Source
TWS	GRACE (CSR-RL06M)	0.5° × 0.5° (monthly)	2003–2020	https://grace.jpl.nasa.gov/data/get-data/ (accessed on 11 March 2022).
SMS, SWE	GLDAS	1° × 1° (monthly)	2003–2020	https://disc.gsfc.nasa.gov/datasets (accessed on 11 March 2022).
Precipitation	TRMM 3B43	0.25° × 0.25° (monthly)	2003–2019	https://gpm.nasa.gov/data/sources (accessed on 21 March 2022).
	ERA5	0.25° × 0.25° (monthly)	2003–2020	https://cds.climate.copernicus.eu/cdsapp#!/home (accessed on 23 March 2022).
	CMA	0.5° × 0.5° (monthly)	2003–2020	http://data.cma.cn/ (accessed on 23 March 2022).
	PENG	0.01° × 0.01° (monthly)	2003–2020	http://poles.tpdc.ac.cn/zh-hans/data/faae7605-a0f2-4d18-b28f-5cee413766a2/ (accessed on 2 April 2022).
AET	GLEAM	0.25° × 0.25° (monthly)	2003–2019	https://www.gleam.eu/ (accessed on 16 March 2022).
	ERA5	0.25° × 0.25° (monthly)	2003–2020	https://cds.climate.copernicus.eu/cdsapp#!/home (accessed on 23 March 2022).
	MOD16	0.01° × 0.01° (monthly)	2003–2020	https://search.earthdata.nasa.gov/search?q=MOD16A2+V006 (accessed on 16 March 2022).
Groundwater level	Ground observations	Stations (monthly)	2005–2014	China Geological Environment Monitoring Institute
			2014, 2016, 2018–2020	Ministry of Water Resources

(2) GRACE-derived GWSA data

In this study, version 2 of the GRACE (January 2003–June 2017) and GRACE-FO (June 2018–December 2020) RL06 mascon solution, processed by the Jet Propulsion Laboratory (JPL), was used to estimate the monthly TWS changes. The JPL mascon solution uses an area concentration function to constrain the mass variation, which is expressed as a 3° × 3° spherical caps of equal global area. Although this solution provides TWS data with a 0.5° × 0.5° spatial sampling resolution, the values of each grid cell in the same cap were identical and were further resampled to 1° × 1° [50]. More information on the data processing of mason solutions and the leakage error reduction of the JPL mascon solution can be found in Watkins et al. [54], Wiese et al. [55], and Save et al. [56]. The GRACE data short-term gaps were filled using the linear difference method during 2003–2020, except for the 11-month gap (July 2017–May 2018) between GRACE and GRACE-FO.

The monthly data of soil moisture storage (SMS) and snow water equivalent (SWE) were obtained from the Global Land Data Assimilation System (GLDAS) Noah V2.1 model with a spatial resolution of 1° × 1° because no obvious differences in SMS and SWE were found in four land surface models (VIC, CLM, CLSM, and Noah) [39]. The SMS and SWE were also estimated by removing the average values between 2004 and 2009 to keep

consistency with the GRACE-derived TWS anomalies. Considering the scarcity of surface water in the arid area of the NCP, the surface water storage capacity of rivers, lakes, and reservoirs was neglected [57]. Therefore, we could obtain GWSA by deducting the amount of soil water anomalies and snow anomalies in the area from the TWS anomalies.

(3) Precipitation and actual evapotranspiration data

The precipitation data were obtained from the Tropical Rainfall Measuring Mission (TRMM) 3B43, China Meteorological Administration (CMA), European Centre for Medium-Range Weather Forecasts ERA5 reanalysis, and PENG datasets [58]. Actual evapotranspiration (AET) datasets could be obtained from the Global Land-surface Evaporation: the Amsterdam Methodology (GLEAM) V3.5 dataset, ERA5 reanalysis, and Moderate-Resolution Imaging Spectroradiometer (MODIS) ET algorithm (MOD16) AET datasets. The different datasets are compared in Figure 2. In the precipitation products, the TRMM and ERA5 data were similar, while the PENG datasets were relatively small. The AET products varied greatly in summer, and MOD16 overestimated the values from May to August but underestimated the values in other months. The correlations between the different rainfall and evapotranspiration datasets were all close to 1. We discuss the impacts of the uncertainty of the different datasets on the model calculation results.

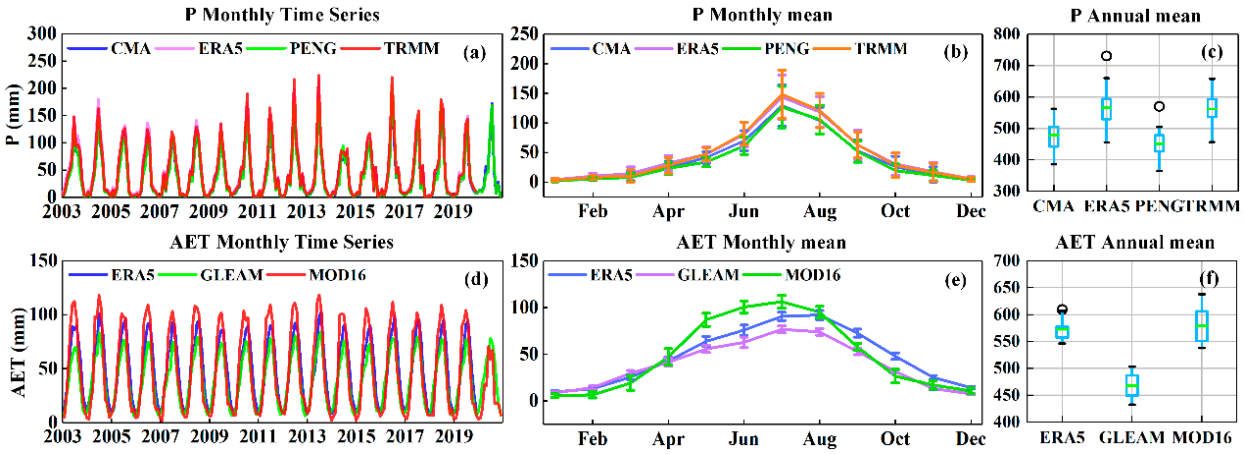


Figure 2. Plots summarizing datasets describing variables. From top to bottom, the panels depict precipitation (P), AET. (**a**,**d**) represent monthly time series, (**b**,**e**) represent monthly means, and (**c**,**f**) represent annual means.

3. Methods

3.1. Downscaling Method

(1) Groundwater balance

In the study area, precipitation infiltration accounted for 86.3% of the total groundwater recharge based on the groundwater budget analysis [52,59]. Groundwater discharge is primarily caused by groundwater withdrawal and evapotranspiration (ET). Groundwater withdrawal has become a major method for groundwater discharge, supplying water for agricultural, municipal, industrial, and ecological uses [59,60], with more than 70% of the exploited groundwater being used for agricultural irrigation [61]. In the study area, most of the irrigation water is lost through ET, and the irrigation return flow is considered internalized water redistribution, which is implicitly accounted for in Equation (1):

$$P = ET + R + \Delta TWS \quad (1)$$

where P is the precipitation, ET is the actual evapotranspiration, R is the net runoff, and ΔTWS is the total water storage change, including soil and groundwater storage changes.

Considering the scarcity of surface water in the arid area of the BTH region, the contribution of the surface water storage capacity, including rivers, lakes, and reservoirs,

is neglected [7]. The change in the long-term average soil storage is relatively small in the NCP, and the water storage change is primarily driven by the groundwater storage change [62]. Therefore, the changes in groundwater storage are mainly subject to changes in precipitation and *ET*:

$$\Delta GWS = P - ET + r \tag{2}$$

where ΔGWS is groundwater storage change and r is the error due to the simplification.

(2) Groundwater storage model equations

The aquifer system of the study area was divided into rectangular cells, the spatial resolution of which can be flexible: either 3°, the same size as the output from the GRACE data, or 0.05° or smaller. Based on Darcy's law and the water balance principle, the groundwater storage changes in a certain cell should be equal to the lateral flux changes from its neighboring cells and the vertical flux changes from precipitation infiltrating and actual evapotranspiration, which leads to

$$\sum_i T_{ei} \frac{h_{ei}^{n+1} - h_e^{n+1}}{d_{ei1} + d_{ei2}} L_{ei} + A_e \sum_e Q^n = S_{ye} \frac{h_e^{n+1} - h_e^n}{\Delta t} A_e \quad i = 1, 2, 3, 4 \tag{3}$$

$$\sum_e Q^n = \alpha_e Q_p^n - \beta_e Q_w^n + Q_c^n \tag{4}$$

where i is the number of cells around cell e; n is the nth time step; e_i represents the ith cell around cell e; T_{ei} is the transmissivity for ith cell around cell e [L^2T^{-1}]; h_e^n and h_{ei}^n are the groundwater level for cell e and its neighboring cell ei at the nth time step, respectively [L]; d_{ei1} and d_{ei2} are the distance from the center of the cell and that of the cell neighboring from the adjacent boundary, respectively [L]; L_{ei} is the lateral width of the aquifer [L]; A_e is the area of cell e [L^2]; α_e is the infiltration rate of precipitation (dimensionless); Q_p^n is the monthly precipitation [LT^{-1}]; β_e is the evaporation rate of the groundwater (dimensionless); Q_w^n is the monthly evapotranspiration [LT^{-1}]; Q_c^n is additional monthly groundwater recharge and discharge from such activities as local pumping [LT^{-1}]; S_{ye} is the comprehensive specific yield (dimensionless); and Δt is a time step chosen to be 1 month, the same as the time interval of the GRACE data [T].

The GRACE-derived groundwater storage anomaly is defined as the difference between the indicated groundwater storage and the average groundwater storage from January 2004 to December 2009 [L], which is denoted as (GW_g) in the equation. It is assumed that GWSA is accurate and error-free. The relationship between GWL and GW_g at a certain cell e can be expressed as:

$$GW_{ge}^n = S_{ye}(h_e^n - h_{avge}) \tag{5}$$

where h_{avge} is the average GWL for cell e from January 2004 to December 2009 [L].

Hence, Equation (3) can be rewritten as the following equations:

$$GW_{ge}^{n+1}(-\sum_i \frac{T_{ei}L_{ei}}{S_{ye}(d_{ei1} + d_{ei2})} - \frac{A_e}{\Delta t}) = -\sum_i \frac{GW_{gei}^{n+1} T_{ei} L_{ei}}{S_{yei}(d_{ei1} + d_{ei2})} - \sum_i T_{ei} \frac{h_{avgei} - h_{avge}}{d_{ei1} + d_{ei2}} L_{ei} - A_e \sum_e Q^n - \frac{GW_{ge}^n}{\Delta t} A_e \tag{6}$$

It should be noted that observation wells are very limited and unevenly distributed in most conditions and that the initial hydraulic gradient in groundwater systems is difficult to achieve. Therefore, it is assumed that initial hydraulic gradient can be estimated by multiplying the surface slope by an unknown coefficient, which is denoted as hydraulic gradient coefficient.

$$T_{ei} \frac{h_{avgei} - h_{avge}}{d_{ei1} + d_{ei2}} L_{ei} \approx T_{ei} \frac{Z_{avgei} - Z_{avge}}{d_{ei1} + d_{ei2}} C_h L_{ei} \tag{7}$$

where Z_{avge} and Z_{avgei} are the average surface elevations at cell e and the neighboring cell e_i, respectively, and C_h is the hydraulic gradient coefficient.

For each grid cell, we obtained a series of equations that may be resolved by considering the boundaries and initial conditions. Thus, the GW_g in each cell can be obtained from the equation. The principle of the model has also been applied in the Shiyang River basin in northwest China [63].

(3) Downscaling

The downscaling can be divided into two steps. The first step is the groundwater storage model construction, calibration, and verification. The study area was divided into $1° \times 1°$ grids in the horizontal direction. The number of simulation layers was one, and it was divided into 81 grids in total (Figure 3). The GWSA for January 2003 was set as the initial condition. The cells along all the boundaries (G50–G81) and cells G28, G34, G35, and G42 (which are close to the sea) are represented by the Dirichlet boundary with the GRACE-derived GWSA. The precipitation and actual evapotranspiration data were used to develop the model. According to regional groundwater studies in China [64], four types of aquifers are present in the study area (Figure 3). The parameters of the model included the comprehensive specific yield (S_y), transmissivity (T), precipitation infiltration coefficient (α), evapotranspiration coefficient (β), and hydraulic gradient coefficient (C_h), which were determined according to empirical values from the different aquifers. The Morris one-at-a-time (MOAT) method was used for parameter calibration owing to its simplicity and efficiency [65,66]. The period from January 2003 to June 2017 was set as the calibration period, and June 2018 to December 2020 was set as the verification period (without considering the gap between GRACE and GRACE-FO). The time step was selected to be one month, which is consistent with GRACE data. The fitting objects of the model were the GRACE-derived time-series GWSA in 45 grid cells, except for the Dirichlet boundary grid cells. The groundwater storage model code (NGFLOW-GRACE) was used to calculate the GWSA for the study area.

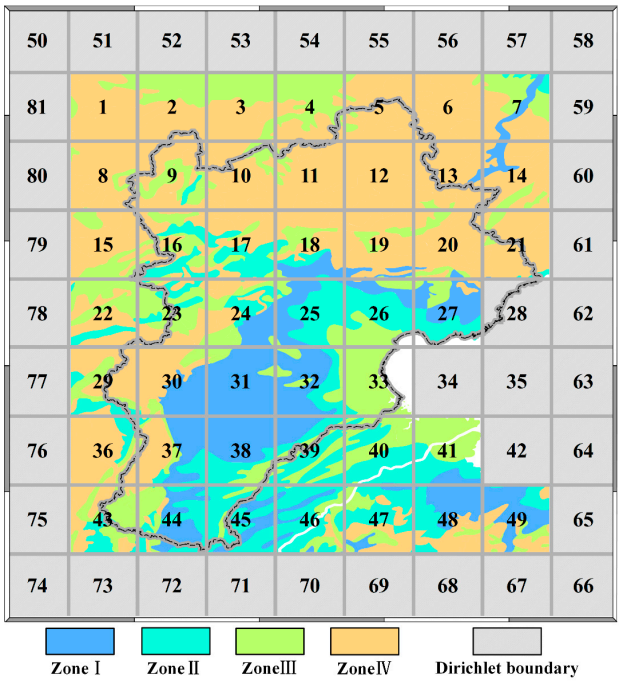


Figure 3. Parameter partitioning and boundary conditions of the study area. Zone I, Zone II, Zone III, and Zone IV represent the pore water of the alluvial–pluvial layer in the piedmont depositional plain, pore water of the alluvial–pluvial layer of the plain, pore water of the alluvial–pluvial layer of the intermountain basin or the coastal plain, and fissure water of the clastic rock in the mountains and hills or the hilly plain, respectively. The grey grid cells are the Dirichlet boundary.

The second step was to run the model with a refined mesh (with a spatial resolution of 0.05°) to obtain a higher-resolution GWSA. During the two processes, the distribution

of the hydrogeological parameters remained the same and was only estimated in the first step, further spreading to 0.05° space. The model parameters for the final calibration are discussed in Section 4.3. The precipitation and evapotranspiration data with a resolution of 0.05° were used to drive the model and obtain the GWSA at small scales.

3.2. Evaluation Methods

The GRACE-derived GWSA was marked as GW_{g_obs}, and the GWSA simulated from the NGFLOW-GRACE model was denoted as GW_{g_sim}. The model performance was measured using total root mean square error (Φ), as shown below:

$$\Phi = \sqrt{\frac{1}{N}\sum_{m=1}^{N}\sum_{n=1}^{Ndt}\left(GW_{mg_sim}^{n} - GW_{mg_obs}^{n}\right)^2} \quad (8)$$

where N is the number of cells with an unknown GWSA in the model area. The time step is Ndt, and m and n are the continuous counts of grid cells and time steps, respectively.

A shuffled complex evolution method was used for parameter estimation [65]. The parameter optimization strategy is to minimize Φ.

The root mean squared error (RMSE) and Nash–Sutcliffe efficiency coefficient (NSE) were used to assess the performance of the model data products. The RMSE was used to further measure the deviation between the simulated and GRACE-derived and ground-based observations in each grid cell. The NSE was used to evaluate the degree of fit of the regression models. An RMSE close to 0 and NSE close to 1.0 indicate that the model performance is reliable.

$$RMSE = \sqrt{\frac{1}{n}\sum_{1}^{n}(G_i - P_i)^2} \quad (9)$$

$$NSE = 1 - \frac{\sum_{i=1}^{n}(G_i - P_i)^2}{\sum_{i=1}^{n}\left(G_i - \overline{G_i}\right)^2} \quad (10)$$

where n is the total amount of data, G_i is the GRACE-derived results or ground-based observations, P_i is the model-simulated results, and $\overline{G_i}$ is the average value.

4. Results

4.1. Model Evaluation

The calibration period of the model was from January 2003 to June 2017. The TRMM (precipitation) and GLEAM (AET) data were first used to drive the model, in which TRMM data have been proven to match the in situ observations in the study area well [7]. Figure 4 illustrates that the GWSA calculated by the model and the GWSA by GRACE-derived data are consistent during the calibration period. Figure 5 shows the model evaluation metrics RMSEs and NSEs for the 45 grid cells, where the RMSEs change from 0.81 cm EWH to 5 cm EWH with an average value of 2.61 cm EWH and the NSEs vary from 0.41 to 0.96 with an average value of 0.84. During the validation period from June 2018 to December 2020, the trend of the simulated values and GRACE-derived values remain similar, with a minimum RMSE of 1 cm EWH and a maximum NSE of 0.83. Overall, the simulated results agree well with the GRACE-derived results. The simulated GWSA matched the GRACE-derived results the best for the northwestern mountainous area, where the changes in GWSA were less affected by human activities.

4.2. Downscaled GWS Changes

After estimating the hydrogeological parameters, the model was refined from a spatial resolution of 1° (Figure S1a) to 0.05° (Figure S1b). Therefore, a grid cell at a spatial scale of 1° was refined to 400 sub-grid cells at a spatial scale of 0.05°, and the spatial parameters in Figure 3 were averaged from a resolution of 1° to 0.05°. For the model with a spatial scale of 0.05°, the PENG precipitation and MOD16 AET data were used to drive the model. The initial conditions and boundary conditions were the same as those used in the calibrated

model. The simulation period for the downscaled models was from January 2003 to December 2020 without considering the gap between the GRACE and GRACE-FO data.

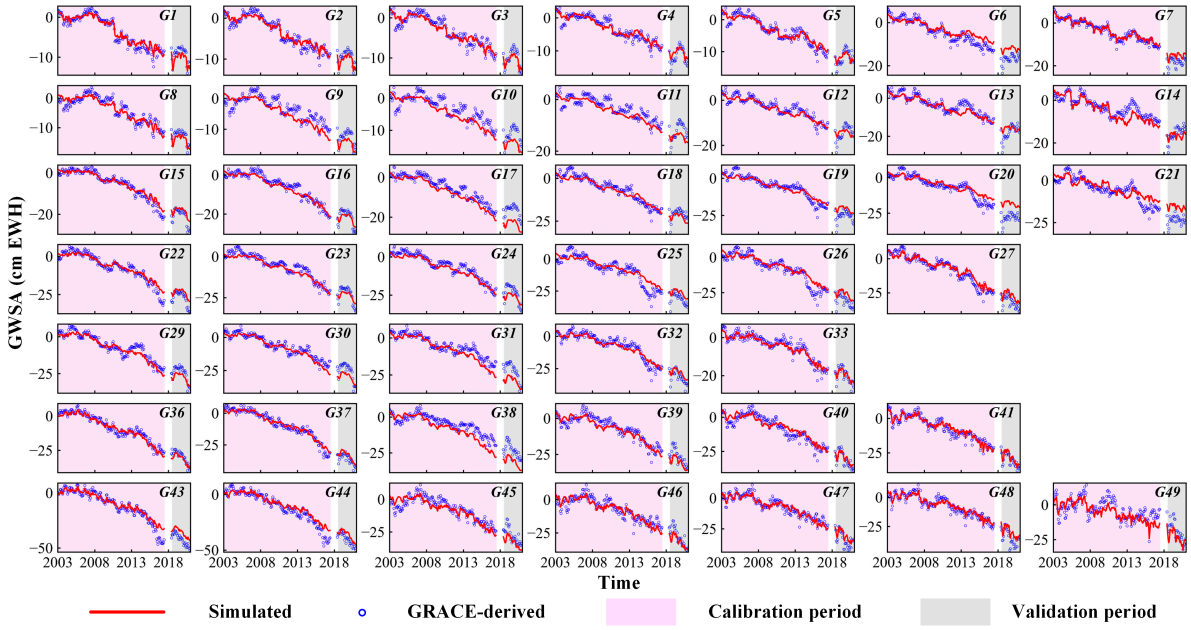


Figure 4. Comparison of GRACE-derived and simulated GWSA for 45 grid cells for both the calibration and validation periods. The pink range is the model calibration period, the gray range is the model validation period, the red lines are the model calculation result, and the blue scatters are the GRACE-derived results.

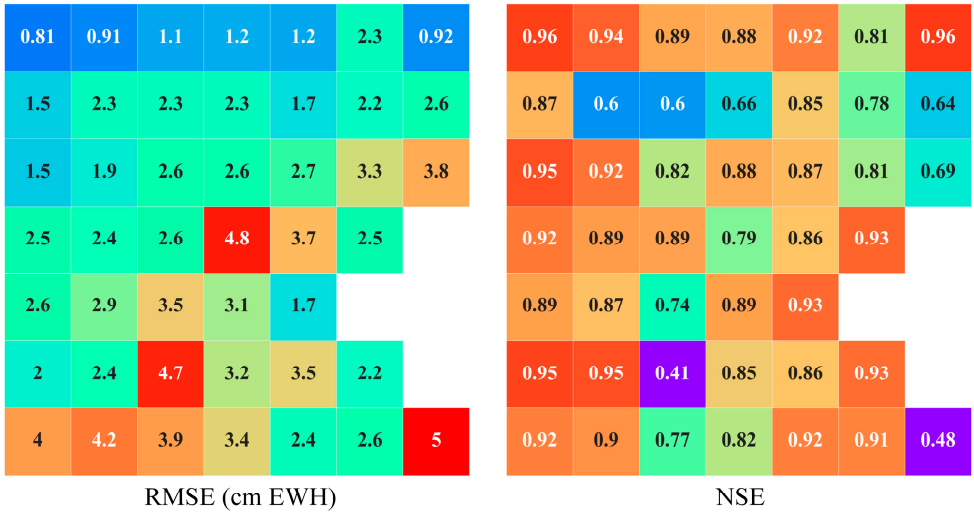


Figure 5. Distribution of RMSEs and NSEs over 45 grid cells for the calibration period.

To better demonstrate the characteristics of the downscaled results, the GWSA of a 1° grid cell was compared with that with a spatial resolution of 0.05° for each of the 45 grid cells (Figure S2). The downscaled GWSA maintained the same pattern as that of the same 1° grid. The correlation coefficient of the 0.05° average GWSA within each grid and the GRACE-derived GWSA in the same 1° grid were above 0.9. Figure 6 contrasts the GRACE-derived GWSA variations with the downscaled GWSA variations on a spatial scale for 2019. Three transects (P1–P2, P3–P4, and P5–P6 across 4°, 4°, and 1°, respectively) demonstrate the high-resolution GWSA data along the transect capture the spatial variations of the GWSA, which appear to be uniform in the coarse-resolution GRACE data. High-resolution GWSA data are particularly useful for small regions in which the downscaled GWS data

capture the distribution variations. For example, the downscaled data reflect a strong spatial heterogeneity in the GWSA variations (−22.98 to −39.05 cm EWH) in transect P5–P6 with a substantial decline observed, whereas the GRACE-derived GWSA is represented by homogenized pixels. Therefore, the downscaling results not only exhibit a time series variation trend consistent with that of the GRACE-derived values but also exhibit finer spatial heterogeneity.

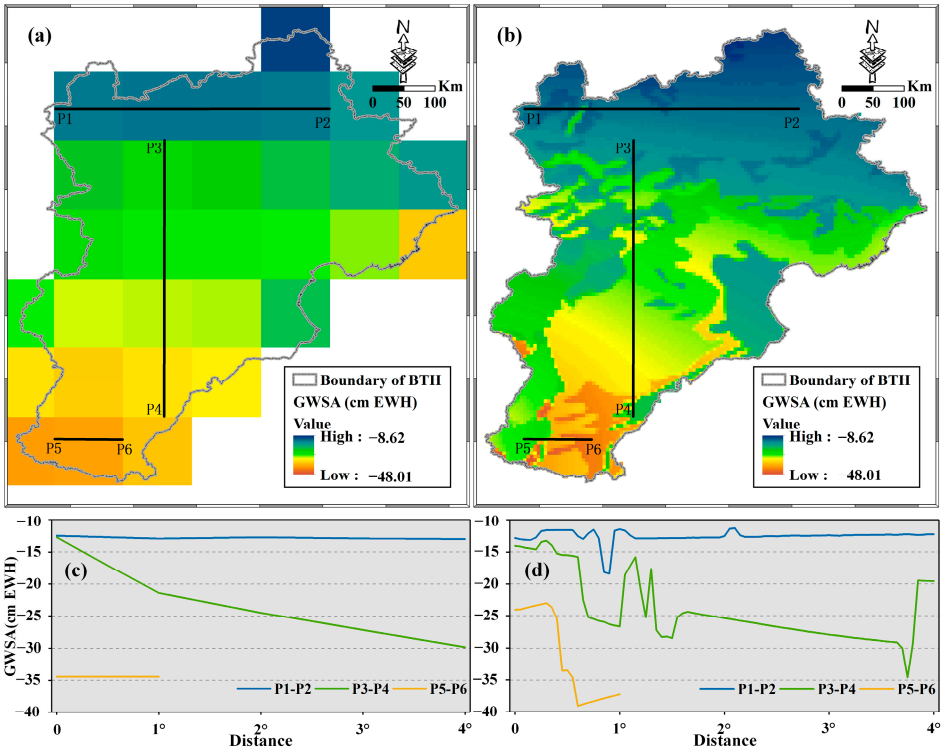


Figure 6. Comparison of GRACE-derived GWSA and downscaled GWSA at the resolution of 1° in the area for 2019. (**a**,**b**) represent the spatial distribution of GRACE-derived GWSA with the resolution of 1° and 0.05°, and (**c**,**d**) represent the change of GWSA in different transects with the resolution of 1° and 0.05°, respectively.

4.3. Validation of Downscaling Results

The time series of the GRACE-derived and downscaled GWSA are compared with the in situ GWL data in Figure 7. The result showed that the GWSA agrees well with the in situ GWL observations, that is, the downscaled results were also capable of capturing the significant groundwater change information in most cases. The GWSA had a significant downward trend from 2005 to 2014, which can be attributed to excessive demand for groundwater required for agriculture during the 2003–2009 precipitation deficiency period and average annual groundwater withdrawal of 2.5×10^{10} m^3, resulting in unsustainable recovery of groundwater storage [7,11].

Figure 8 quantifies the comparison between the GRACE-derived, downscaled GWSA and in situ observed GWL in terms of correlation coefficients (CC) and RMSE. For the computation of RMSE and CC, two time series were normalized to [−1,1] because their units were inconsistent. Downscaling the GRACE data led to an increase in the CC and a decrease in the RMSE, indicating the downscaling method improved GWSA computation. On average, higher correlations are observed after downscaling, which are improved from 0.78 to 0.85. The best cases occur for the H52 well, with the CC increasing from 0.74 to 0.91 and the RMSE decreasing from 0.34 to 0.21 in the B4 well. Overall, the results indicate that the downscaled process increases the correlation coefficients between the GRACE-derived GWSA and in situ GWL measurements by 0.06 (on average) and reduces the RMSE by 8.85% (on average). The least agreement between the downscaled results and GWL was

found with the shallow monitoring wells. The key reasons for this could be due to the exploitation of the deep groundwater.

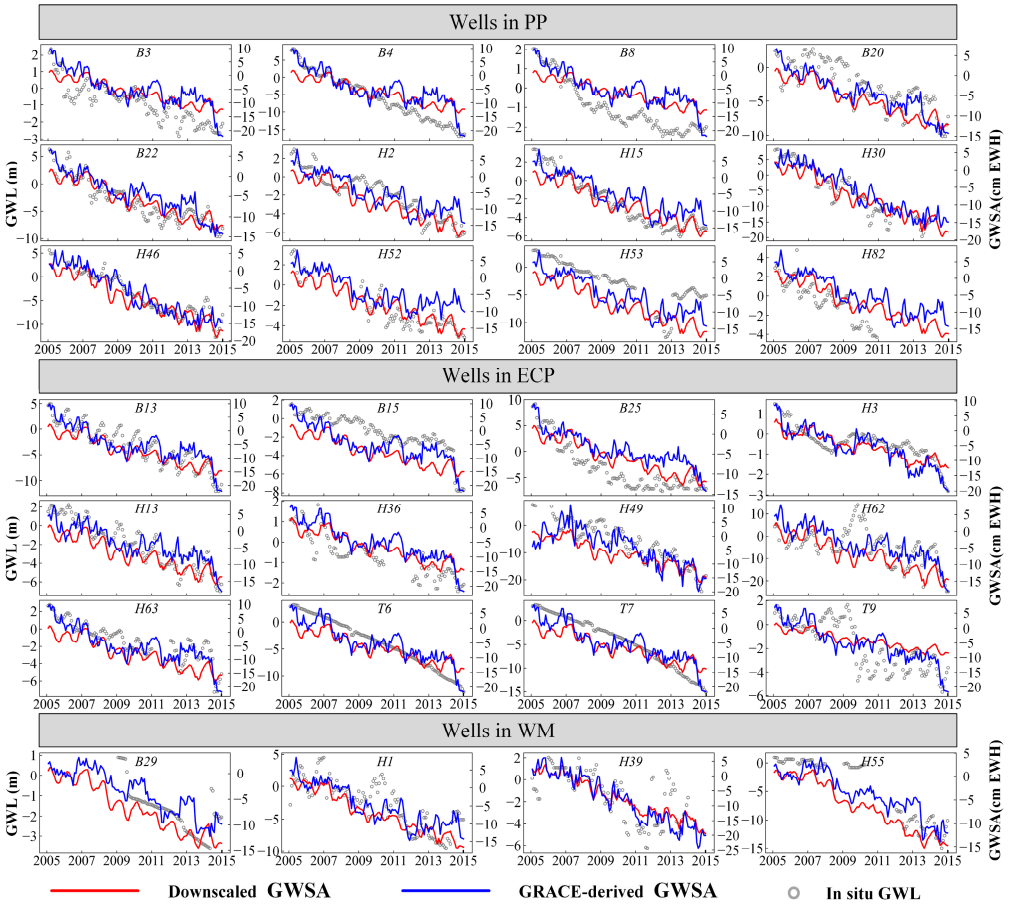


Figure 7. The gray scatters are the monitoring well GWL observations, which have been removed from the average from 2005 to 2009 (orange red dots shown in Figure 1); the red and blue lines represent corresponding monitoring well locations with downscaled GWSA in the 0.05° grid and GRACE-derived GWSA in the 1° grid.

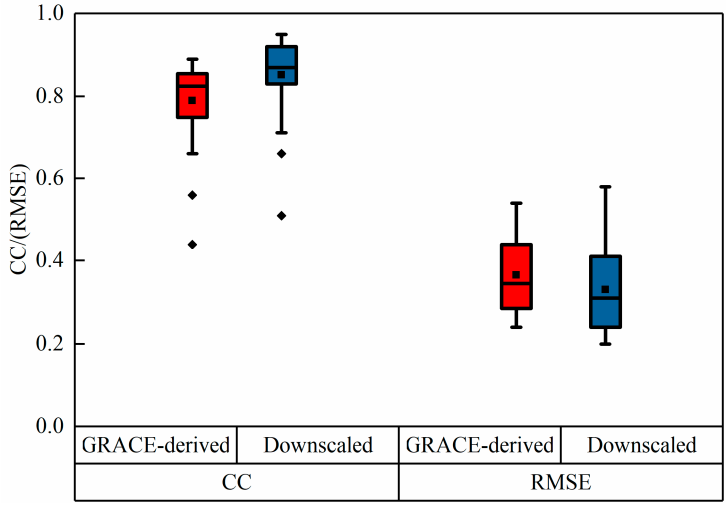


Figure 8. The correlation coefficients and RMSEs from monitoring well GWL observations with corresponding monitoring well locations with downscaled GWSA in the 0.05° grid and GRACE-derived GWSA in the 1° grid after normalization.

5. Discussion

5.1. Spatiotemporal Changes of Downscaled GWSA in Subregions

The linear least squares regression method was used to calculate the trends of the variables in the current study. Figure 9 presents the temporal fluctuations of the monthly GWSA in the BTH region and three subregions (PP, ECP, and WM) from January 2003 to December 2020. The GWS declines over the BTH with different decline rates during the four periods, i.e., January 2003–December 2008 (−0.06 cm/month), January 2009–December 2014 (−0.097 cm/month), January 2015–June 2017 (−0.216 cm/month), June 2018–December 2020 (−0.143 cm/month). The time series demonstrate that the three subregions exhibit similar variations; GWSA has a decreasing trend, ranging from −0.053 to −0.071 cm/month in January 2003– December 2008, followed by a larger fluctuation and decrease rate in the PP from January 2009 to December 2014. However, GWS has a continuously accelerated decline with a large difference in the subregions from January 2015 to June 2017, and the rate of decrease is more than two times those of the previous period, which are −0.313 cm/month, −0.183 cm/month, −0.239 cm/month, and −0.216 cm/month in the PP, WM, ECP, and BTH, respectively; similar trends have been reported in the literature [67,68]. The western mountainous areas showed the slowest decline in the different periods. Since 2018, the rate of decline has slowed in the BTH and the three subregions.

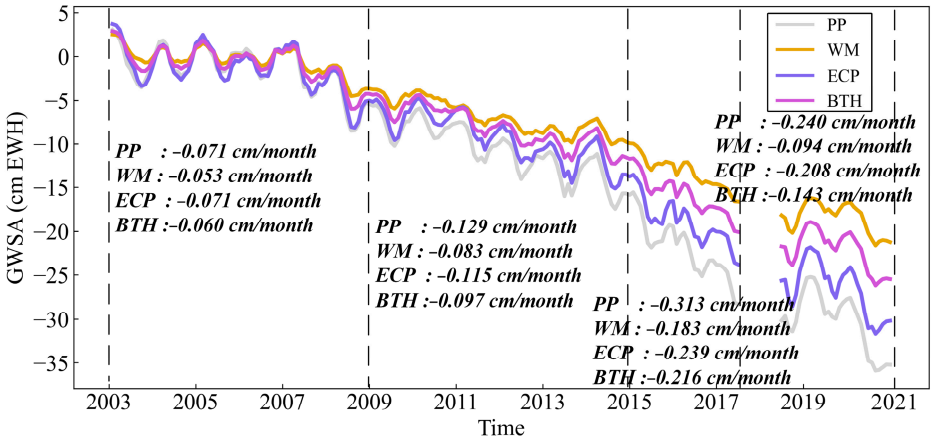


Figure 9. The GWSA trends in different subregions over time; PP, ECP, WM, BTH represent the piedmont plain, eastern central plains, western mountainous area, and the entire study area, respectively.

Figure 10 shows the spatial patterns of the linear trends of the 0.05° GWSA with the same legend range from −0.562 to −0.084 cm/month. For the spatial distribution, the GWSA generally declined from 2003 to 2020, while areas of groundwater storage decline are more visible in the PP. The downward trend gradually increased from northeast to southwest, and the maximum decline rate appeared in Handan City, Hebei Province. From January 2015 to June 2017, the decline rate of the GWSA in the PP was the most serious and was mainly distributed in Handan, central Xingtai, eastern Shijiazhuang, and western Cangzhou. In addition, the decline in the GWSA in the Zhangjiakou area intensified during this period. From June 2018 to December 2020, especially in Beijing, the GWS decline slowed. The coastal areas south of Tianjin and west of Cangzhou experienced stable GWS changes over the years. In the BTH region, the decline rate of the GWS gradually decreased owing to the groundwater overexploitation control policy in recent years.

5.2. Comparison of GWS with GWL Observations in Administrative Regions

To demonstrate the relationship between the GWL changes and GWS changes in the study area, the downscaled GWS changes were compared using the GWL changes from 843 observation wells in the subregions, including Cangzhou, Xingtai, Handan, Tangshan in Hebei Province, Beijing, and Tianjin. Because of the different time series of the collected monitoring well data, the GWL and GWS data for Cangzhou, Xingtai, Handan,

and Tangshan from December 2018 to December 2019 were selected for comparison. The monitoring wells were divided into shallow and deep wells. Beijing and Tianjin monitoring wells from December 2014 to December 2016 were shallow wells. The GWS was not represented by layers (because the GWS derived from GRACE was difficult to divide vertically). It should be noted that owing to the true value of specific yield is hard to obtain, we only compared groundwater level data from observation wells with groundwater storage changes in the pixel including the wells. The GWS changes legend for Beijing and Tianjin were different from those in the other three subregions.

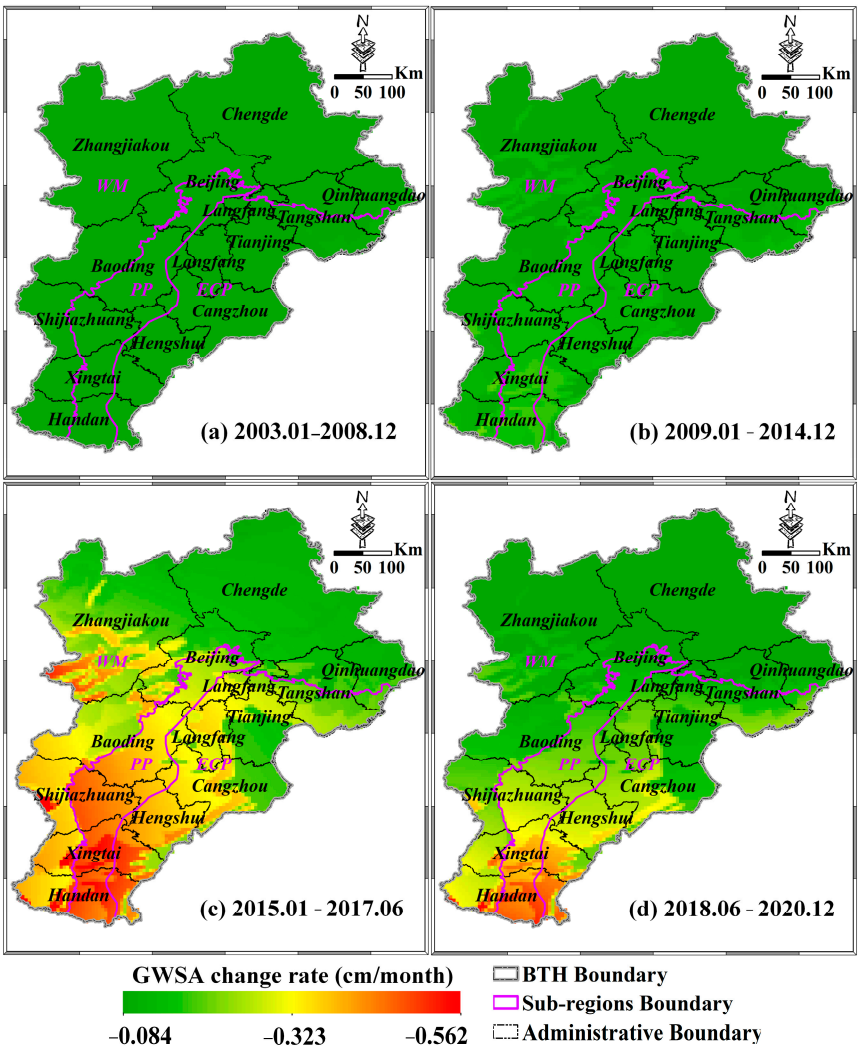


Figure 10. Spatial distribution of GWSA trends in different periods.

Figure 11 shows that the change trend of the deep monitoring wells and that of the GWS changes in the spatial location correspond well; in other words, at locations where the GWL drops more, the GWS decreases more, such as at the locations in the red dotted circle. However, there are also inconsistencies, such as the locations in the blue dotted circle; the GWL drops are the largest, but the GWS decline is relatively small. This could be attributed to the local groundwater extraction, which is not considered in the model due to the lack of detailed information.

Figure 12 shows that the GWS and GWL changes also exhibited consistent trends in the monthly variations. As shown in Figure 12a–d,f, from March to June, the GWS gradually decreased and the GWL also gradually decreased, mainly due to groundwater exploitation for agricultural irrigation. With the increase in precipitation in July, the GWS and GWL showed an increasing trend. In August, the GWL demonstrated an upward

trend, but the GWS continued to decrease, which is difficult to explain. From October to December, when groundwater exploitation was slow, the GWS gradually recovered and the GWL also slowly increased. However, Figure 12e shows that the variation trends of GWS and GWL are inconsistent, which indicates that in the absence of deep-monitoring GWL verifications, the variation trend of the GWS cannot adequately explain the variation in the shallow GWL in Beijing.

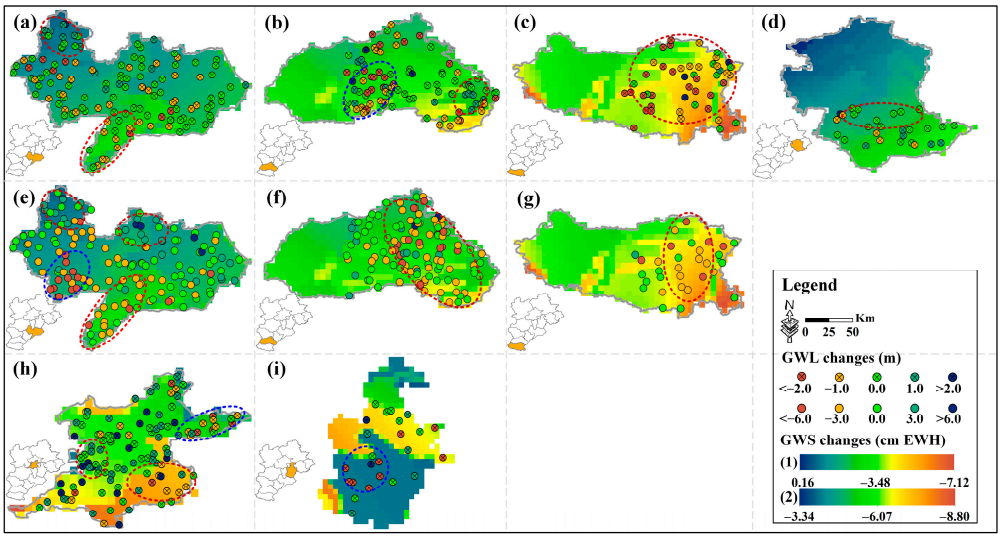


Figure 11. Comparison of GWL and GWS changes in different subregions. (**a**–**d**) represent the GWL changes in the shallow aquifers in Cangzhou, Xingtai, Handan, and Tangshan, (**e**–**g**) represent the GWL changes of deep aquifers (without Tangshan), and GWS changes with color bar (1) from December 2018 to December 2019; (**h**,**i**) represent the GWL changes in the shallow aquifer, and GWS changes with color bar (2) from December 2014 to December 2016 in Beijing and Tianjin.

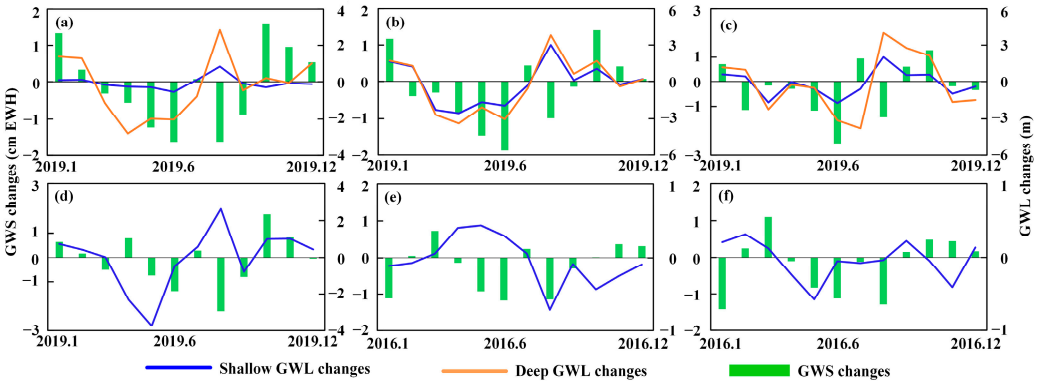


Figure 12. Comparison of monthly GWL changes and GWS changes. (**a**–**d**) represent the monthly GWL changes and GWS changes in Cangzhou, Xingtai, Handan, and Tangshan, respectively, in 2019; (**e**,**f**) represent Beijing and Tianjin in 2016, respectively.

Table 2 summarizes the GWL changes in the monitoring wells and GWS changes in each subregion during the different time periods. Between December 2018 and December 2019, the GWL in the shallow aquifers of Tangshan, Cangzhou, Xingtai, and Handan dropped by 0.42 m, 0.71 m, 1.15 m, and 2.36 m, respectively, and the average GWL in the deep aquifers dropped by 2.30 m, 3.63 m, and 3.76 m (without the Tangshan), respectively. The decrease in the GWL in the deep aquifers was generally greater than that in the shallow aquifers, indicating that the main groundwater consumption was from the deep confined aquifer. The GWS decreased on average by 1.21 cm EWH, 1.37 cm EWH, 2.55 cm EWH, and 3.65 cm EWH, respectively. The good correspondence between the GWL and GWS changes indicated that groundwater consumption was the most serious in Handan.

Table 2. Statistics of GWL observations and GWS changes in the subregions.

Time	December 2018–December 2019				December 2014–December 2016	
Sub-Regions	**Cang Zhou**	**Xing Tai**	**Han Dan**	**Tang Shan**	**Tian Jin**	**Bei Jing**
Number of shallow wells (up)	16	21	4	7	17	69
Number of shallow wells (down)	118	90	48	23	26	67
Number of deep wells (up)	31	13	1	—-	—-	—-
Number of deep wells (down)	145	117	30	—-	—-	—-
GWL changes in shallow aquifers (m)	−0.71	−1.15	−2.36	−0.42	−0.98	0.29
GWL changes in deep aquifers (m)	−2.30	−3.63	−3.76	—-	—-	—-
GWS changes (cm EWH)	−1.37	−2.55	−3.65	−1.21	−5.06	−6.14

From December 2014 to December 2016, the GWS in Beijing and Tianjin decreased by 6.14 cm EWH and 5.06 cm EWH, respectively, and the GWS in southwest Beijing and central Tianjin declined the most. From the perspective of GWL changes, the average GWL in Tianjin dropped by 0.98 m, while the average GWL in Beijing rose by 0.29 m, which is closer to the 0.43 m increase from 2014 to 2016 listed in the Beijing Water Statistical Yearbook in 2016 [69]. The GWL and GWS changes in Beijing exhibited completely different trends, which were similar to the results of GRACE-derived GWS changes and GWS changes estimated using GWL by Zhang et al. [50]. There may be two reasons for this result. First, the uncertainty of the GRACE data because the subregion GRACE data are contaminated by signal leakage and amplitude-damping effects [11,22,50]. Second, the water from the South-to-North Water Diversion began to enter Beijing in December 2014, replacing the exploitation of some groundwater sources, and the surplus water of the South-to-North Water Diversion was used for groundwater replenishment. The GWL in Beijing has gradually recovered since 2015; however, the effects related to human factors were not considered in the model.

5.3. Evaluation of Hydrogeological Parameters

Figure S3 presents the spatial distribution of the relevant parameters in the downscaled model. The comprehensive specific yield ranged from 8.0×10^{-4} to 1.0×10^{-3} in the piedmont plain gradually decreased to 4.0×10^{-5} to 1.0×10^{-4} in the coastal areas, and the intermountain basins had the smallest comprehensive specific yield of less than 4.0×10^{-5}. This paper compares the parameters results with previous studies. Previous research has shown that the specific yield of shallow aquifers in the plain area of the study area is between 0.025 and 0.29, the storage coefficient in the deep aquifers decreases from 1.0×10^{-3} to 8.0×10^{-3} in alluvial plains to 4.0×10^{-4} to 5.0×10^{-4} in coastal plains, and the average storage coefficient of deep aquifers is 0.00125, approximately 2% of that of the shallow aquifer area-weighted specific yield (0.075) [22,47,50,59]. Shao et al. [59] determined that the specific yield of the aquifer exhibits a downward trend from the piedmont alluvial plain to both sides of the river valley, that the specific yield ranges from 0.03 to 0.23, and that the order of magnitude of the storage coefficient in the confined layer is 10^{-4}–10^{-6}. Li et al. [52] calculated specific yield ranging from 0.025 to 0.2 and storage coefficients in the confined layer ranging from 2.7×10^{-5} to 2.3×10^{-4}, which were provided in the numerical simulation. The model simulated the comprehensive changes of phreatic and confined aquifer. Thus, we redefined traditional specific yield as comprehensive specific yield in this study, as mentioned above. We found that our optimized storage coefficient has the same order of magnitude as the elastic storage coefficient of confined water. The groundwater exploitation in the study area is dominated by deep confined water. Therefore, we think the optimized comprehensive water supply is reasonable.

The precipitation infiltration coefficients in the study area ranged from 0.02 to 0.30. The precipitation infiltration coefficients were greater than 0.2 for the piedmont plains, between 0.05 and 0.2 for the central and eastern plains, and less than 0.03 for the mountainous and coastal plains. This result is similar to the parameters of the plain area optimized by

Li et al. [52], which was spatially distributed within the range of 0.17 to 0.25. The evapotranspiration coefficient ranged from 0.03 to 0.66. The evapotranspiration coefficients were greater than 0.50 in the piedmont plains—less than 0.1 in most areas of the intermountain basin, 0.1–0.25 in the coastal areas, and 0.25–0.5 in most of the central and eastern plains, which is consistent with the parameters in the *Handbook of Hydrogeology* [70].

5.4. Uncertainty Analysis

The accuracy of a model is restricted by uncertainties in climate forcing (particularly precipitation and AET) and model parameters as well as deficiencies in the model structure. The MOAT method with 1000 samplings for each parameter was used to determine the sensitivity of the model to the aquifer parameters. As shown in Figure S4, the hydraulic gradient coefficient and comprehensive specific yield for the porous aquifer had the highest degrees of sensitivity. The parameter with the lowest degree of sensitivity was transmissivity. Given that precipitation and evapotranspiration are the main forcing data of the model, 12 scenarios with different precipitation and AET data inputs were set to examine the influences of the model input data. Figure S5 illustrates that the GWSA trends from the twelve scenarios exhibit the same patterns over the entire area, and the maximum absolute difference in the GWS changes is less than 1 cm EWH, indicating that although there are some differences between precipitation and evapotranspiration from different data sources, the different forcing data may not have substantially influenced the groundwater storage changes. Therefore, the average groundwater storage values for the different forcing data were not further implemented.

5.5. Advantages and Limitations

The groundwater storage model was established using the free GRACE data applied to improve the GWSA resolution from 1° to 0.05° and assess the groundwater storage changes at small scales in this study. This study provides a new method for the downscale study of GRACE data, which is based on the hydrological process of groundwater. This is an example of the combination of remote sensing, in situ monitoring, and hydrological models. Remote sensing data provide driving data and constraints for the model, and in situ monitoring provides an important data source for model verification. This study can provide a new idea for the combination of remote sensing data and numerical models and form an important dataset of groundwater storage changes for water resource management. However, the present study has some limitations. First, the established model considered the GRACE-derived GWSA as the fitting target; therefore, the accuracy of the GRACE-derived GWSA data also affects the accuracy of the downscaling results [71]. In this model, it is assumed that the GRACE-derived GWSA data are error free. Second, after generalizing the water balance in the study area, this study only considered groundwater storage changes driven by natural factors, such as precipitation and actual evapotranspiration. Surface water changes, such as reservoir scheduling, water transfer projects [53], and local groundwater extraction, were not fully considered in the model. Although the research results are relatively reasonable, to provide more comprehensive scientific guidance for policy making, the impact of human activities on groundwater storage changes must be considered in the model [72]. Third, comprehensive specific yield and hydraulic gradient coefficient were the most sensible variables. Whether those sensitive parameters will change with time and downscaled spatial is not discussed in this study. Furthermore, the observational data used for the evaluation were limited, and those will be the focus of future research.

6. Conclusions

Because the use of GRACE products for smaller areas is limited due to its coarse spatial resolution, this study utilizes a dynamic downscaling method to improve the GWSA resolution from 1° to 0.05° by constructing a groundwater storage numerical model with GWSA as the key variable. First, we constructed groundwater storage model and

performed model calibration and validation and verified the downscaling results. Then, the downscaling results were used to evaluate the temporal and spatial changes in groundwater storage at both the regional and local scales and were compared to the groundwater level changes in different administrative regions. In addition, the model parameters and uncertainty analysis were discussed. This study demonstrates the potential of downscaling GWSA data for local-scale applications. The following conclusions were drawn from this study:

(1) The study area is divided into $1^\circ \times 1^\circ$ grids with one layer in the horizontal direction, and the initial GWSA from January 2003, Dirichlet boundary, and initial parameters are given. After parameter optimization, the simulated GWSA at the spatial resolution of 1° had an average RMSE of 2.61 cm EWH and average NSE of 0.84 for the calibration period and the minimum RMSE of 1.0 cm EWH and maximum NSE of 0.83 for the validation period. The calibrated parameter distributions are consistent with those from the field investigation.

(2) The downscaled results exhibited a time series variation trend that was consistent with that of the GRACE-derived results, of which the correlation coefficient between the average GWSA within 0.05° grid and the GRACE-derived GWSA in the same 1° grid was above 0.9. The results also showed finer spatial heterogeneity. Compared to the GRACE-derived GWSA, the downscaled process increased the correlation coefficients between the GWSA and in situ measurements of the GWL by 0.06 (on average) and reduced the RMSE by 8.85% (on average).

(3) Based on the downscaling results, the variation pattern of groundwater storage demonstrated that the downward trend gradually increased from northeast to southwest in the study area, while groundwater storage in the piedmont plains declined significantly by 0.313 cm/month from January 2015 to June 2017, and the western mountainous areas declined the slowest during the different periods. However, the decline rate of the GWS has gradually decreased owing to the groundwater overexploitation control policy in recent years.

(4) The comparison between the GWL and GWS changes for different administrative regions showed good correspondence in both the spatial and monthly variations and indicated that groundwater consumption was the most serious in Handan, Hebei Province, where the GWL declined by 2.36 m and 3.76 m in the shallow and deep aquifers, respectively, and the GWS decreased by 3.65 cm EWH from December 2018 to December 2019, followed by Xingtai, Hebei Province.

(5) The MOAT method with 1000 samplings for each parameter in the model was used to carry out a sensitivity analysis, which indicated that the hydraulic gradient coefficient and comprehensive specific yield have the highest sensitivity and that transmissivity has the lowest sensitivity. In addition, different forcing data had no substantial influence on the groundwater storage changes.

The proposed method takes GRACE-derived groundwater storage data as the fitting variable of the model, which can be applied in remote areas where observation well data are lacking. Meanwhile, a dataset of groundwater storage changes with a higher resolution can be obtained and may be public for the further analysis of groundwater-related problems. For future work, the downscaling of groundwater storage changes at refined time scales will be conducted, as will the improvement of model reliability at refined spatial scales. Then, the relationship between downscaled groundwater storage changes and groundwater discharges under climate changes will further be explored to estimate the true groundwater withdrawals for effective groundwater management.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/rs15061490/s1, Figure S1. Comparisons of the two spatial scales, (a) 1° and (b) 0.05° in grid cells. Figure S2. Time series changes in GWSA with a spatial downscaling scale of 0.05° at a spatial resolution of 1° for 45 cells. The blue scatters are GRACE-derived GWSA at 1° grid, the grey band are the range of GWSA variation for all 0.05° grids at 1° grid, red lines are the average GWSA of all 0.05° grids at 1° grids. Figure S3. Spatial distribution of (a) precipitation

infiltration coefficients and (b) evapotranspiration coefficient and (c) comprehensive specific yield in the study area. Figure S4. Sensitivities of hydrogeology parameters, u^* and δ represent mean and standard deviation. T, α, β, and S_y represent the transmissivity, precipitation infiltration coefficient, evapotranspiration coefficient and comprehensive specific yield of the aquifer in the Zone I, respectively, and the subscripts a, b, and c represent the parameters in the Zone II, Zone III, and Zone IV, respectively, C_h represents the hydraulic gradient coefficient in the study area. Figure S5. Comparisons of the average GWSA for 12 scenarios with different driving data in this study area.

Author Contributions: J.S.: data collection and analysis, writing—original draft, writing-review and editing; L.H.: conceptualization, methodology, writing—review and editing; F.C. and L.Y., discussed and revised the manuscript; X.L. and K.S.: writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant number 41877173 and U2167211.

Data Availability Statement: This study provides the model program and input data files as well as the downscaled GWSA datasets, which can be downloaded from https://figshare.com/articles/dataset/Downscaled_GWSA_dataset/20406309 (uploaded on 30 July 2022) or are available from the first author upon request.

Acknowledgments: We appreciate the free access of the GRACE mascon solutions and GLDAS, TRMM, ERA5, GLEAM, MODIS, and PENG meteorological and hydrological datasets. GRACE mascon solutions: ftp://podaac-ftp.jpl.nasa.gov/allData/tellus/L3/land%20mass/RL06 (accessed on 11 March 2022); GLDAS: http://disc.gsfc.nasa.gov/hydrology/data-holdings (accessed on 11 March 2022); GLEAM: http://www.gleam.eu (accessed on 16 March 2022); MODIS: https://doi.org/10.5067/MODIS/MCD12Q1.006 (accessed on 16 March 2022); TRMM: https://gpm.nasa.gov/category/mission-affiliation/trmm (accessed on 21 March 2022); ERA5: https://cds.climate.copernicus.eu/cdsapp#!/home (accessed on 23 March 2022); PENG: http://poles.tpdc.ac.cn/zh-hans/data/faae7605-a0f2-4d18-b28f-5cee413766a2/ (accessed on 2 April 2022); CMA: http://data.cma.cn/ (accessed on 23 March 2022).

Conflicts of Interest: All authors declare that no conflict of interest exists.

References

1. Döll, P.; Müller Schmied, H.; Schuh, C.; Portmann, F.T.; Eicker, A. Global-scale assessment of groundwater depletion and related groundwater abstractions: Combining hydrological modeling with information from well observations and GRACE satellites. *Water Resour. Res.* **2014**, *50*, 5698–5720. [CrossRef]
2. de Graaf, I.E.; Gleeson, T.; Sutanudjaja, E.H.; Bierkens, M.F. Environmental flow limits to global groundwater pumping. *Nature* **2019**, *574*, 90–94. [CrossRef] [PubMed]
3. Wada, Y.; Van Beek, L.P.; Van Kempen, C.M.; Reckman, J.W.; Vasak, S.; Bierkens, M.F. Global depletion of groundwater resources. *Geophys. Res. Lett.* **2010**, *37*, L20402. [CrossRef]
4. Long, D.; Pan, Y.; Zhou, J.; Chen, Y.; Hou, X.; Hong, Y.; Scanlon, B.R.; Longuevergne, L. Global analysis of spatiotemporal variability in merged total water storage changes using multiple GRACE products and global hydrological models. *Remote Sens. Environ.* **2017**, *192*, 198–216. [CrossRef]
5. Alley, W.M.; Healy, R.W.; LaBaugh, J.W.; Reilly, T.E. Flow and storage in groundwater systems. *Science* **2002**, *296*, 1985–1990. [CrossRef]
6. Jasechko, S.; Perrone, D. Global groundwater wells at risk of running dry. *Science* **2021**, *372*, 418–421. [CrossRef]
7. Yin, W.; Han, S.C.; Zheng, W.; Yeo, I.Y.; Hu, L.; Tangdamrongsub, N.; Ghobadi-Far, K. Improved water storage estimates within the North China Plain by assimilating GRACE data into the CABLE model. *J. Hydrol.* **2020**, *590*, 125348. [CrossRef]
8. Tapley, B.D.; Bettadpur, S.; Ries, J.C.; Thompson, P.F.; Watkins, M.M. GRACE measurements of mass variability in the Earth system. *Science* **2004**, *305*, 503–505. [CrossRef]
9. Tapley, B.D.; Watkins, M.M.; Flechtner, F.; Reigber, C.; Bettadpur, S.; Rodell, M.; Sasgen, I.; Famiglietti, J.S.; Landerer, F.W.; Chambers, D.P.; et al. Contributions of GRACE to understanding climate change. *Nat. Clim. Chang.* **2019**, *9*, 358–369. [CrossRef]
10. Landerer, F.W.; Flechtner, F.M.; Save, H.; Webb, F.H.; Bandikova, T.; Bertiger, W.I.; Bettadpur, S.V.; Byun, S.H.; Dahle, C.; Dobslaw, H.; et al. Extending the global mass change data record: GRACE Follow-On instrument and science data performance. *Geophys. Res. Lett.* **2020**, *47*, e2020GL088306. [CrossRef]
11. Feng, W.; Zhong, M.; Lemoine, J.M.; Biancale, R.; Hsu, H.T.; Xia, J. Evaluation of groundwater depletion in North China using the Gravity Recovery and Climate Experiment (GRACE) data and ground-based measurements. *Water Resour. Res.* **2013**, *49*, 2110–2118. [CrossRef]

12. Hu, L.; Jiao, J.J. Calibration of a large-scale groundwater flow model using GRACE data: A case study in the Qaidam Basin, China. *Hydrogeol. J.* **2015**, *23*, 1305–1317. [CrossRef]
13. Rodell, M.; Velicogna, I.; Famiglietti, J.S. Satellite-based estimates of groundwater depletion in India. *Nature* **2009**, *460*, 999–1002. [CrossRef] [PubMed]
14. Longuevergne, L.; Scanlon, B.R.; Wilson, C.R. GRACE Hydrological estimates for small basins: Evaluating processing approaches on the High Plains Aquifer, USA. *Water Resour. Res.* **2010**, *46*, W1151. [CrossRef]
15. Scanlon, B.R.; Longuevergne, L.; Long, D. Ground referencing GRACE satellite estimates of groundwater storage changes in the California Central Valley, USA. *Water Resour. Res.* **2012**, *48*, W04520. [CrossRef]
16. Tran, H.; Zhang, J.; O'Neill, M.M.; Ryken, A.; Condon, L.E.; Maxwell, R.M. A hydrological simulation dataset of the Upper Colorado River Basin from 1983 to 2019. *Sci. Data* **2022**, *9*, 16. [CrossRef]
17. Garćia-Garćia, D.; Ummenhofer, C.C.; Zlotnicki, V. Australian water mass variations from GRACE data linked to Indo-Pacific climate variability. *Remote Sens. Environ.* **2011**, *115*, 2175–2183. [CrossRef]
18. Forootan, E.; Awange, J.; Kusche, J.; Heck, B.; Eicker, A. Independent patterns of water mass anomalies over Australia from satellite data and models. *Remote Sens. Environ.* **2012**, *124*, 427–443. [CrossRef]
19. Voss, K.A.; Famiglietti, J.S.; Lo, M.; De Linage, C.; Rodell, M.; Swenson, S.C. Groundwater depletion in the Middle East from GRACE with implications for transboundary water management in the Tigris-Euphrates-Western Iran region. *Water Resour. Res.* **2013**, *49*, 904–914. [CrossRef]
20. Moiwo, J.P.; Tao, F.; Lu, W. Analysis of satellite-based and in situ hydro-climatic data depicts water storage depletion in North China Region. *Hydrol. Process.* **2013**, *27*, 1011–1020. [CrossRef]
21. Tang, Q.; Zhang, X.; Tang, Y. Anthropogenic impacts on mass change in North China. *Geophys. Res. Lett.* **2013**, *40*, 3924–3928. [CrossRef]
22. Huang, Z.; Pan, Y.; Gong, H.; Yeh, P.J.F.; Li, X.; Zhou, D.; Zhao, W. Subregional-scale groundwater depletion detected by GRACE for both shallow and deep aquifers in North China Plain. *Geophys. Res. Lett.* **2015**, *42*, 1791–1799. [CrossRef]
23. Pan, Y.; Zhang, C.; Gong, H.; Yeh, P.J.F.; Shen, Y.; Guo, Y.; Huang, Z.; Li, X. Detection of human-induced evapotranspiration using GRACE satellite observations in the Haihe River basin of China. *Geophys. Res. Lett.* **2017**, *44*, 190–199. [CrossRef]
24. Zhang, X.; Ren, L.; Feng, W. Comparison of the shallow groundwater storage change estimated by a distributed hydrological model and GRACE satellite gravimetry in a well-irrigated plain of the Haihe River basin, China. *J. Hydrol.* **2022**, *610*, 127799. [CrossRef]
25. Sun, A.; Green, R.; Swenson, S.; Rodell, M. Toward calibration of regional groundwater models using GRACE data. *J. Hydrol.* **2012**, *422*, 1–9. [CrossRef]
26. Iqbal, N.; Hossain, F.; Lee, H.; Akhter, G. Integrated groundwater resource management in Indus basin using satellite gravimetry and physical modeling tools. *Environ. Monit. Assess.* **2017**, *189*, 128. [CrossRef]
27. Swenson, S.; Yeh, P.J.-F.; Wahr, J.; Famiglietti, J. A comparison of terrestrial water storage variations from grace with in situ measurements from Illinois. *Geophys. Res. Lett.* **2006**, *33*, 627–642. [CrossRef]
28. Chen, J.L.; Rodell, M.; Wilson, C.R.; Famiglietti, J.S. Low degree spherical harmonic influences on Gravity Recovery and Climate Experiment (GRACE) water storage estimates. *Geophys. Res. Lett.* **2005**, *32*, L14405. [CrossRef]
29. Abhishek; Kinouchi, T. Synergetic application of GRACE gravity data, global hydrological model, and in-situ observations to quantify water storage dynamics over Peninsular India during 2002–2017. *J. Hydrol.* **2021**, *596*, 126069. [CrossRef]
30. Tangdamrongsub, N.; Steele-Dunne, S.C.; Gunter, B.C.; Ditmar, P.G.; Sutanudjaja, E.H.; Sun, Y.; Xia, T.; Wang, Z. Improving estimates of water resources in a semi-arid region by assimilating GRACE data into the PCR-GLOBWB hydrological model. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 2053–2074. [CrossRef]
31. Niu, J.; Shen, C.; Li, S.-G.; Phanikumar, M.S. Quantifying storage changes in regional Great Lakes watersheds using a coupled subsurface-land surface process model and GRACE, MODIS products. *Water Resour. Res.* **2014**, *50*, 7359–7377. [CrossRef]
32. Gong, H.; Pan, Y.; Zheng, L.; Li, X.; Zhu, L.; Zhang, C.; Huang, Z.; Li, Z.; Wang, H.; Zhou, C. Long-term groundwater storage changes and land subsidence development in the North China Plain (1971–2015). *Hydrogeol. J.* **2018**, *26*, 1417–1427. [CrossRef]
33. Alshehri, F.; Mohamed, A. Analysis of groundwater storage fluctuations using GRACE and remote sensing data in Wadi As-Sirhan, Northern Saudi Arabia. *Water* **2023**, *15*, 282. [CrossRef]
34. Mohamed, A.; Abdelrady, A.; Alarifi, S.S.; Othman, A. Geophysical and remote sensing assessment of Chad's groundwater resources. *Remote Sens.* **2023**, *15*, 560. [CrossRef]
35. Famiglietti, J.S.; Rodell, M. Water in the balance. *Science* **2013**, *340*, 1300–1301. [CrossRef]
36. Fowler, H.J.; Blenkinsop, S.; Tebaldi, C. Linking climate change modelling to impacts studies: Recent advances in downscaling techniques for hydrological modelling. *Int. J. Climatol.* **2007**, *27*, 1547–1578. [CrossRef]
37. Wilby, R.L.; Wigley, T.M.L.; Conway, D.; Jones, P.D.; Hewitson, B.C.; Main, J.; Wilks, D.S. Statistical downscaling of general circulation model output: A comparison of methods. *Water Resour. Res.* **1998**, *34*, 2995–3008. [CrossRef]
38. Ning, S.W.; Ishidaira, H.; Wang, J. Statistical downscaling of GRACE-derived terrestrial water storage using satellite and GLDAS products. *J. Jpn. Soc. Civ. Eng.* **2014**, *70*, 133–138. [CrossRef]
39. Yin, W.; Hu, L.; Zhang, M.; Wang, J.; Han, S.C. Statistical downscaling of GRACE-derived groundwater storage using ET data in the North China Plain. *J. Geophys. Res. Atmos.* **2018**, *123*, 5973–5987. [CrossRef]
40. Sun, A.Y. Predicting groundwater level changes using GRACE data. *Water Resour. Res.* **2013**, *49*, 5900–5912. [CrossRef]

41. Humphrey, V.; Gudmundsson, L. GRACE-REC: A reconstruction of climate-driven water storage changes over the last century. *Earth Syst. Sci. Data* **2019**, *11*, 1153–1170. [CrossRef]
42. Eicker, A.; Schumacher, M.; Kusche, J.; Döll, P.; Schmied, H.M. Calibration/Data assimilation approach for integrating GRACE data into the WaterGAP Global Hydrology Model (WGHM) using an Ensemble Kalman Filter: First results. *Surv. Geophys.* **2014**, *35*, 1285–1309. [CrossRef]
43. Zhong, D.; Wang, S.; Li, J. A self-calibration variance-component model for spatial downscaling of GRACE observations using land surface model outputs. *Water Resour. Res.* **2021**, *57*, e2020WR028944. [CrossRef]
44. Tangdamrongsub, N.; Han, S.C.; Tian, S.; Müller Schmied, H.; Sutanudjaja, E.H.; Ran, J.; Feng, W. Evaluation of groundwater storage variations estimated from GRACE data assimilation and state-of-the-art land surface models in Australia and the North China Plain. *Remote Sens.* **2018**, *10*, 483. [CrossRef]
45. Schumacher, M.; Forootan, E.; van Dijk, A.I.; Schmied, H.M.; Crosbie, R.S.; Kusche, J.; Döll, P. Improving drought simulations within the Murray-Darling Basin by combined calibration/assimilation of GRACE data into the WaterGAP Global Hydrology Model. *Remote Sens. Environ.* **2018**, *204*, 212–228. [CrossRef]
46. Tian, S.; Van Dijk, A.I.; Tregoning, P.; Renzullo, L.J. Forecasting dryland vegetation condition months in advance through satellite data assimilation. *Nat. Commun.* **2019**, *10*, 469. [CrossRef] [PubMed]
47. Cao, G.; Zheng, C.; Scanlon, B.R.; Liu, J.; Li, W. Use of flow modeling to assess sustainability of groundwater resources in the North China Plain. *Water Resour. Res.* **2013**, *49*, 159–175. [CrossRef]
48. Yang, C.; Li, H.; Fang, Y.; Cui, C.; Wang, T.; Zheng, C.; Leung, L.R.; Maxwell, R.M.; Zhang, Y.; Yang, X. Effects of groundwater pumping on ground surface temperature: A regional modeling study in the North China Plain. *J. Geophys. Res.-Atmos.* **2019**, *125*, e2019JD031764. [CrossRef]
49. Zhang, C.; Duan, Q.; Yeh, P.J.; Pan, Y.; Gong, H.; Gong, W.; Di, Z.; Lei, X.; Liao, W.; Huang, Z.; et al. The effectiveness of the South-to-North Water Diversion Middle Route Project on water delivery and groundwater recovery in North China Plain. *Water Resour. Res.* **2020**, *56*, e2019WR026759. [CrossRef]
50. Zhang, C.; Duan, Q.; Yeh, P.J.-F.; Pan, Y.; Gong, H.; Moradkhani, H.; Gong, W.; Lei, X.; Liao, W.; Xu, L.; et al. Sub-regional groundwater storage recovery in North China Plain after the South-to-North water diversion project. *J. Hydrol.* **2021**, *597*, 126156. [CrossRef]
51. Wang, L.; Jia, B.; Xie, Z.; Wang, B.; Liu, S.; Li, R.; Liu, B.; Wang, Y.; Chen, S. Impact of groundwater extraction on hydrological process over the Beijing-Tianjin-Hebei region, China. *J. Hydrol.* **2022**, *609*, 127689. [CrossRef]
52. Li, X.; Ye, S.; Wei, A.; Zhou, P.; Wang, L. Modelling the response of shallow groundwater levels to combined climate and water-diversion scenarios in Beijing-Tianjin-Hebei Plain, China. *Hydrogeol. J.* **2017**, *25*, 1733–1744. [CrossRef]
53. Long, D.; Yang, W.; Scanlon, B.R.; Zhao, J.; Liu, D.; Burek, P.; Pan, Y.; You, L.; Wada, Y. South-to-North Water Diversion stabilizing Beijing's groundwater levels. *Nat. Commun.* **2020**, *11*, 3665. [CrossRef] [PubMed]
54. Watkins, M.M.; Wiese, D.N.; Yuan, D.N.; Boening, C.; Landerer, F.W. Improved methods for observing Earth's time variable mass distribution with GRACE using spherical cap mascons. *J. Geophys. Res. Solid Earth* **2015**, *120*, 2648–2671. [CrossRef]
55. Wiese, D.N.; Landerer, F.W.; Watkins, M.M. Quantifying and reducing leakage errors in the JPL RL05M GRACE mascon solution. *Water Resour. Res.* **2016**, *52*, 7490–7502. [CrossRef]
56. Save, H.; Bettadpur, S.; Tapley, B.D. High-resolution CSR GRACE RL05 mascons. *J. Geophys. Res. Solid Earth* **2016**, *121*, 7547–7569. [CrossRef]
57. Wang, J.; Song, C.; Reager, J.T.; Yao, F.; Famiglietti, J.S.; Sheng, Y.; MacDonald, G.M.; Brun, F.; Schmied, H.M.; Marston, R.A.; et al. Recent global decline in endorheic basin water storages. *Nat. Geosci.* **2018**, *11*, 926–932. [CrossRef]
58. Peng, S.; Ding, Y.; Liu, W.; Li, Z. 1 km monthly temperature and precipitation dataset for China from 1901 to 2017. *Earth Syst. Sci. Data* **2019**, *11*, 1931–1946. [CrossRef]
59. Shao, J.; Cui, Y.; Hao, Q.; Han, Z.; Cheng, T. Study on the estimation of groundwater withdrawals based on groundwater flow modeling and its application in the North China Plain. *J. Earth Sci.* **2014**, *25*, 1033–1042. [CrossRef]
60. Cao, G.; Han, D.; Song, X. Evaluating actual evapotranspiration and impacts of groundwater storage change in the North China Plain. *Hydrol. Process.* **2014**, *28*, 1797–1808. [CrossRef]
61. Zhang, X.; Ren, L.; Kong, X. Estimating spatiotemporal variability and sustainability of shallow groundwater in a well-irrigated plain of the Haihe River basin using SWAT model. *J. Hydrol.* **2016**, *541*, 1221–1240. [CrossRef]
62. Moiwo, J.P.; Yang, Y.; Han, S.; Lu, W.; Yan, N.; Wu, B. A method for estimating soil moisture storage in regions under water stress and storage depletion: A case study of Hai River Basin, North China. *Hydrol. Process.* **2011**, *25*, 2275–2287. [CrossRef]
63. Sun, J.; Hu, L.; Liu, X.; Sun, K. Enhanced understanding of groundwater storage changes under the influence of river basin governance using GRACE data and downscaling model. *Remote Sens.* **2022**, *14*, 4719. [CrossRef]
64. Zhang, Z.; Li, L. *Groundwater Resources of China*; China Cartographic Publishing House: Beijing, China, 2005. (In Chinese)
65. Wang, C.; Duan, Q.; Tong, C.H.; Di, Z.; Gong, W. A GUI platform for uncertainty quantification of complex dynamical models. *Environ. Model Softw.* **2016**, *76*, 1–12. [CrossRef]
66. Morris, M.D. Factorial sampling plans for preliminary computational experiments. *Technometrics* **1991**, *33*, 161–174. [CrossRef]
67. Zhao, Q.; Zhang, B.; Yao, Y.; Wu, W.; Meng, G.; Chen, Q. Geodetic and hydrological measurements reveal the recent acceleration of groundwater depletion in North China Plain. *J. Hydrol.* **2019**, *575*, 1065–1072. [CrossRef]

68. Liu, B.; Zou, X.; Yi, S.; Sneeuw, N.; Cai, J.; Li, J. Identifying and separating climate-and human-driven water storage anomalies using GRACE satellite data. *Remote Sens. Environ.* **2021**, *263*, 112559. [CrossRef]
69. Beijing Water Authority. *Beijing Water Resource Statistics Year Book 2016*; Beijing Water Authority: Beijing, China, 2016. (In Chinese)
70. China Geology Survey. *Handbook of Hydrogeology*, 2nd ed.; Geological Publishing House: Beijing, China, 2012. (In Chinese)
71. Girotto, M.; Reichle, R.H.; Rodell, M.; Liu, Q.; Mahanama, S.; De Lannoy, G.J.M. Multi-sensor assimilation of SMOS brightness temperature and GRACE terrestrial water storage observations for soil moisture and shallow groundwater estimation. *Remote Sens. Environ.* **2019**, *227*, 12–27. [CrossRef]
72. Liu, J.; Jiang, L.; Zhang, X.; Druce, D.; Kittel, C.M.; Tøttrup, C.; Bauer-Gottwein, P. Impacts of water resources management on land water storage in the North China Plain: Insights from multi-mission earth observations. *J. Hydrol.* **2021**, *603*, 126933. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.