A new GRACE downscaling approach for deriving high-resolution groundwater storage changes using ground-based scaling factors

Huixiang Li[1], Yun Pan[1], Zhiyong Huang[2], Chong Zhang[3], Li Xu[4], Huili Gong[1], and James S. Famiglietti[4]

[1]Capital Normal University
[2]Changsha University of Science & Technology
[3]Beijing Normal University
[4]University of Saskatchewan

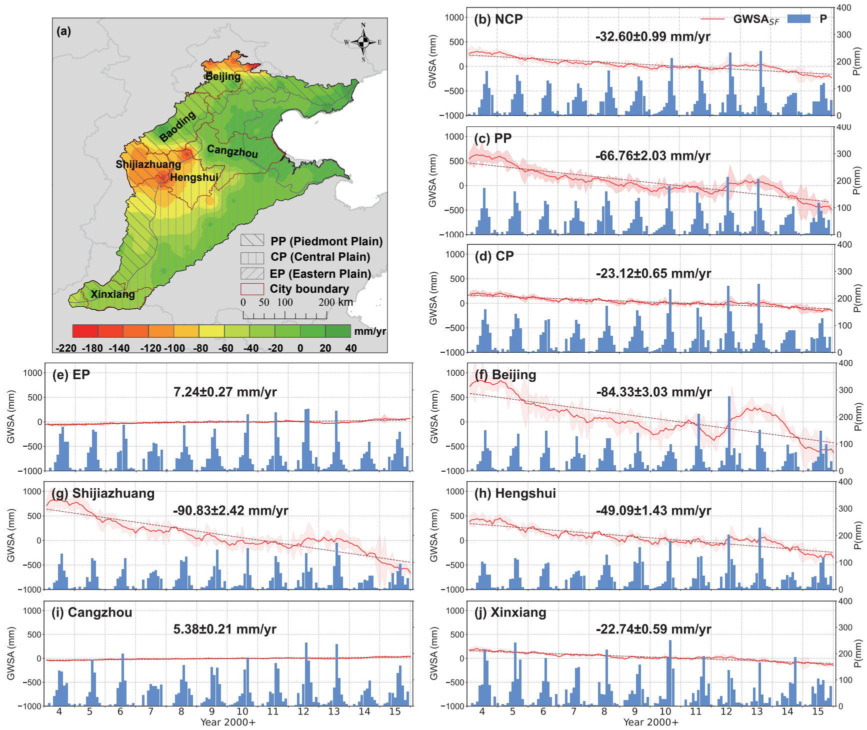
May 8, 2023

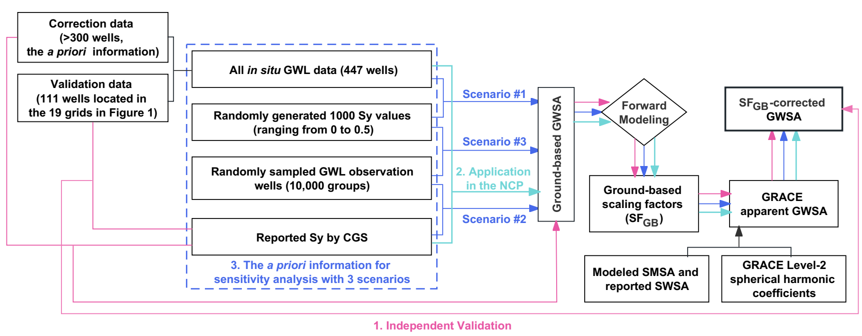
Abstract

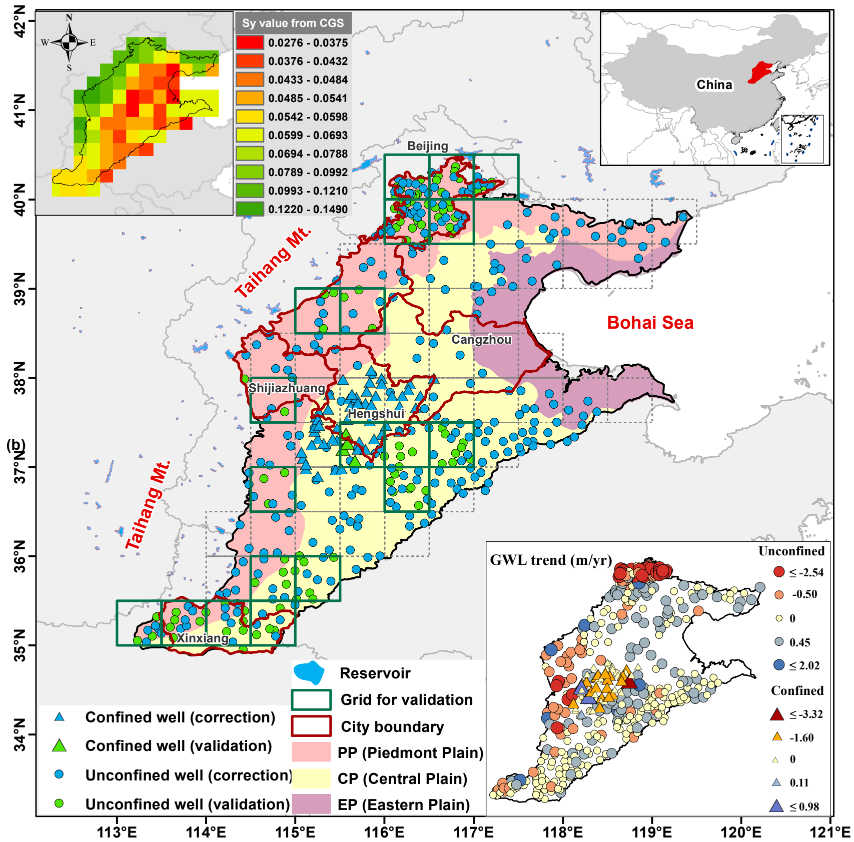
To compensate for the intrinsic coarse spatial resolution of groundwater storage (GWS) anomalies (GWSA) from the Gravity Recovery and Climate Experiment (GRACE) satellites and make better use of current dense in situ groundwater-level data in some regions, a new statistical downscaling method was proposed to derive high-resolution GRACE GWS changes. A ground-based scaling factor (SFGB) method was proposed to downscale GRACE GWS changes that were corrected using gridded scaling factors estimated from ground-based GWS changes through forward modeling. The proposed method was applied in the North China Plain (NCP), where many observation wells and consistently measured specific yield are available. Importantly, the sensitivity of the proposed method was explored considering the uncertainties of in situ GWS changes due to variable specific yield and/or number of observation wells. Independent validation shows that SFGB can effectively recover GRACE GWSA at the 0.5º grid scale (r = 0.81, root mean square error = 40.51 mm/yr). The SFGB-corrected GWSA in the NCP was -32.60{plus minus}0.99 mm/yr (-4.6{plus minus}0.14 km3/yr) during 2004-2015, showing contrasting GWS trends in the piedmont west (loss) and the coastal east (gains). Uncertainties in SFGB-corrected GWSA arising from specific yield, groundwater-level, and both can be reduced by 90%, 65%, and 84%, respectively relative to ground-based GWSA. This study highlights the potential value of jointly using GRACE and in situ observation data to improve the accuracy of GRACE-derived GWSA at smaller scales. The new downscaling method and the improved groundwater storage change estimates would facilitate better groundwater management.

Hosted file

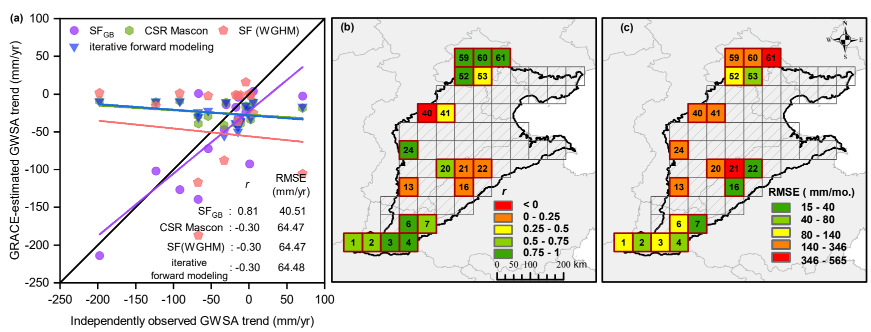
`962014_0_art_file_10931440_rtpdcx.docx` available at https://authorea.com/users/615116/articles/641684-a-new-grace-downscaling-approach-for-deriving-high-resolution-groundwater-storage-changes-using-ground-based-scaling-factors


(a)
Beijing
Baoding
Cangzhou
Shijiazhuang
Hengshui
Xinxiang
PP (Piedmont Plain)
CP (Central Plain)
EP (Eastern Plain)
City boundary
0 50 100 200 km
-220 -180 -140 -120 -100 -80 -60 -40 -20 0 20 40
mm/yr
(b) NCP
-32.60±0.99 mm/yr
GWSA$_{SF}$
P
(c) PP
-66.76±2.03 mm/yr
(d) CP
-23.12±0.65 mm/yr
(e) EP
7.24±0.27 mm/yr
(f) Beijing
-84.33±3.03 mm/yr
(g) Shijiazhuang
-90.83±2.42 mm/yr
(h) Hengshui
-49.09±1.43 mm/yr
(i) Cangzhou
5.38±0.21 mm/yr
(j) Xinxiang
-22.74±0.59 mm/yr
GWSA (mm)
P(mm)
Year 2000+

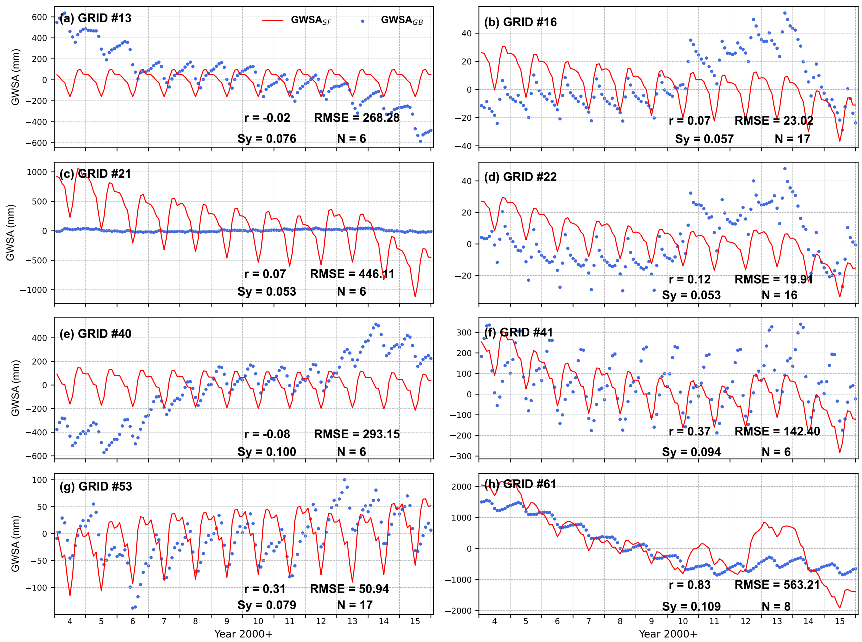

Correction data (>300 wells, the *a priori* information)
Validation data (111 wells located in the 19 grids in Figure 1)
All *in situ* GWL data (447 wells)
Randomly generated 1000 Sy values (ranging from 0 to 0.5)
Randomly sampled GWL observation wells (10,000 groups)
Reported Sy by CGS
3. The *a priori* information for sensitivity analysis with 3 scenarios
Scenario #1
Scenario #3
2. Application in the NCP
Scenario #2
Ground-based GWSA
Forward Modeling
Ground-based scaling factors (SF_{GB})
SF_{GB}-corrected GWSA
GRACE apparent GWSA
Modeled SMSA and reported SWSA
GRACE Level-2 spherical harmonic coefficients
1. Independent Validation

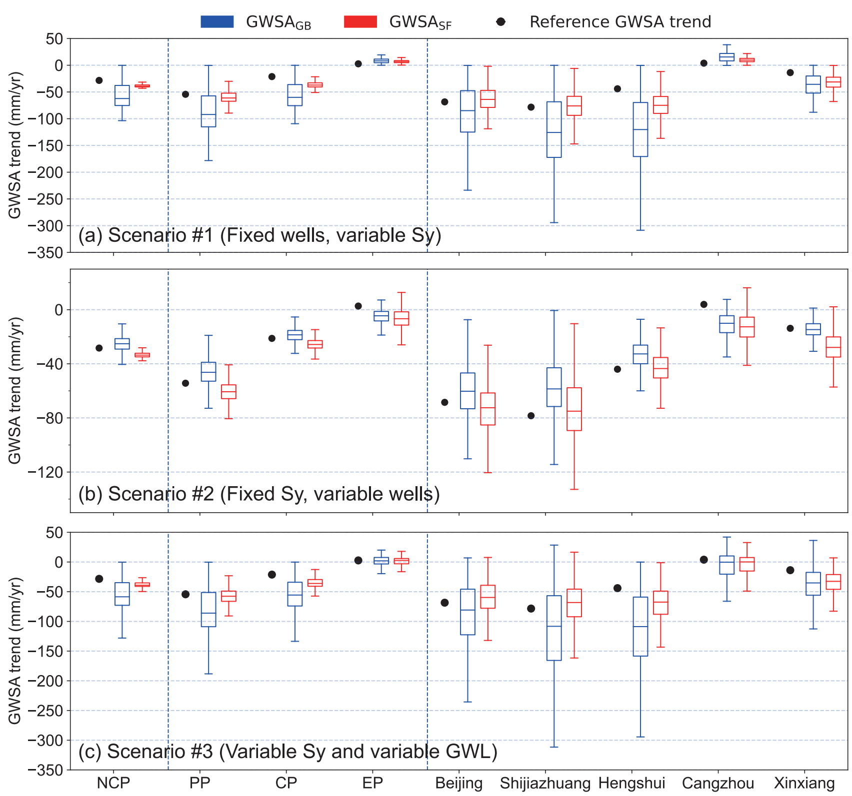


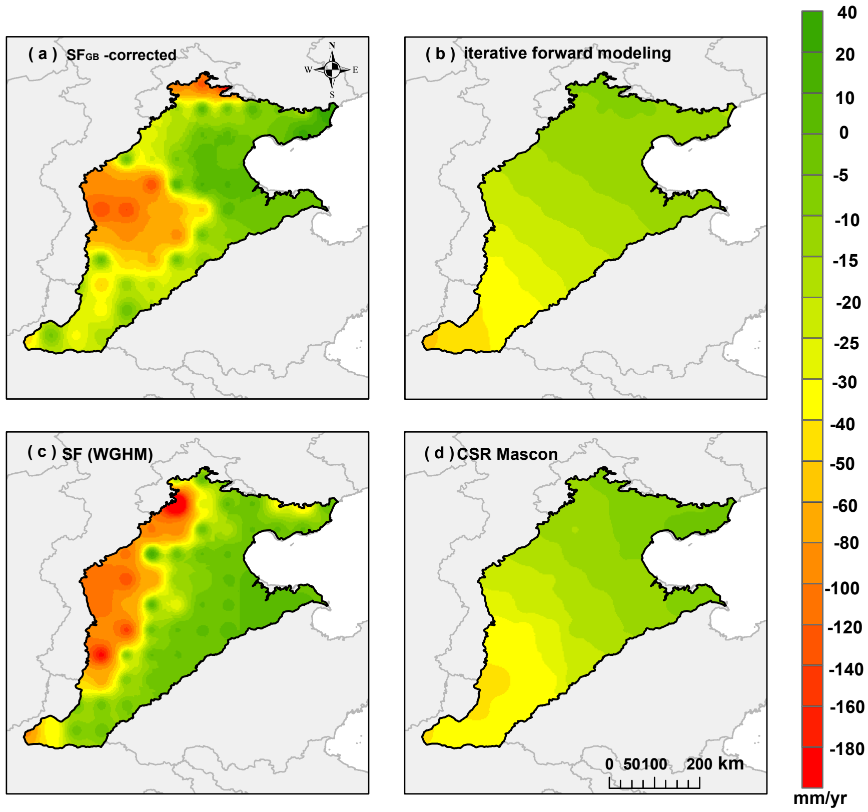
Sy value from CGS
0.0276 - 0.0375
0.0376 - 0.0432
0.0433 - 0.0484
0.0485 - 0.0541
0.0542 - 0.0598
0.0599 - 0.0693
0.0694 - 0.0788
0.0789 - 0.0992
0.0993 - 0.1210
0.1220 - 0.1490
China
Beijing
Taihang Mt.
Bohai Sea
Cangzhou
Shijiazhuang
Hengshui
Xinxiang
(b)
GWL trend (m/yr)
Unconfined
≤ -2.54
-0.50
0
0.45
≤ 2.02
Confined
≤ -3.32
-1.60
0
0.11
≤ 0.98
Reservoir
Grid for validation
City boundary
PP (Piedmont Plain)
CP (Central Plain)
EP (Eastern Plain)
Confined well (correction)
Confined well (validation)
Unconfined well (correction)
Unconfined well (validation)
42°N
41°N
40°N
39°N
38°N
37°N
36°N
35°N
34°N
113°E
114°E
115°E
116°E
117°E
118°E
119°E
120°E
121°E



(a)
SF_GB
CSR Mascon
SF (WGHM)
iterative forward modeling
GRACE-estimated GWSA trend (mm/yr)
Independently observed GWSA trend (mm/yr)
r
RMSE (mm/yr)
SF_GB : 0.81 40.51
CSR Mascon : -0.30 64.47
SF(WGHM) : -0.30 64.47
iterative forward modeling : -0.30 64.48
(b)
< 0
0 - 0.25
0.25 - 0.5
0.5 - 0.75
0.75 - 1
0 50 100 200 km
(c)
RMSE (mm/mo.)
15 - 40
40 - 80
80 - 140
140 - 346
346 - 565


(a) GRID #13
GWSA$_{SF}$
GWSA$_{GB}$
r = -0.02 RMSE = 268.28
Sy = 0.076 N = 6
(b) GRID #16
r = 0.07 RMSE = 23.02
Sy = 0.057 N = 17
(c) GRID #21
r = 0.07 RMSE = 446.11
Sy = 0.053 N = 6
(d) GRID #22
r = 0.12 RMSE = 19.91
Sy = 0.053 N = 16
(e) GRID #40
r = -0.08 RMSE = 293.15
Sy = 0.100 N = 6
(f) GRID #41
r = 0.37 RMSE = 142.40
Sy = 0.094 N = 6
(g) GRID #53
r = 0.31 RMSE = 50.94
Sy = 0.079 N = 17
(h) GRID #61
r = 0.83 RMSE = 563.21
Sy = 0.109 N = 8
GWSA (mm)
Year 2000+


GWSA$_{GB}$
GWSA$_{SF}$
Reference GWSA trend
GWSA trend (mm/yr)
(a) Scenario #1 (Fixed wells, variable Sy)
(b) Scenario #2 (Fixed Sy, variable wells)
(c) Scenario #3 (Variable Sy and variable GWL)
NCP
PP
CP
EP
Beijing
Shijiazhuang
Hengshui
Cangzhou
Xinxiang



Hosted file

962014_0_supp_10931944_rt09w3.docx available at https://authorea.com/users/615116/articles/641684-a-new-grace-downscaling-approach-for-deriving-high-resolution-groundwater-storage-changes-using-ground-based-scaling-factors

A new GRACE downscaling approach for deriving high-resolution groundwater storage changes using ground-based scaling factors

Huixiang Li[1,2,3], Yun Pan[1,2,3*], Zhiyong Huang[4,5,6*], Chong Zhang[1,2,3], Li Xu[7,8], Huili Gong[1,2,3], James S. Famiglietti[7,8, 9]

[1]Beijing Laboratory of Water Resources Security, Capital Normal University, Beijing, China

[2]MOE Key Laboratory of Mechanism, Prevention, and Mitigation of Land Subsidence, Capital Normal University, Beijing, China

[3]Hebei Cangzhou Groundwater and Land Subsidence National Observation and Research Station, Cangzhou, China

[4]School of Hydraulic and Environmental Engineering, Changsha University of Science & Technology, Changsha, China

[5]Key Laboratory of Water-Sediment Sciences and Water Disaster Prevention of Hunan Province, Changsha, China

[6]Key Laboratory of Dongting Lake Aquatic Eco-Environmental Control and Restoration of Hunan Province, Changsha, China

[7]Global Institute for Water Security, University of Saskatchewan, Saskatoon, Canada

[8]School of Environment and Sustainability, University of Saskatchewan, Saskatoon, Canada

[9]School of Sustainability, Arizona State University, Tempe, USA

*Corresponding authors:

Yun Pan (pan@cnu.edu.cn); Zhiyong Huang (huangzy9084@126.com)

Key points:

1. A new statistical downscaling method to improve GRACE groundwater storage change estimates was proposed and proved to be effective.

2. Leakage-corrected GRACE groundwater storage anomalies in North China Plain showed losses (gains) in the Piedmont Plain (Coastal Plain).

3. Uncertainty in leakage-corrected groundwater storage anomaly arising from specific yield and/or groundwater level can be reduced by 65-90%.

Abstract: To compensate for the intrinsic coarse spatial resolution of groundwater storage (GWS) anomalies (GWSA) from the Gravity Recovery and Climate Experiment (GRACE) satellites and make better use of current dense *in situ* groundwater-level data in some regions, a new statistical downscaling method was proposed to derive high-resolution GRACE GWS changes. A ground-based scaling factor (SF_{GB}) method was proposed to downscale GRACE GWS changes that were corrected using gridded scaling factors estimated from ground-based GWS changes through forward modeling. The proposed method was applied in the North China Plain (NCP), where many observation wells and consistently measured specific yield are available. Importantly, the sensitivity of the proposed method was explored considering the uncertainties of *in situ* GWS changes due to variable specific yield and/or number of observation wells. Independent validation shows that SF_{GB} can effectively recover GRACE GWSA at the 0.5º grid scale (r = 0.81, root mean square error = 40.51 mm/yr). The SF_{GB}-corrected GWSA in the NCP was -32.60±0.99 mm/yr (-4.6±0.14 km^3/yr) during 2004-2015, showing contrasting GWS trends in the piedmont west (loss) and the coastal east (gains). Uncertainties in SF_{GB}-corrected GWSA arising from specific yield, groundwater-level, and both can be reduced by 90%, 65%, and 84%, respectively relative to ground-based GWSA. This study highlights the potential value of jointly using GRACE and *in situ* observation data to improve the accuracy of GRACE-derived GWSA at smaller scales. The new downscaling method and the improved groundwater storage change estimates would facilitate better groundwater management.

Keywords: GRACE; *In situ* observations; Statistical downscaling; Groundwater level;

Specific yield; Scaling factor.

1. Introduction

Groundwater is a critical resource. It is important to understand how its quantity, *i.e.*, groundwater storage (GWS) changes at different spatial and temporal scales. *In situ* groundwater-level (GWL) observations and Gravity Recovery and Climate Experiment (GRACE) satellite measurements are widely used to monitor GWS changes. However, both of them have strengths and limitations (Alley and Konikow 2015; Famiglietti et al. 2015; Famiglietti and Rodell 2013; Scanlon et al. 2012). *In situ* GWL observations can be used to estimate GWS changes by multiplying a specific yield (Sy) for unconfined aquifers or storage coefficient for confined aquifers. However, the actual Sy or storage coefficient is not always spatially available or is difficult to be accurately estimated in some cases (Gehman et al. 2022; Lv et al. 2021; Rodell et al. 2007). A regional mean Sy referenced from soil lithology was thus often used in previous studies (Bhanja et al. 2016; Famiglietti et al. 2011; Leblanc et al. 2009). Furthermore, the observation wells can be insufficient in many regions, resulting in large uncertainties in *in-situ*-measured GWS changes (Chen et al. 2016; Hachborn et al. 2017; Henry et al. 2011).

GRACE is capable of capturing GWS changes independent of *in situ* information (Rodell et al. 2009). GRACE data has been successfully applied worldwide (Chandanpurkar et al. 2021; Huang et al. 2015; Panda and Wahr 2016; Reager and Famiglietti 2009; Richey et al. 2015; Shamsudduha et al. 2012; Strassberg et al. 2009; Syed et al. 2008; Syed et al. 2009; Xiang et al. 2016). The most widely used GRACE solutions include the level-2 spherical harmonic (SH) solutions and level-3 mascon solutions. Mascon solutions can be directly used to estimate terrestrial water storage anomaly (TWSA) time series, but mascons are intrinsic global solutions and are not designed for a specific region (Zhang et al. 2019). SH solutions need complex processing and further signal leakage correction to reduce noise in higher degree coefficients. However, by employing various correction methods, SH data can be applied at sub-regional scales below the GRACE footprint. The iterative forward modeling and scaling factor are two commonly used approaches for this data processing (Chen et al. 2009; Chen et al. 2015; Landerer and Swenson 2012; Longuevergne et al. 2010). However, scaling factors may be sensitive to models used, because they are derived from the *a priori* information which is mostly obtained from model-simulated TWS anomalies (*i.e.*, TWSA) or GWS anomalies (GWSA) (Huang et al. 2019; Landerer and Swenson 2012; Liu and Zou 2019). Many existing global land surface models such as those in the Global Land Data Assimilation System (GLDAS) and global hydrological models such as WaterGAP (WGHM) have a common lack of groundwater modules or many of them are not well-suited to represent local human activities. It is difficult to invert the spatial pattern of mass changes in the study area by using the iterative forward modeling method. Yi et al. (2016) and Vishwakarma et al. 2017 proposed a multi-basin inversion method and data-driven approach respectively, but these methods also had difficulties in inverting the spatial distribution of signal variations.

Based on the limitations of the above correction methods, we attempt to improve the resolution of GRACE by using a new downscaling approach. Downscaling is widely used in the fields of remote sensing, climate, and hydrology (Atkinson 2013; Peng et al. 2017; Quintana Seguí et al. 2010; Saikrishna et al. 2022; Xu et al. 2019). Downscaling methods can be categorized into two types: dynamical downscaling and statistical downscaling. Low-resolution data patterns are nested inside high-resolution data patterns during dynamic downscaling (Saikrishna et al. 2022; Adachi and Tomita 2020; Brown et al. 2008). Because a model is based on physical principles, the physical interpretation of the dynamically downscaled results is simplified. However, computational intensity limits the application of this time-consuming method (Adachi and Tomita 2020; Jyolsna et al. 2021). Statistical downscaling involves establishing a statistical relationship between small-scale observation data and large-scale data (Tang et al. 2016). For example, in the statistical downscaling of soil moisture (Peng et al. 2017), multi-source satellite data with varying resolutions, geographic information data related to soil moisture, and model data can all be used for statistical downscaling. Therefore, many kinds of data can be used as variables in statistical downscaling. Common methods of statistical downscaling include multiple regression, machine learning, *etc*. Compared with dynamical downscaling, statistical downscaling is simpler and requires much less computational time (Chen et al. 2010).

Statistical downscaling methods are also widely used in GRACE downscaling (Pulla et al. 2023; Arshad et al. 2022; Yin et al. 2022). Most studies construct correlations with GRACE GWSA using data related to changes in GWS, such as precipitation, soil moisture, evapotranspiration, and soil lithology information (Pulla et al. 2023; Chen et al. 2019).

However, it is difficult for these downscaling applications to reflect changes in GWS in areas with high levels of human activity and groundwater extraction (Miro and Famiglietti 2018; Sahour et al. 2020; Yin et al. 2022; Zhang et al. 2019). Meanwhile, statistical downscaling applications that use other variable data like evapotranspiration to establish statistical relationships with GRACE data (e.g., Yin et al. 2018) cannot be applied to regions where there is no strong correlation between GWS and evapotranspiration. There have been few studies on GRACE downscaling that use *in situ* GWL data as variables, and most studies use *in situ* data as prediction targets and validation data (Seyoum et al. 2019; Zhang et al. 2021).

In situ observation data are not only often used in statistical downscaling as prediction target data and validation data (Liu et al. 2020; Xu et al. 2020; Pulla et al. 2023), but are also input into models as variables (Duan and Bastiaanssen 2013; Hunink et al. 2014; López López et al. 2018; Samadi et al. 2013; Shen et al. 2021; Teng et al. 2014; Xu et al. 2018). A large number of studies have shown that when combined with *in situ* observation data, the results of downscaling studies are better (Duan and Bastiaanssen 2013; López López et al. 2018). Although groundwater monitoring on the ground is not an easy task, many countries or regions have constructed dense *in situ* GWL monitoring networks, such as California Central Valley, India, the Bengal Basin, and North China Plain. Previous studies have demonstrated the capability of GRACE satellites in monitoring GWS changes in those aquifers or regions by using *in situ* GWL measurements as validation data (Shamsudduha et al. 2012c; Bhanja and Mukherjee 2019; Huang et al. 2015; Scanlon et al. 2012b;). However, none of those studies retrieved a high-resolution GWS change map using the *in situ* GWL data and/or

GRACE data. Therefore, a large research opportunity still exists in those aquifers or regions where making better use of the dense *in situ* GWL datasets can compensate for the coarse GRACE data by deriving high-resolution GWS changes through statistical downscaling.

In this study, a new statistical downscaling approach was proposed by jointly using *in situ* GWL observations and GRACE data. Ground-based scaling factors were derived to downscale GRACE GWS changes through forward modeling with *in situ* GWL observations as the *a priori* information. The North China Plain (NCP) was selected as the study area considering the great significance of groundwater resources for agricultural and societal development, and the sufficient GWL observation wells and reported Sy data in the region, which can support a variety of validation and sensitivity analyses. To demonstrate the reliability and effectiveness of the proposed downscaling method, this study was conducted in three steps. As the flowchart in Figure 1 shows that an independent validation was first performed using selected validation data. Then, the proposed method was directly applied in the NCP to improve the GWS change estimates at smaller scales. Finally, a sensitivity analysis was performed to investigate how Sy values and/or the number and location of GWL observation wells influence the results of scaling factors through forward modeling.

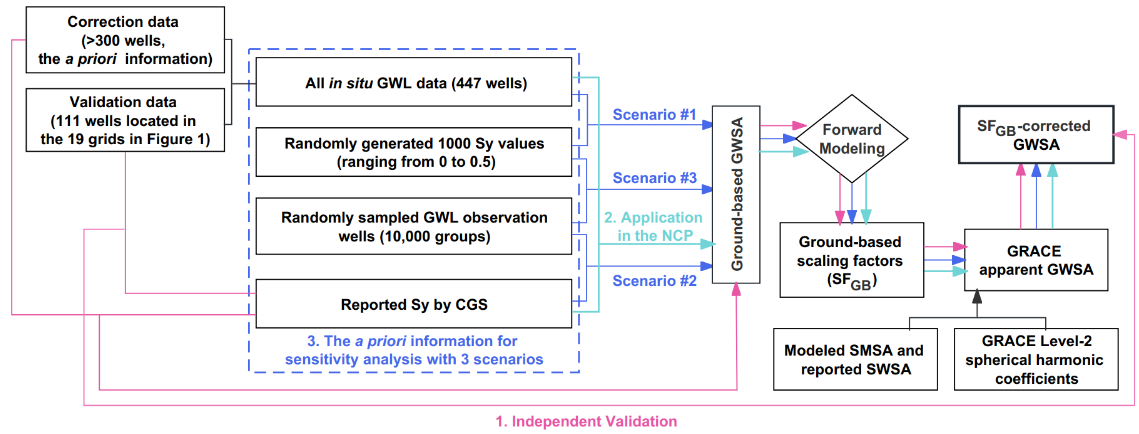


Figure 1. Schematic flowchart of technical and analytical methods used in this study. CGS Sy refers to the widely used Sy of NCP released by the China Geology Survey (Zhang *et al*., 2009). The sensitivity analysis and independent validation were implemented using different *prior* information as that in the application to the NCP.

2. Study area and data

2.1. Study area

A case study was applied to the North China Plain (NCP, Figure 2) at multiple spatial scales, *i.e.*, the whole region (~14×10^4 km²), sub-regions, cities, and the 0.5° grid cells. The sub-regions cover the Piedmont Plain (PP) (~5.4×10^4 km²), the Central Plain (CP) (~6.6×10^4 km²), and the Eastern Plain (EP) (~2.0×10^4 km²). Cities in the plain area include Beijing (0.60×10^4 km²), Shijiazhuang (0.66×10^4 km²), Cangzhou (1.42×10^4km²), Hengshui (0.88×10^4 km²), and Xinxiang (0.67×10^4 km²).

The NCP is characterized by cold-dry winters (December-March) and hot-humid summers (July-September). The annual precipitation in the NCP gradually reduces from 1,200 mm in the southeast to 400 mm in the northwest (Xing et al. 2013). Groundwater in this region is the main source of water supply for agricultural production which contributes ~10% to China's grain production, including ~30% of the total wheat production. Groundwater abstraction had been intensified since the 1970s. Intensive groundwater use started to drop in recent years due to regulations and technical innovation (Gong et al. 2018b). The NCP comprises three distinct hydrogeological settings within the Quaternary aquifer system, including PP (fluvial fans distributed along the Taihang Mountains in the west, and composed of clayey gravel, and medium-coarse sand), CP (alluvial sediments composed of clay, silty clay, and fine-medium sand), and EP (marine and alluvial sediments along the coastal area of Bohai Sea, characterized by fine sand, silt, sandy clay, and silty clay). The grain size of sediment particles and the permeability decrease from the west (PP) to the east (EP). Groundwater flows locally from the top of alluvial fans and regionally from the west to the east (Xing et al. 2013).

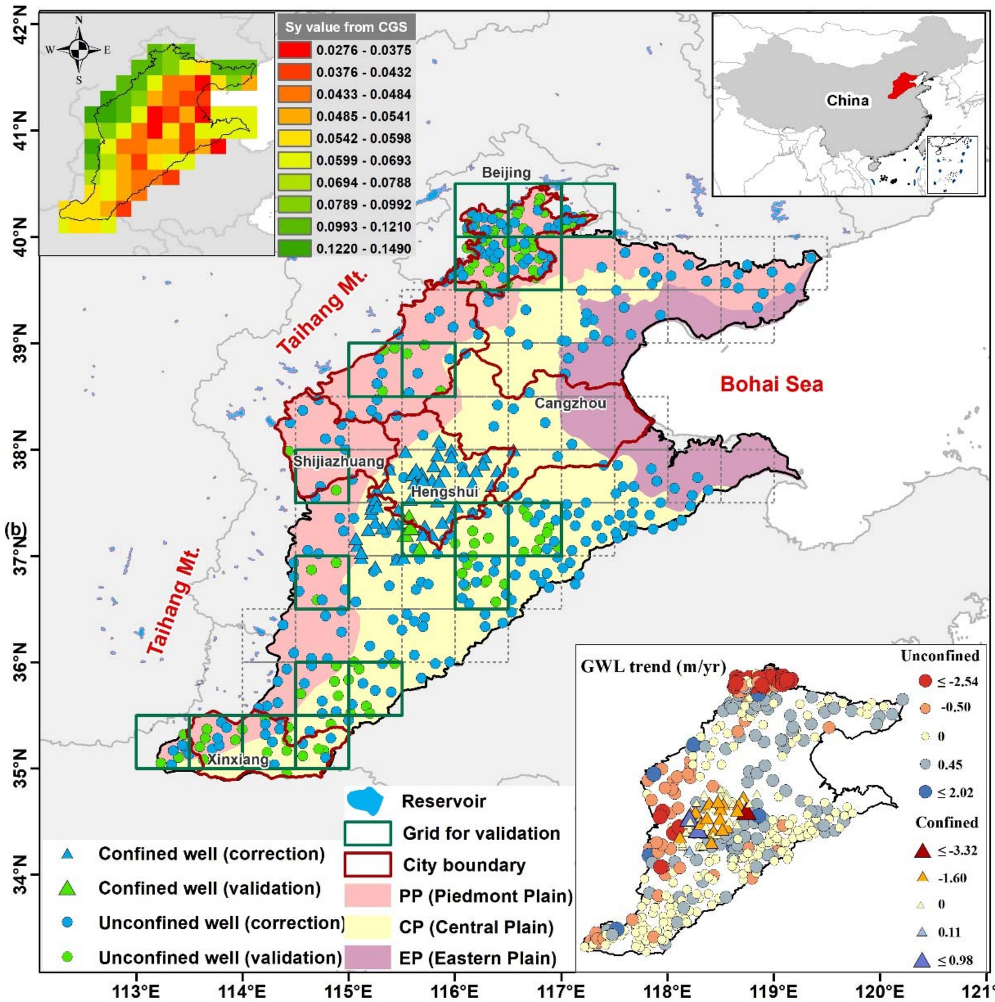


Figure 2. Study area with location of observation wells. The GWL trend map of observation wells was shown on the lower right for a better understanding of GWS changes at local scales. The inserted map on the upper left shows the spatial distribution of Sy reported by CGS (Zhang et al. 2009).

2.2. Data

The GRACE data used for TWSA estimation were the RL06 level-2 spherical harmonic (SH) solutions from the Center for Space Research (CSR) at the University of Texas, covering the period from January 2003 to December 2016. To estimate GWSA from GRACE TWSA, we subtracted soil moisture storage anomalies (SMSA) simulated by GLDAS-1 Community Land Model (CLM) and surface water storage anomalies (SWSA) of 14 major reservoirs from the China Water Annual Report and Beijing Water Resources Bulletin. The CLM SMSA was used considering its best performance among 4 GLDAS model simulations (CLM, VIC, Noah, and Mosaic) with *in situ* observations (Zhang et al., 2021).

GWL data from a total of 447 observation wells were obtained from the Groundwater Level Yearbook (January 2003 to December 2016) compiled by the China Institute of Geological and Environmental Monitoring. The compiled dataset includes 389 wells located at unconfined aquifers and 58 wells at confined aquifers (Figure 2). The Sy data (Figure 2) were obtained from China Geological Survey (CGS), which was empirically estimated by the aquifer soil texture and GWL fluctuations (Zhang et al. 2009). Generally, Sy refers to the aquifer storage coefficient for unconfined aquifers, while the storativity which is much smaller than Sy is usually used for confined aquifers. It is however difficult to identify Sy or storativity for the unconfined or confined aquifers in the NCP because most of the pumping wells abstract water from both unconfined and confined aquifers. In this study, the CGS Sy was used, but attention was paid to its uncertainty that may result from hydrogeological complexities.

3. Methods

3.1. Correcting GRACE GWSA using ground-based scaling factors

The SF_{GB} method in this study used ground-based GWSA (hereafter denoted as $GWSA_{GB}$) as the *a priori* information to derive scaling factors which were then applied to correct the leakage error in GRACE-derived GWSA. Key estimation consists of two parts: disaggregation of apparent GWSA and correction. To estimate TWSA, the CSR SH solutions were truncated at degree and order 60 and filtered by a 300-km Gaussian smoother (Swenson et al. 2003; Swenson and Wahr 2002). To disaggregate GWSA from GRACE TWSA, SMSA, and SWSA were forward-modeled and then subtracted from GRACE TWSA to derive the apparent GWSA (Rodell and Famiglietti 2002). Different from Huang et al. (2015) which used regionally averaged long-term groundwater depletion rate as the *a priori* information, the *a priori* information in this study has a spatial resolution of 0.5°. As a result, this study focuses on the performance of SF_{GB} not only at regional and sub-regional scales but also at city and grid scales.

In this study, we corrected trends and seasonal signals of TWSA individually as they may be weakened during the post-processing of GRACE SH solutions (Landerer and Swenson 2012). In doing so, the *a priori* information (i.e., the $GWSA_{GB}$ time series) was first decomposed into trends and seasonal components using the additive decomposition method (Perfilieva et al. 2013). The retrieved trends and seasonal signals of the *a priori* information were individually forward-modeled to calculate the scaling factors. The scaling factors were then applied to correct the corresponding trends and seasonal components of GRACE apparent GWSA. Next, the corrected trends and seasonal components were aggregated to obtain the SF_{GB}-corrected GWSA (hereafter denoted as $GWSA_{SF}$) time series. Given that the influence of signals outside the studied area should also be considered in the correction, this study calculated the trends of

$GWSA_{GB}$ outside the NCP using the data obtained from the Monthly Report of Groundwater Dynamics released by the Ministry of Water Resources of China. Since there would be a 6-month gap at the beginning and the end of the studied period after the decomposition of the trends and seasonal signals, this study only analyzed the GWSA from January 2004 to December 2015 to ensure the consistency of the calculated GWSA in the time series.

3.2. Independent Validation and application in the NCP

To achieve a more effective validation, the *in situ* GWL observation wells were divided into two independent groups, *i.e.*, the correction data used to produce scaling factors for correcting GRACE GWSA and the validation data used for verification of the method reliability. The division of the two independent data groups avoids the repeated use of GWL data for correction or validation and hence avoids the reliance of validation results on the correction data. The correction and validation data groups are determined as follows. (1) Rescale the entire NCP into 61 grids at 0.5° resolution (see the dashed grids in Figure 2). Most grids encompass several GWL observation wells. (2) Select those grids with densely- and evenly-distributed wells which can be better served as validation grids. A total of 19 grids (see the dark green grids in Figure 2) were selected as validation grids within which the ratio of temporal sampling of GWL data (2003-2016) was greater than 85% (sufficient enough for validation purposes). (3) The validation data were determined by selecting ~50% of wells within each of the 19 grids. Finally, a total of 111 wells were selected as independent validation data, as marked by the green points in Figure 2. (4) The remaining >300 GWL observation wells were divided as correction data.

The correction data were used as *in situ a priori* information to derive scaling factors for correcting GRACE GWSA. The validation data were used to independently evaluate the performance of GRACE GWSA derived from various methods, including $GWSA_{SF}$, and the results obtained from CSR mascon solution, iterative forward modeling, and WGHM-based scaling factors.

After verifying the reliability and effectiveness of the proposed SF_{GB} downscaling method, an application was conducted under a realistic situation when all GWL observation wells and the reported Sy by CGS were considered altogether to derive SF_{GB} for correcting GRACE GWSA. The improved GWS changes were estimated at regional, subregional, city, and grid scales with a comparison to the results and findings in previous studies.

3.3. Numerical experiments for sensitivity analysis of the SF_{GB} downscaling method

As the mass change derived from the scaling factor method is affected by the *a priori* information (Huang et al. 2015; Landerer and Swenson 2012; Long et al. 2015), $GWSA_{SF}$ can be biased by the uncertainties generated by Sy (due to limited knowledge on hydrogeological conditions) and GWL (due to uneven spatial distribution and/or insufficient observation wells). Three scenarios were designed to investigate the effects of individual Sy, GWL, and both of them on uncertainties. The CGS Sy and all (100%) observation wells were used to estimate the referenced GWSA for comparison when conducting numerical experiments for sensitivity analysis. The sensitivities were analyzed using the standard deviations of multiple $GWSA_{SF}$ that result from using different Sy and GWL as the *a priori* information.

Scenario #1: "*Fixed wells, variable Sy*". The number of wells was fixed considering all the *in situ* GWL data from 447 observation wells, but the Sy values were variable by randomly generating 1,000 groups of values ranging from 0 to 0.5. The *a priori* $GWSA_{GB}$ can be estimated using the randomly generated Sy and GWL data.

Scenario #2: "*Fixed Sy, variable wells*". The Sy values were fixed using the data from CGS, but the number of wells was variable by randomly generating 10,000 groups at a 0.5 grid scale. The *a priori* $GWSA_{GB}$ can be estimated using the randomly sampled GWL and the CGS Sy.

Scenario #3: "*Variable Sy, variable GWL*". Sy values are variable by randomly generating 1,000 groups of values ranging from 0 to 0.5, and the number of wells is variable by randomly generating 10,000 groups at a 0.5 grid scale. The *a priori* $GWSA_{GB}$ can be estimated from both randomly generated Sy and GWL.

Notably, when testing a random generation of 1,000 groups of wells, almost identical results were achieved relative to the random generation of 10,000 groups of wells. Therefore, we finally used 1,000 groups of results in both Scenarios #2 and #3.

3.4. Error estimation

In this study, we estimated errors from the GRACE measurements, hydrological model simulations, and *in situ* GWSA. Following Chen et al. (2009), GRACE measurement error was calculated using the residual signals in the Pacific Ocean at the same latitude. The error from model-simulated SMSA was estimated using the standard deviation of the SMSA simulations of four models (i.e., CLM, VIC, MOS, and NOAH) after forward modeling. An error of 10% was specified for $GWSA_{GB}$. The final error of apparent GWSA and $GWSA_{SF}$ was calculated using the error propagation principle of addition and multiplication given by equations (1) and (2):

$$\sigma_{\text{apparent GWSA}} = \sqrt{\sigma_{TWSA}^2 + \sigma_{SMSA}^2} \tag{1}$$

$$\sigma_{GWSA_{SF}} = \sqrt{(\sigma_{SF})^2 \times (apparent\ GWSA)^2 + (\sigma_{apparent\ GWSA})^2 \times (SF)^2} \tag{2}$$

where σ is the estimated uncertainty of corresponding variables, SF is the scaling factor, and the upper hyphen "—" means calculating the derivative.

4. Results and discussion

4.1. Validation of GRACE-derived GWSA against independent measurements

Figure 3a shows the spatial correlations (statistics at 0.5° grid scale) between GRACE-derived GWSAs and independent validation data. Overall, $GWSA_{SF}$ represents a good consistency with independent validation data, with a correlation coefficient of 0.81 and root mean square error (RMSE) of 40.5 mm/yr. However, GWSA trends (Figure S1) derived from CSR mascon, iterative forward modeling, and WGHM-based scaling factor have weak correlations with independent validation data, with a correlation coefficient of -0.30, -0.30, and -0.07, respectively.

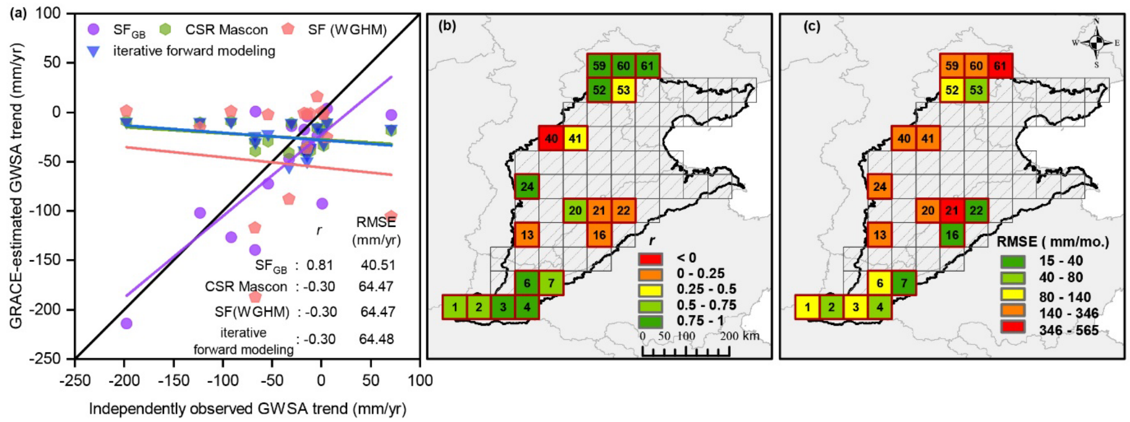


Figure 3. (a) Comparison between GRACE-derived GWSA trends and independent measurements. Four GRACE-estimated GWSA trends are inter-compared, *i.e.*, the SF_{GB}, iterative forward modeling, CSR Mascon solution, and WGHM-based scaling factor. The statistics for correlation coefficient (r), and root mean square error (RMSE) were performed at 0.5° grid scale. (b-c) Grid-scale statistics of correlation coefficients (*r*, b) and RMSE (c) between monthly GWSA corrected by SF_{GB} and the independent measurements.

Figure 3b-c shows the spatial distribution of the statistical metrics between $GWSA_{SF}$ and the independent measurements at the grid scale. Figure 4 plots the monthly time series of $GWSA_{SF}$ with large discrepancies in the grid cells of Figure 3b-c compared to the independent validation dataset. Differences between the two types of data are caused by the mismatch of data caused by several reasons. Firstly, the mismatch can be caused by the uncertainties of *in situ* observations, *e.g.*, insufficient observation wells in the unconfined aquifers (such as GRID #13, #40, #41, and #53) and biased Sy in the confined aquifers (such as GRID #16, #21, and #22). Such uncertainties may generate biased scaling factors. Secondly, the mismatch of spatial resolutions between GRACE and *in situ* observations further results in discrepancies between $GWSA_{SF}$ and the validation dataset. For example, the area of Beijing (GRID #61) is below the footprint of GRACE, but $GWSA_{SF}$ shows a rapid increase in Beijing during July-December, 2012 in contrast to the stable fluctuations reflected by the validation data. The extreme rainfall event in July 2012 (Zhang et al. 2013; Zhang et al. 2015) is believed to be the cause for that mismatch, as GRACE and *in situ* observations may capture different mass variations during the flood event at different spatial resolutions.

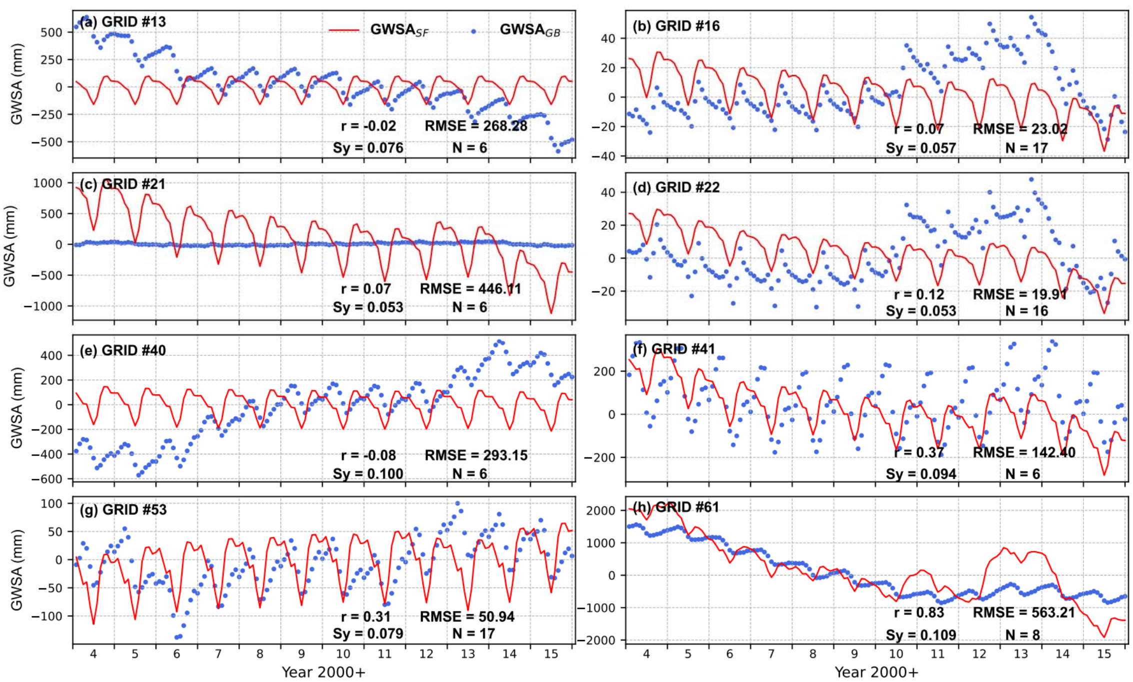


Figure 4. Comparison between monthly $GWSA_{SF}$ and independent measurements at the grid cells showing large discrepancies (r < 0.5, RMSE > 500 mm/mo.). The Sy and number of total observation wells (N) at each grid cell are shown above for understanding uncertainties of *in situ* data.

4.2. Divergent GWSA in the NCP based on the SF_{GB} downscaling method

In this section, GWL data from all available observation wells and the CGS Sy were used to derive scaling factors for the NCP. Results show that GWSA in the entire NCP represents a decreasing rate of -32.60±0.99 mm/yr (4.56±0.14 Gt/yr) during 2004-2015 (**Error! Reference source not found.**5), which was more significant than the rate (-18.6±0.8 mm/yr) estimated by the Mascon solution during 2003-2015 (Gong et al. 2018) and the rate (-28.57 mm/yr) simulated by the MODFLOW model during 1960-2008 (Cao et al. 2013).

Compared with previous GRACE studies in the NCP (Feng et al. 2013; Gong et al. 2018; Huang et al. 2015), this study showed remarkable spatial variations in GWSA at the sub-regional scale, demonstrating groundwater loss in the piedmont west and gains in the coastal east. Significant depletion rates were found in PP (-66.76±2.03 mm/yr) and CP (-23.12±0.65 mm/yr), while the GWSA in EP showed a slightly increasing trend (7.24±0.27 mm/yr) (Figure 5). In addition, the magnitude of the GWSA trend was identified as decreasing from PP in the west to the eastern coastal area (EP). Such spatial variations echo previous studies reporting the conditions of groundwater pumping and hydrogeological characteristics from the PP to the EP (Huang et al. 2015).

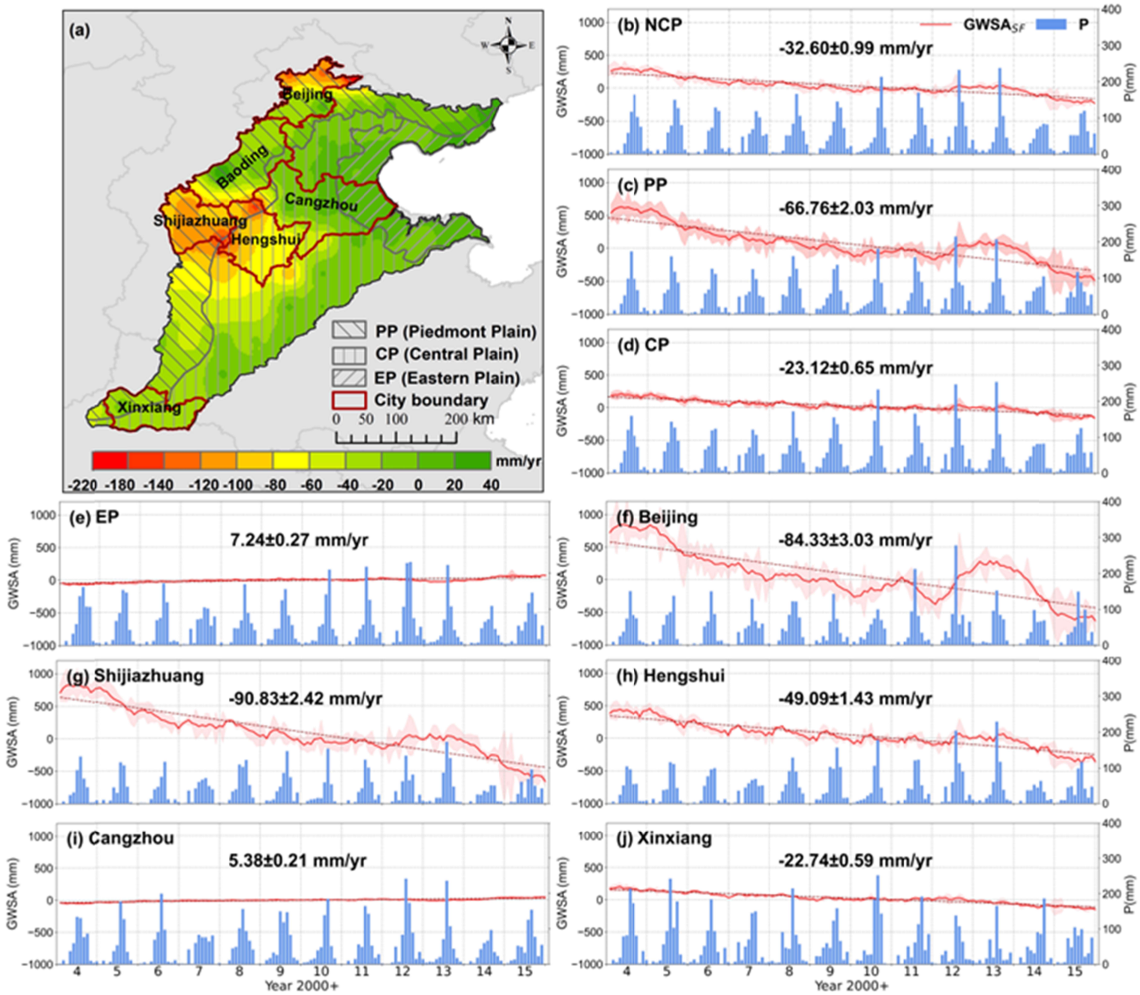


Figure 5. The spatial distribution of GWSA trend in the NCP during 2004-2015 corrected by the SF_{GB} (a), and SF_{GB}-corrected monthly GWSA averaged for the entire NCP (b), the sub-regions, PP (c), CP (d), and EP (e), and the cities, Beijing (f), Shijiazhuang (g), Hengshui (h), Cangzhou (i), and Xinxiang (j).

At the city level, Shijiazhuang, Beijing, and Hengshui cities are identified as the regions that have suffered the most significant groundwater depletion, with a rate of -90.83±2.42 mm/yr (-0.61±0.02 Gt/yr), -84.33±3.03 mm/yr (-0.55±0.02 Gt/yr), and -49.09±1.43 mm/yr (-0.43±0.01 Gt/yr), respectively (Figure 5). These cities have the most intensive groundwater pumping activities over the NCP (Zhang et al. 2021). In contrast, Cangzhou City exhibited a slightly increasing trend of 5.38±0.21 mm/yr (or 0.08±0.01 Gt/yr), which agrees with the GWSA recovery period since 2005 revealed by the InSAR measurements (Jiang et al. 2018). Such an increasing trend could be attributed to the strict groundwater management policies issued by the local government, indicating that the proposed methods in this study may be effective in revealing the actual groundwater dynamics under regional water management. Furthermore, the estimated GWSA reflects the impact of spatial-temporal variations of precipitation. For instance, the abnormal increase of GWSA in Beijing in 2012 was believed to be related to extreme rainfall (Zhang et al. 2013; Zhang et al. 2015). A noticeable decrease in GWSA during 2014-2015, especially in the unconfined aquifers of the PP (*e.g.*, Beijing and Shijiazhuang), was ascribed to the concurrent severe drought (Zhang et al. 2021).

In addition to the divergent GWSA over the NCP, the SF_{GB}-based results highlight the uncertainties in estimated GWSA in some regions. For example, significant differences are recognized among different estimates in Beijing (SF_{GB}-based rate: -84.33±3.03 mm/yr, *in situ*-based rate: -68.49±3.86 mm/yr, and reported rate from water resources bulletin: -43.76 mm/yr by). The larger depletion rate revealed by the SF_{GB}-based estimate may be resulted from overestimated Sy and/or insufficient observation wells, as to be discussed in Section 4.3. However, such a discrepancy may be reasonable considering the lack of *in situ* Sy in Beijing, and it is possible to use the SF_{GB} to further estimate Sy at the sub-regional scale in Beijing.

4.3. Sensitivity analysis

Given the uncertainties in the *a priori* information due to inaccurate Sy and/or the scarcity and uneven distribution of GWL observation wells, both SF_{GB}- and *in situ*-based method may overestimate or underestimate GWSA. Thus, a set of numerical experiments was conducted to investigate the sensitivity of the SF_{GB} downscaling method and the *in situ*-based method to the variability of Sy and the number of GWL observation wells.

The statistical results of numerical experiments at different spatial scales under different scenarios are shown in Figure 6 and Table S1-S3. The magnitude of sensitivity (or uncertainty) can be measured by the distance between the upper and lower quartile lines in Figure 6. Compared to Scenario #1 ("*fixed wells, variable Sy*", Figure 6a) and Scenario #3 ("*variable Sy, variable wells*", Figure 6c), the trends of $GWSA_{SF}$ and $GWSA_{GB}$ in Scenario #2 ("*fixed Sy, variable wells*", Figure 6b) are closer to the reference GWSA trend. Meanwhile, the $GWSA_{SF}$ and $GWSA_{GB}$ are less sensitive to the variability of the number of observation wells at regional and subregional scales in Scenario #2. Even under an extreme situation when only 5 wells (1% of the total number of wells) are selected, relatively good results still can be obtained. The sensitivity of $GWSA_{SF}$ to the variability of wells increases slightly relative to $GWSA_{GB}$ when the spatial scale decreased to the city scale. Both $GWSA_{SF}$ and $GWSA_{GB}$ are more sensitive to the variability of Sy than the variability of wells.

In Scenario #2, the reported Sy from CGS was used, while in Scenarios #1 and #3, the random Sy between 0 and 0.5 was used. The random Sy in Scenarios #1 and #3 influenced the accuracy of both $GWSA_{SF}$ and $GWSA_{GB}$ considerably, but the $GWSA_{SF}$ trends at almost all spatial scales were closer to the reference GWSA trend than $GWSA_{GB}$, indicating the effectiveness of SF_{GB} in correcting the leakage error from grid scale to the entire regional scale.

$GWSA_{SF}$ has a different magnitude of sensitivity to the variability of Sy and/or wells at different spatial scales, and the optimization ratio of $GWSA_{SF}$ compared to $GWSA_{GB}$ varies among different scenarios. At regional, subregional, and city scales, $GWSA_{SF}$ shows lower sensitivity than GWS_{GB} in Scenario #1 and Scenario #3. In Scenario #1 (or #3), the uncertainty in $GWSA_{SF}$ trends decreased significantly at the three scales, with the largest optimization ratio of 90% (#3: 84%) in the NCP, followed by 81% (#3: 73%) in the CP, 74% in the PP (#3: 69%), and 69% in Hengshui (#3: 61%) (Table S1 and S3, Figure 6). Relative low optimization ratios were found in the EP in Scenario #1 and #3, and in Xinxiang in Scenario #1 (see Table S1 and S3, Figure 6). In Scenario #2, optimization of $GWSA_{SF}$ was found in the NCP (by 65%), PP (by 4%), and CP (by 3%) under the low sensitivity to the variability of wells (Table S2, Figure 6). Overall, the SF_{GB} downscaling method is capable of reducing the uncertainty in GWSA induced by the random Sy at any spatial scales as indicated in Scenarios #1 and #3, and this method can be used to estimate large-scale GWSA when reliable Sy is already known but the *in situ* observation wells are insufficient as indicated in Scenario #2.

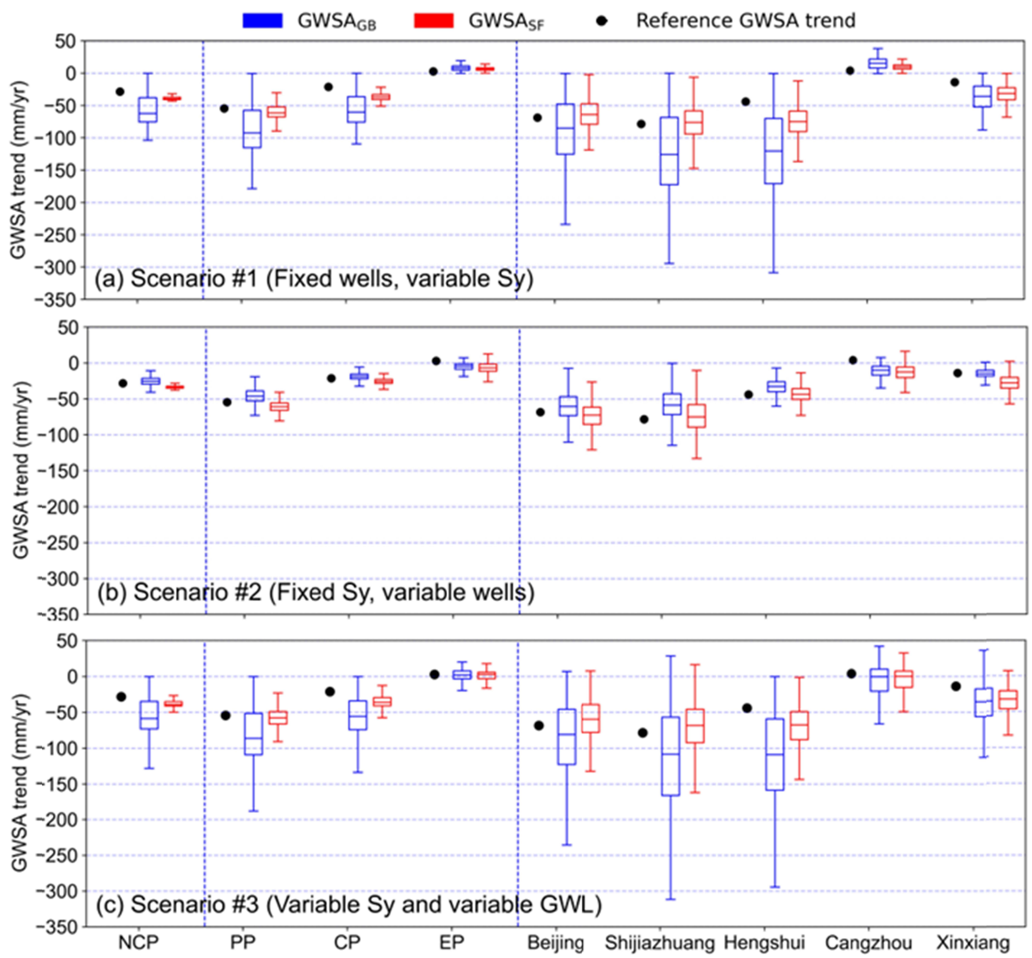


Figure 6. Multi-scale result statistics of numerical experiments for sensitivity analysis of SF_{GB}-corrected and *in situ*-based GWSAs under three scenarios as indicated in Section 3.3. The statistics were performed at varying spatial scales including regional scale (NCP), sub-regional scales (PP, CP, and EP), and city scales (Beijing, Shijiazhuang, Hengshui, Cangzhou, and Xinxiang). The error bars indicate the standard deviations in the 1,000 estimated GWSAs. The upper, middle, and lower caps of the rectangles represent the upper quartile, median, and lower quartile of the 1,000 different estimates, respectively.

4.4 Caveats

The proposed SF_{GB} downscaling method is applicable under the situation when observation wells are insufficient and the reliable Sy values are simultaneously unknown. However, a premise should be considered that the available insufficient observation wells should be relatively evenly distributed and can be interpolated into grid cells based on which the forward modeling is performed to estimate scaling factors. The SF_{GB} downscaling method is less sensitive to the variability of observation wells than to the Sy. In other words, $GWSA_{SF}$ relies less on the number of wells. Nevertheless, attention should be paid to the heterogeneity of GWL variability within one grid. If there is a limited number of wells in one grid (*e.g.*, 0.5°×0.5°) with high heterogeneity of hydrogeology conditions, the GWL data of the limited wells may not be able to represent the average GWL in that grid. Such a situation may result in a large bias in the estimated $GWSA_{SF}$, especially at the city and grid scales.

5. Conclusions

This study proposed a new statistical downscaling approach to derive high-resolution GWS changes using the ground-based scaling factor (SF_{GB}) method by jointly using GRACE and *in situ* GWL data. The proposed method takes advantage of these two datasets by combining the high-resolution *in situ* GWL data with the independent measurements of mass changes from GRACE. Independent validation and numerical experiments demonstrated the effectiveness of SF_{GB} in reducing uncertainties arising from Sy and improving the spatial resolution of estimated GWSA. In areas with a certain amount of observation wells, the SF_{GB}-based statistical downscaling method could provide better estimates of GWSA than using *in situ* GWL data or GRACE data alone. The improved GWSA estimates would be beneficial to water resources management departments which usually have a desire for higher-resolution and lower-uncertainty GWSA datasets for policy-making.

Previous studies have revealed an overall groundwater depletion over the NCP. However, the findings based on the SF_{GB} downscaling method in this study provided updated information against those traditional insights, revealing obvious GWSA variability at the sub-regional and city scales. The most significant GWS loss occurred in Shijiazhuang in the piedmont west with a depletion rate of -90.83±2.42 mm/yr (-0.61±0.02 Gt/yr), in contrast to the slightly increasing trend (5.38±0.21 mm/yr or 0.08±0.01 Gt/yr) in Cangzhou city in the coastal east under the water resources management policies. The GWSA estimates in the NCP at a higher resolution obtained in this study would promote our understanding of the impacts of climate, hydrogeology, and human intervention on regional groundwater resources, and help identify priorities for regional groundwater management practices.

Acknowledgments

This study was jointly supported by the National Natural Science Foundation of China (No. 42071397, 41771456, 42201345), the Second Tibetan Plateau Scientific Expedition and Research Program (STEP, No. 2019QZKK0207-02), and Beijing Natural Science Foundation (No. 8232021).

References:

Adachi, S.A., & Tomita, H. (2020). Methodology of the Constraint Condition in Dynamical Downscaling for Regional Climate Evaluation: A Review. *Journal of Geophysical Research: Atmospheres, 125*

Alley, W.M., & Konikow, L.F. (2015). Bringing GRACE Down to Earth. *Groundwater, 53*, 826-829

Arshad, A., Mirchi, A., Samimi, M., & Ahmad, B. (2022). Combining downscaled-GRACE data with SWAT to improve the estimation of groundwater storage and depletion variations in the Irrigated Indus Basin (IIB). *SCIENCE OF THE TOTAL ENVIRONMENT, 838*, 156044

Atkinson, P.M. (2013). Downscaling in remote sensing. *International Journal of Applied Earth Observation and Geoinformation, 22*, 106-114

Bhanja, S.N., Mukherjee, A., Saha, D., Velicogna, I., & Famiglietti, J.S. (2016). Validation of GRACE based groundwater storage anomaly using in-situ groundwater level measurements in India. *JOURNAL OF HYDROLOGY, 543*, 729-738

Bhanja, S.N., & Mukherjee, A. (2019). In situ and satellite-based estimates of usable groundwater storage across India: Implications for drinking water supply and food security. *ADVANCES IN WATER RESOURCES, 126*, 15-23

Brown, C.M., Greene, A.M., Block, P.J., & Giannini, A. (2008). Review of Downscaling Methodologies for Africa Climate Applications. In

Cao, G., Zheng, C., Scanlon, B.R., Liu, J., & Li, W. (2013). Use of flow modeling to assess sustainability of groundwater resources in the North China Plain. *WATER RESOURCES RESEARCH, 49*, 159-175

Chandanpurkar, H.A., Reager, J.T., Famiglietti, J.S., Nerem, R.S., Chambers, D.P., Lo, M.H., Hamlington, B.D., & Syed, T.H. (2021). The Seasonality of Global Land and Ocean Mass and the Changing Water Cycle. *GEOPHYSICAL RESEARCH LETTERS, 48*

Chen, J., Famigliett, J.S., Scanlon, B.R., & Rodell, M. (2016). Groundwater Storage Changes: Present Status from GRACE Observations. *SURVEYS IN GEOPHYSICS, 37*, 397-417

Chen, J.L., Wilson, C.R., Blankenship, D., & Tapley, B.D. (2009). Accelerated Antarctic ice loss from satellite gravity measurements. *Nature Geoscience, 2*, 859-862

Chen, J.L., Wilson, C.R., Li, J., & Zhang, Z. (2015). Reducing leakage error in GRACE-observed long-term ice mass change: a case study in West Antarctica. *JOURNAL OF GEODESY, 89*, 925-940

Chen, J.L., Wilson, C.R., Tapley, B.D., Yang, Z.L., & Niu, G.Y. (2009). 2005 drought event in the Amazon River basin as measured by GRACE and estimated by climate models. *Journal of Geophysical Research, 114*

Chen, L., He, Q., Liu, K., Li, J., & Jing, C. (2019). Downscaling of GRACE-Derived Groundwater Storage Based on the Random Forest Model. *Remote Sensing, 11*, 2979

Chen, S., Yu, P., & Tang, Y. (2010). Statistical downscaling of daily precipitation using support vector machines and multivariate analysis. *JOURNAL OF HYDROLOGY, 385*, 13-22

Duan, Z., & Bastiaanssen, W.G.M. (2013). First results from Version 7 TRMM 3B43 precipitation product in combination with a new downscaling‐calibration procedure. *REMOTE SENSING OF ENVIRONMENT, 131*, 1-13

Famiglietti, J.S., Cazenave, A., Eicker, A., Reager, J.T., Rodell, M., & Velicogna, I. (2015). Satellites provide the big picture. *SCIENCE, 349*, 684-685

Famiglietti, J.S., Lo, M., Ho, S.L., Bethune, J., Anderson, K.J., Syed, T.H., Swenson, S.C., de Linage, C.R., & Rodell, M. (2011). Satellites measure recent rates of groundwater depletion in California's Central Valley. *GEOPHYSICAL RESEARCH LETTERS, 38*

Famiglietti, J.S., & Rodell, M. (2013). Water in the Balance. *SCIENCE, 340*, 1300-1301

Feng, W., Zhong, M., Lemoine, J.M., Biancale, R., Hsu, H.T., & Xia, J. (2013). Evaluation of groundwater depletion in North China using the Gravity Recovery and Climate Experiment (GRACE) and ground-based measurements. In, *Egu General Assembly Conference*

Gehman, C.L., Harry, D.L., Sanford, W.E., Stednick, J.D., & Beckman, N.A. (2022). Estimation of Specific Yield for Regional Groundwater Models: Pitfalls, Ramifications, and a Promising Path Forward. *WATER RESOURCES RESEARCH, v. 58*, e2021W-e30761W

Gong, H., Pan, Y., Zheng, L., Li, X., Zhu, L., Zhang, C., Huang, Z., Li, Z., Wang, H., & Zhou, C. (2018a). Long-term groundwater storage changes and land subsidence development in the North China Plain (1971-2015). *HYDROGEOLOGY JOURNAL, 26*, 1417-1427

Gong, H., Pan, Y., Zheng, L., Li, X., Zhu, L., Zhang, C., Huang, Z., Li, Z., Wang, H., & Zhou, C. (2018b). Long-term groundwater storage changes and land subsidence development in the North China Plain (1971-2015). *HYDROGEOLOGY JOURNAL, 26*, 1417-1427

Hachborn, E., Berg, A., Levison, J., & Ambadan, J.T. (2017). Sensitivity of GRACE-derived estimates of groundwater-level changes in southern Ontario, Canada. *HYDROGEOLOGY JOURNAL, 25*, 2391-2402

Henry, C.M., Allen, D.M., & Huang, J. (2011). Groundwater storage variability and annual recharge using well-hydrograph and GRACE satellite data. *HYDROGEOLOGY JOURNAL, 19*, 741-755

Huang, Z., Jiao, J.J., Luo, X., Pan, Y., & Zhang, C. (2019). Sensitivity Analysis of Leakage Correction of GRACE Data in Southwest China Using A-Priori Model Simulations: Inter-Comparison of Spherical Harmonics, Mass Concentration and In Situ Observations. *SENSORS, 19*

Huang, Z., Pan, Y., Gong, H., Yeh, P.J., Li, X., Zhou, D., & Zhao, W. (2015). Subregional-scale groundwater depletion detected by GRACE for both shallow and deep aquifers in North China Plain. *GEOPHYSICAL RESEARCH LETTERS, 42*, 1791-1799

Hunink, J.E., Immerzeel, W.W., & Droogers, P. (2014). A High-resolution Precipitation 2-step mapping Procedure (HiP2P): Development and application to a tropical mountainous area. *REMOTE SENSING OF ENVIRONMENT, 140*, 179-188

Jiang, L., Bai, L., Zhao, Y., Cao, G., Wang, H., & Sun, Q. (2018). Combining InSAR and Hydraulic Head Measurements to Estimate Aquifer Parameters and Storage Variations of Confined Aquifer system in Cangzhou North China Plain. *WATER RESOURCES RESEARCH, 54*, 8234-8252

Jyolsna, P.J., Kambhammettu, B.V.N.P., & Gorugantula, S. (2021). Application of random forest and multi-linear regression methods in downscaling GRACE derived groundwater storage changes. *Hydrological sciences journal, 66*, 874-887

Landerer, F.W., & Swenson, S.C. (2012). Accuracy of scaled GRACE terrestrial water storage estimates. *WATER RESOURCES RESEARCH, 48*

Leblanc, M.J., Tregoning, P., Ramillien, G., Tweed, S.O., & Fakes, A. (2009). Basin-scale, integrated observations of the early 21st century multiyear drought in southeast Australia. *WATER RESOURCES*

RESEARCH, 45

Liu, B., & Zou, X. (2019). Comparison and Analysis of GRACE Time-Varying Signal Recovery Methods. *Journal of Geodesy and Geodynamics, 39*, 204-209

Liu, Y., Jing, W., Wang, Q., & Xia, X. (2020). Generating high-resolution daily soil moisture by using spatial downscaling techniques: a comparison of six machine learning algorithms. *ADVANCES IN WATER RESOURCES, 141*, 103601

Long, D., Yang, Y., Wada, Y., Hong, Y., Liang, W., Chen, Y., Yong, B., Hou, A., Wei, J., & Chen, L. (2015). 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, 168*, 177-193

Longuevergne, L., Scanlon, B.R., & Wilson, C.R. (2010). GRACE Hydrological estimates for small basins: Evaluating processing approaches on the High Plains Aquifer, USA. *WATER RESOURCES RESEARCH, 46*

López López, P., Immerzeel, W.W., Rodríguez Sandoval, E.A., Sterk, G., & Schellekens, J. (2018). Spatial Downscaling of Satellite-Based Precipitation and Its Impact on Discharge Simulations in the Magdalena River Basin in Colombia. *Frontiers in Earth Science, 6*

Lv, M., Xu, Z., Yang, Z.I., Lu, H., & Lv, M. (2021). A comprehensive review of specific yield in land surface and groundwater studies. *Journal of Advances in Modeling Earth Systems*

Miro, M., & Famiglietti, J. (2018). Downscaling GRACE Remote Sensing Datasets to High-Resolution Groundwater Storage Change Maps of California’s Central Valley. *Remote Sensing, 10*, 143

Panda, D.K., & Wahr, J. (2016). Spatiotemporal evolution of water storage changes in India from the updated GRACE-derived gravity records. *WATER RESOURCES RESEARCH, 52*, 135-149

Peng, J., Loew, A., Merlin, O., & Verhoest, N.E.C. (2017). A review of spatial downscaling of satellite remotely sensed soil moisture. *REVIEWS OF GEOPHYSICS, 55*, 341-366

Perfilieva, I., Yarushkina, N., Afanasieva, T., & Romanov, A. (2013). Time series analysis using soft computing methods. *INTERNATIONAL JOURNAL OF GENERAL SYSTEMS, 42*, 687-705

Pulla, S.T., Yasarer, H., & Yarbrough, L.D. (2023). GRACE Downscaler: A Framework to Develop and Evaluate Downscaling Models for GRACE. *Remote Sensing, 15*, 2247

Quintana Seguí, P., Ribes, A., Martin, E., Habets, F., & Boé, J. (2010). Comparison of three downscaling methods in simulating the impact of climate change on the hydrology of Mediterranean basins. *JOURNAL OF HYDROLOGY, 383*, 111-124

Reager, J.T., & Famiglietti, J.S. (2009). Global terrestrial water storage capacity and flood potential using GRACE. *GEOPHYSICAL RESEARCH LETTERS, 36*

Richey, A.S., Thomas, B.F., Lo, M.H., Reager, J.T., Famiglietti, J.S., Voss, K., Swenson, S., & Rodell, M. (2015). Quantifying renewable groundwater stress withGRACE. *WATER RESOURCES RESEARCH, 51*, 5217-5238

Rodell, M., Chen, J., Kato, H., Famiglietti, J.S., Nigro, J., & Wilson, C.R. (2007). Estimating groundwater storage changes in the Mississippi River basin (USA) using GRACE. *HYDROGEOLOGY JOURNAL, 15*, 159-166

Rodell, M., Velicogna, I., & Famiglietti, J.S. (2009). Satellite-based estimates of groundwater depletion in India. *NATURE, 460*, 999-1002

Rodell, M., & Famiglietti, J.S. (2002). The potential for satellite-based monitoring of groundwater storage changes using GRACE; the High Plains Aquifer, central US. *Journal of hydrology (Amsterdam), 263*, 245-256

Sahour, H., Sultan, M., Vazifedan, M., Abdelmohsen, K., Karki, S., Yellich, J., Gebremichael, E., Alshehri, F., & Elbayoumi, T. (2020). Statistical Applications to Downscale GRACE-Derived Terrestrial Water Storage Data and to Fill Temporal Gaps. *Remote Sensing, 12*, 533

Saikrishna, T.S., Ramu, D.A., Prasad, K.B.R.R., Osuri, K.K., & Rao, A.S. (2022). High resolution dynamical downscaling of global products using spectral nudging for improved simulation of Indian monsoon rainfall. *ATMOSPHERIC RESEARCH, 280*, 106452

Samadi, S., Carbone, G.J., Mahdavi, M., Sharifi, F., & Bihamta, M.R. (2013). Statistical Downscaling of River Runoff in a Semi Arid Catchment. *WATER RESOURCES MANAGEMENT, 27*, 117-136

Scanlon, B.R., Longuevergne, L., & Long, D. (2012a). Ground referencing GRACE satellite estimates of groundwater storage changes in the California Central Valley, USA. *WATER RESOURCES RESEARCH, 48*

Scanlon, B.R., Longuevergne, L., & Long, D. (2012b). Ground referencing GRACE satellite estimates of groundwater storage changes in the California Central Valley, USA. *WATER RESOURCES RESEARCH, 48*

Seyoum, W., Kwon, D., & Milewski, A. (2019). Downscaling GRACE TWSA Data into High-Resolution Groundwater Level Anomaly Using Machine Learning-Based Models in a Glacial Aquifer System. *Remote Sensing, 11*, 824

Shamsudduha, M., Taylor, R.G., & Longuevergne, L. (2012a). Monitoring groundwater storage changes in the highly seasonal humid tropics: Validation of GRACE measurements in the Bengal Basin. *WATER RESOURCES RESEARCH, 48*

Shamsudduha, M., Taylor, R.G., & Longuevergne, L. (2012b). Monitoring groundwater storage changes in the highly seasonal humid tropics: Validation of GRACE measurements in the Bengal Basin. *WATER RESOURCES RESEARCH, 48*

Shamsudduha, M., Taylor, R.G., & Longuevergne, L. (2012c). Monitoring groundwater storage changes in the highly seasonal humid tropics: Validation of GRACE measurements in the Bengal Basin. *WATER RESOURCES RESEARCH, 48*

Shen, Z., Zhang, Q., Singh, V.P., Sun, P., He, C., & Cheng, C. (2021). Station ‐ based non ‐ linear regression downscaling approach: A new monthly precipitation downscaling technique. *INTERNATIONAL JOURNAL OF CLIMATOLOGY, 41*, 5879-5898

Strassberg, G., Scanlon, B.R., & Chambers, D. (2009). Evaluation of groundwater storage monitoring with the GRACE satellite: Case study of the High Plains aquifer, central United States. *WATER RESOURCES RESEARCH, 45*

Swenson, S., Wahr, J., & Milly, P. (2003). Estimated accuracies of regional water storage variations inferred from the Gravity Recovery and Climate Experiment (GRACE). *WATER RESOURCES RESEARCH, 39*

Swenson, S., & Wahr, J. (2002). Methods for inferring regional surface-mass anomalies from Gravity Recovery and Climate Experiment (GRACE) measurements of time-variable gravity. *JOURNAL OF GEOPHYSICAL RESEARCH-SOLID EARTH, 107*

Syed, T.H., Famiglietti, J.S., Rodell, M., Chen, J., & Wilson, C.R. (2008). Analysis of terrestrial water storage changes from GRACE and GLDAS. *WATER RESOURCES RESEARCH, 44*

Syed, T.H., Famiglietti, J.S., & Chambers, D.P. (2009). GRACE-Based Estimates of Terrestrial Freshwater Discharge from Basin to Continental Scales. *JOURNAL OF HYDROMETEOROLOGY, 10*, 22-40

Tang, J., Niu, X., Wang, S., Gao, H., Wang, X., & Wu, J. (2016). Statistical downscaling and dynamical downscaling of regional climate in China: Present climate evaluations and future climate projections. *Journal of Geophysical Research: Atmospheres, 121*, 2110-2129

Teng, H., Shi, Z., Ma, Z., & Li, Y. (2014). Estimating spatially downscaled rainfall by regression kriging using TRMM precipitation and elevation in Zhejiang Province, southeast China. *INTERNATIONAL JOURNAL OF REMOTE SENSING, 35*, 7775-7794

Vishwakarma, B.D., Horwath, M., Devaraju, B., Groh, A., & Sneeuw, N. (2017). A Data-Driven Approach for Repairing the Hydrological Catchment Signal Damage Due to Filtering of GRACE Products. *WATER RESOURCES RESEARCH, 53*, 9824-9844

Xiang, L., Wang, H., Steffen, H., Wu, P., Jia, L., Jiang, L., & Shen, Q. (2016). Groundwater storage changes in the Tibetan Plateau and adjacent areas revealed from GRACE satellite gravity data. *EARTH AND PLANETARY SCIENCE LETTERS, 449*, 228-239

Xing, L., Guo, H., & Zhan, Y. (2013). Groundwater hydrochemical characteristics and processes along flow paths in the North China Plain. *JOURNAL OF ASIAN EARTH SCIENCES, 70-71*, 250-264

Xu, C., Qu, J., Hao, X., Cosh, M., Prueger, J., Zhu, Z., & Gutenberg, L. (2018). Downscaling of Surface Soil Moisture Retrieval by Combining MODIS/Landsat and In Situ Measurements. *Remote Sensing, 10*, 210

Xu, Y., Wang, L., Ma, Z., Li, B., Bartels, R., Liu, C., Zhang, X., & Dong, J. (2020). Spatially Explicit Model for Statistical Downscaling of Satellite Passive Microwave Soil Moisture. *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, 58*, 1182-1191

Xu, Z., Han, Y., & Yang, Z. (2019). Dynamical downscaling of regional climate: A review of methods and limitations. *Science China Earth Sciences, 62*, 365-375

Yi, S., Wang, Q., & Sun, W. (2016). Basin mass dynamic changes in China from GRACE based on a multibasin inversion method. *Journal of Geophysical Research: Solid Earth, 121*, 3782-3803

Yin, W., Hu, L., Zhang, M., Wang, J., & Han, S. (2018). Statistical Downscaling of GRACE-Derived Groundwater Storage Using ET Data in the North China Plain. *Journal of Geophysical Research: Atmospheres, 123*, 5973-5987

Yin, W., Zhang, G., Han, S., Yeo, I., & Zhang, M. (2022). Improving the resolution of GRACE-based water storage estimates based on machine learning downscaling schemes. *JOURNAL OF HYDROLOGY, 613*, 128447

Zhang, C., Duan, Q., Yeh, P.J., Pan, Y., Gong, H., Moradkhani, H., Gong, W., Lei, X., Liao, W., Xu, L., Huang, Z., Zheng, L., & Guo, X. (2021). Sub-regional groundwater storage recovery in North China Plain after the South-to-North water diversion project. *JOURNAL OF HYDROLOGY, 597*

Zhang, D., Lin, Y., Zhao, P., Yu, X., Wang, S., Kang, H., & Ding, Y. (2013). The Beijing extreme rainfall of 21 July 2012: “Right results” but for wrong reasons. *GEOPHYSICAL RESEARCH LETTERS, 40*, 1426-1431

Zhang, D., Liu, X., & Bai, P. (2019). Assessment of hydrological drought and its recovery time for eight tributaries of the Yangtze River (China) based on downscaled GRACE data. *JOURNAL OF HYDROLOGY, 568*, 592-603

Zhang, G., Zheng, W., Yin, W., & Lei, W. (2021). Improving the Resolution and Accuracy of Groundwater Level Anomalies Using the Machine Learning-Based Fusion Model in the North China Plain. *SENSORS, 21*, 46

Zhang, L., Yi, S., Wang, Q., Chang, L., Tang, H., & Sun, W. (2019). Evaluation of GRACE mascon solutions for small spatial scales and localized mass sources. *GEOPHYSICAL JOURNAL INTERNATIONAL, 218*, 1307-1321

Zhang, Y., Hong, Y., Wang, X.G., Gourley, J.J., Xue, X.W., Saharia, M., Ni, G.H., Wang, G.L., Huang, Y., Chen, S., & Tang, G.Q. (2015). Hydrometeorological Analysis and Remote Sensing of Extremes: Was the July 2012 Beijing Flood Event Detectable and Predictable by Global Satellite Observing and Global Weather Modeling Systems? *JOURNAL OF HYDROMETEOROLOGY, 16*, 381-395

Zhang, Z., Fei, Y., & Chen, Z. (2009). *Survey and Evaluation of Ground-Water Sustainable Utilization in North China Plain [in Chinese]*

Figure1.

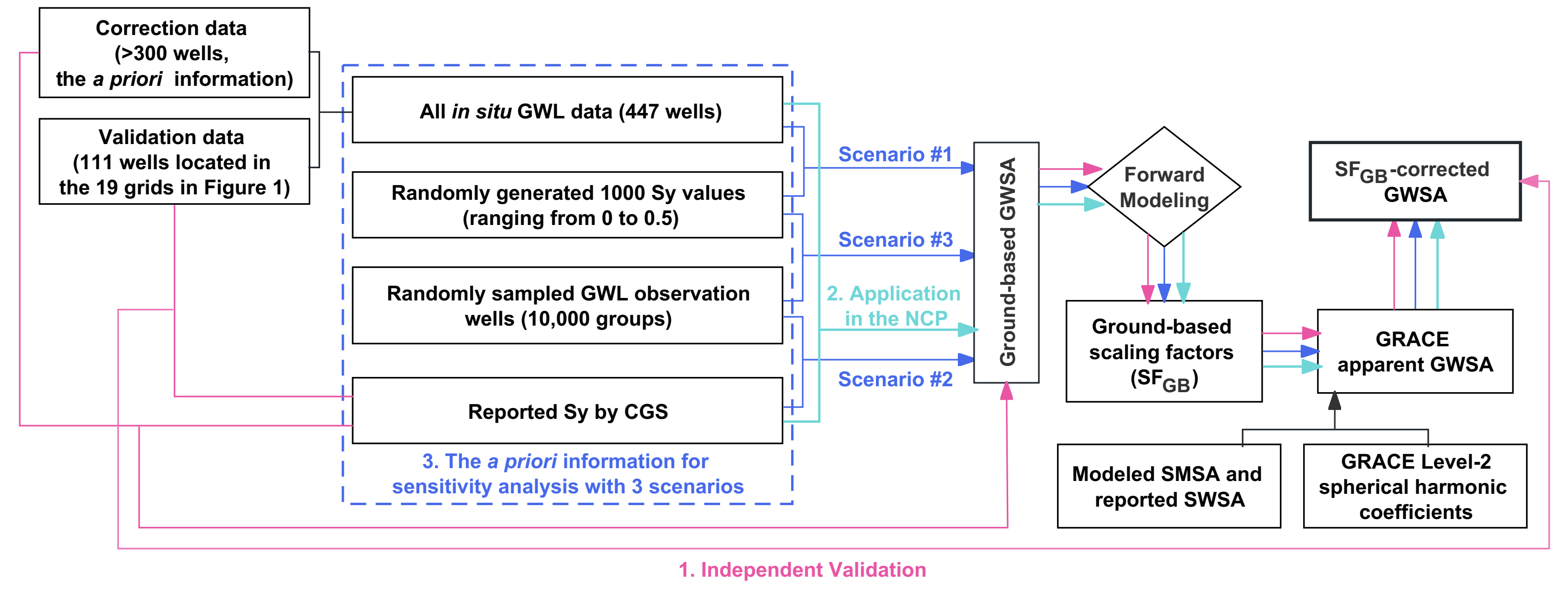

Correction data (>300 wells, the *a priori* information)
Validation data (111 wells located in the 19 grids in Figure 1)
All *in situ* GWL data (447 wells)
Randomly generated 1000 Sy values (ranging from 0 to 0.5)
Randomly sampled GWL observation wells (10,000 groups)
Reported Sy by CGS
3. The *a priori* information for sensitivity analysis with 3 scenarios
Scenario #1
Scenario #3
2. Application in the NCP
Scenario #2
Ground-based GWSA
Forward Modeling
Ground-based scaling factors (SF_{GB})
SF_{GB}-corrected GWSA
GRACE apparent GWSA
Modeled SMSA and reported SWSA
GRACE Level-2 spherical harmonic coefficients
1. Independent Validation

Figure2.

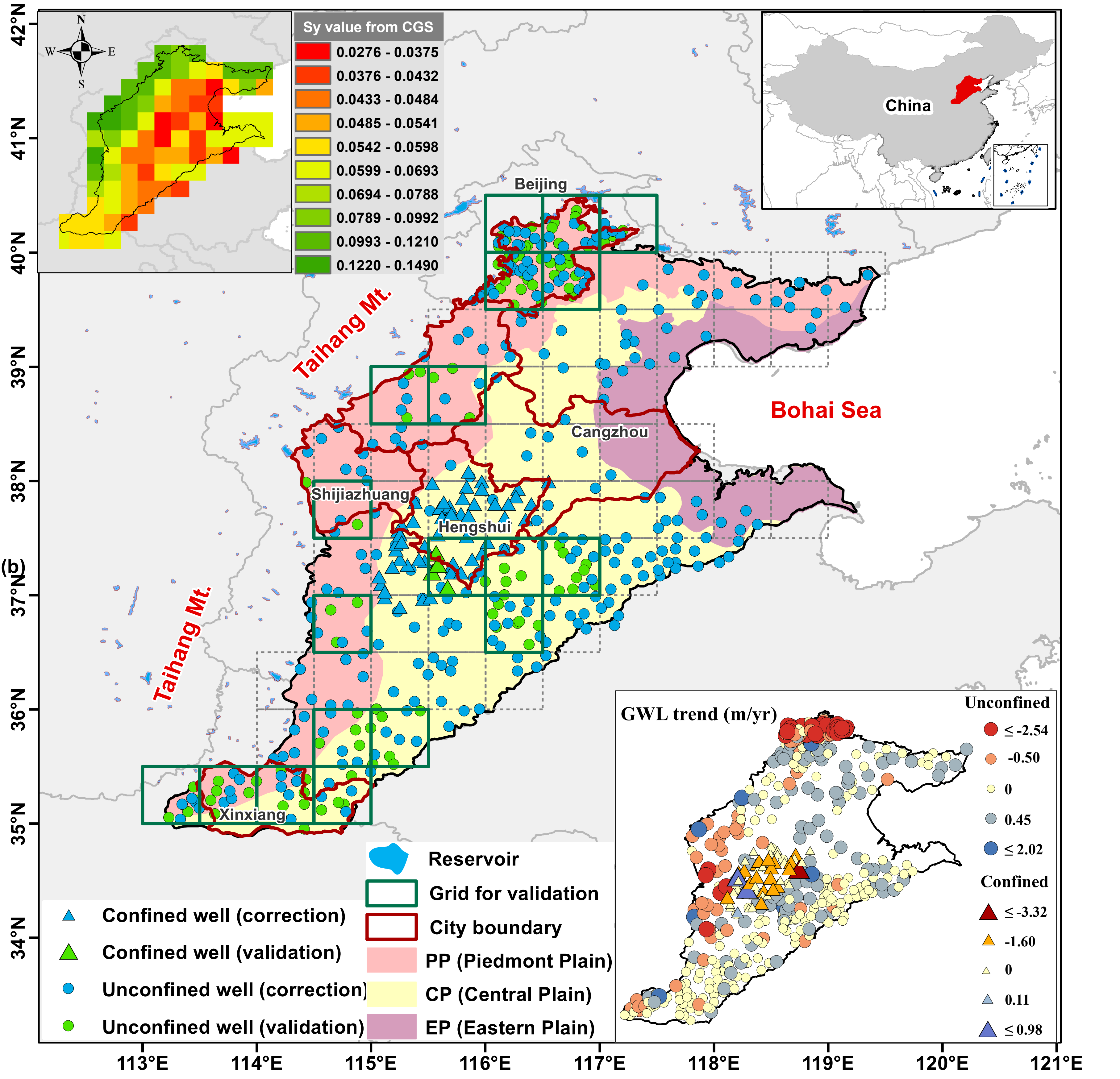

Sy value from CGS
0.0276 - 0.0375
0.0376 - 0.0432
0.0433 - 0.0484
0.0485 - 0.0541
0.0542 - 0.0598
0.0599 - 0.0693
0.0694 - 0.0788
0.0789 - 0.0992
0.0993 - 0.1210
0.1220 - 0.1490
N
S
W
E
China
Beijing
Taihang Mt.
Taihang Mt.
Bohai Sea
Cangzhou
Shijiazhuang
Hengshui
Xinxiang
(b)
42°N
41°N
40°N
39°N
38°N
37°N
36°N
35°N
34°N
113°E
114°E
115°E
116°E
117°E
118°E
119°E
120°E
121°E
GWL trend (m/yr)
Unconfined
≤ -2.54
-0.50
0
0.45
≤ 2.02
Confined
≤ -3.32
-1.60
0
0.11
≤ 0.98
Reservoir
Grid for validation
City boundary
PP (Piedmont Plain)
CP (Central Plain)
EP (Eastern Plain)
Confined well (correction)
Confined well (validation)
Unconfined well (correction)
Unconfined well (validation)

Figure3.

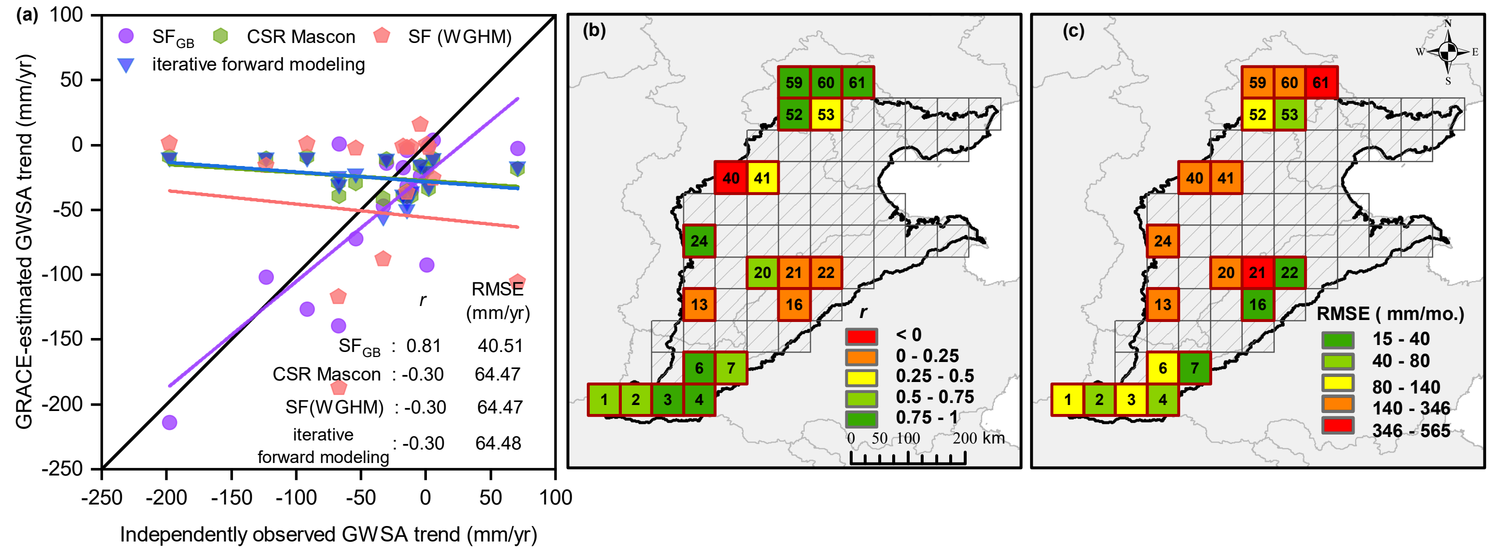
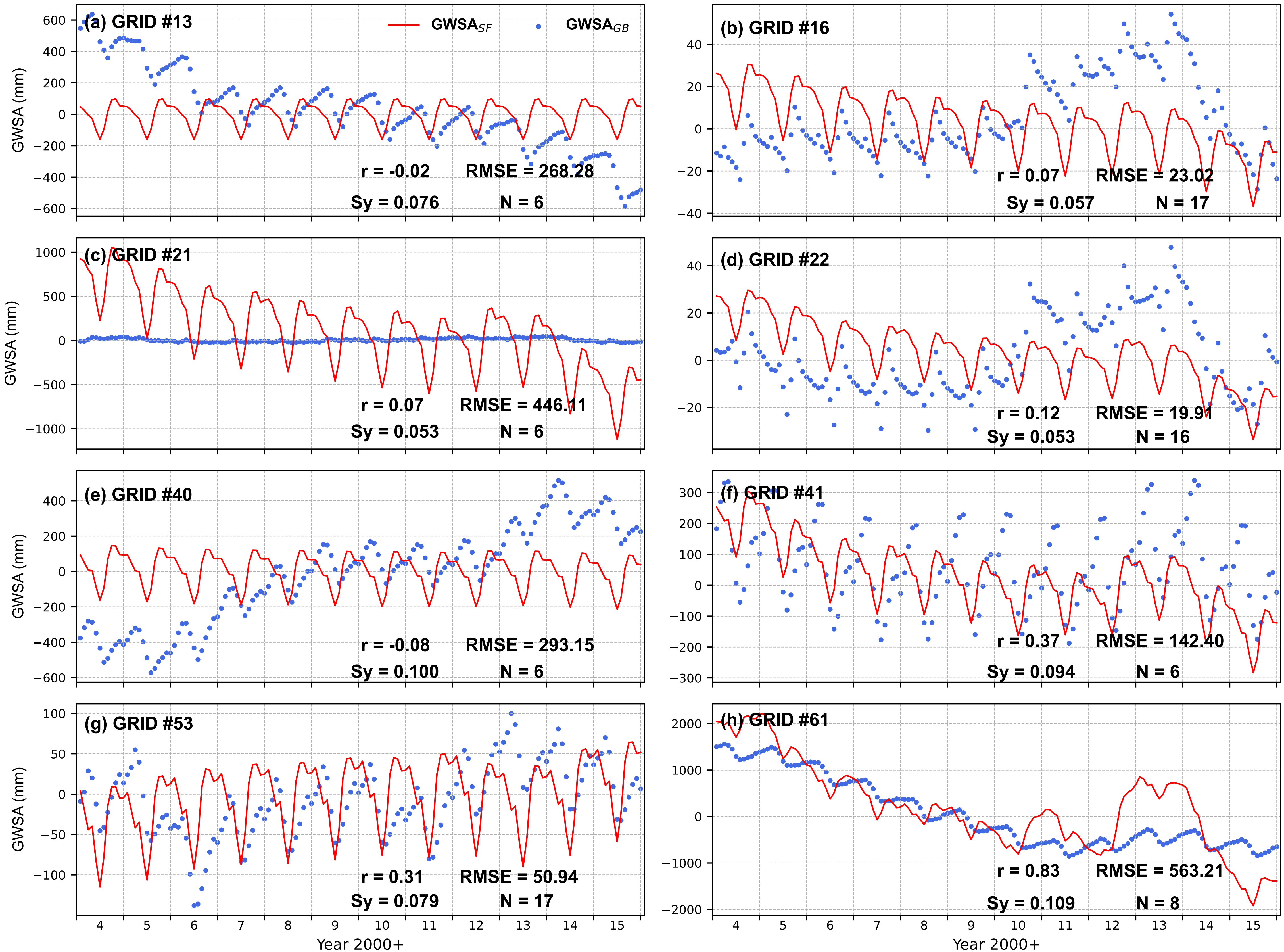

(a)
SF$_{GB}$
CSR Mascon
SF (WGHM)
iterative forward modeling
r
RMSE (mm/yr)
SF$_{GB}$: 0.81 40.51
CSR Mascon : -0.30 64.47
SF(WGHM) : -0.30 64.47
iterative forward modeling : -0.30 64.48
GRACE-estimated GWSA trend (mm/yr)
Independently observed GWSA trend (mm/yr)
(b)
r
< 0
0 - 0.25
0.25 - 0.5
0.5 - 0.75
0.75 - 1
0 50 100 200 km
(c)
RMSE (mm/mo.)
15 - 40
40 - 80
80 - 140
140 - 346
346 - 565
N
W
E
S

Figure4.



(a) GRID #13
GWSA$_{SF}$
GWSA$_{GB}$
r = -0.02
RMSE = 268.28
Sy = 0.076
N = 6
(b) GRID #16
r = 0.07
RMSE = 23.02
Sy = 0.057
N = 17
(c) GRID #21
r = 0.07
RMSE = 446.11
Sy = 0.053
N = 6
(d) GRID #22
r = 0.12
RMSE = 19.91
Sy = 0.053
N = 16
(e) GRID #40
r = -0.08
RMSE = 293.15
Sy = 0.100
N = 6
(f) GRID #41
r = 0.37
RMSE = 142.40
Sy = 0.094
N = 6
(g) GRID #53
r = 0.31
RMSE = 50.94
Sy = 0.079
N = 17
(h) GRID #61
r = 0.83
RMSE = 563.21
Sy = 0.109
N = 8
GWSA (mm)
Year 2000+

Figure5.

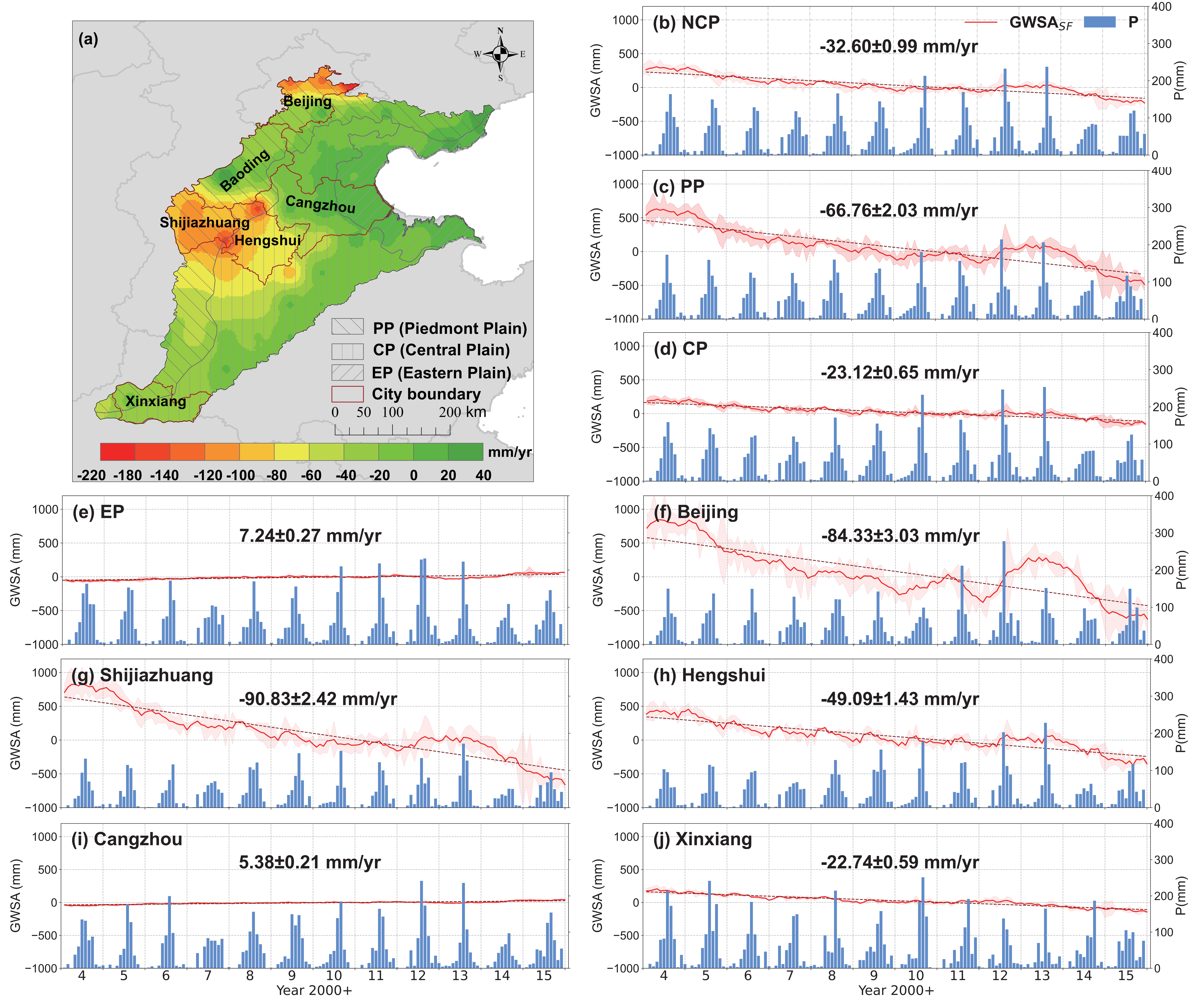
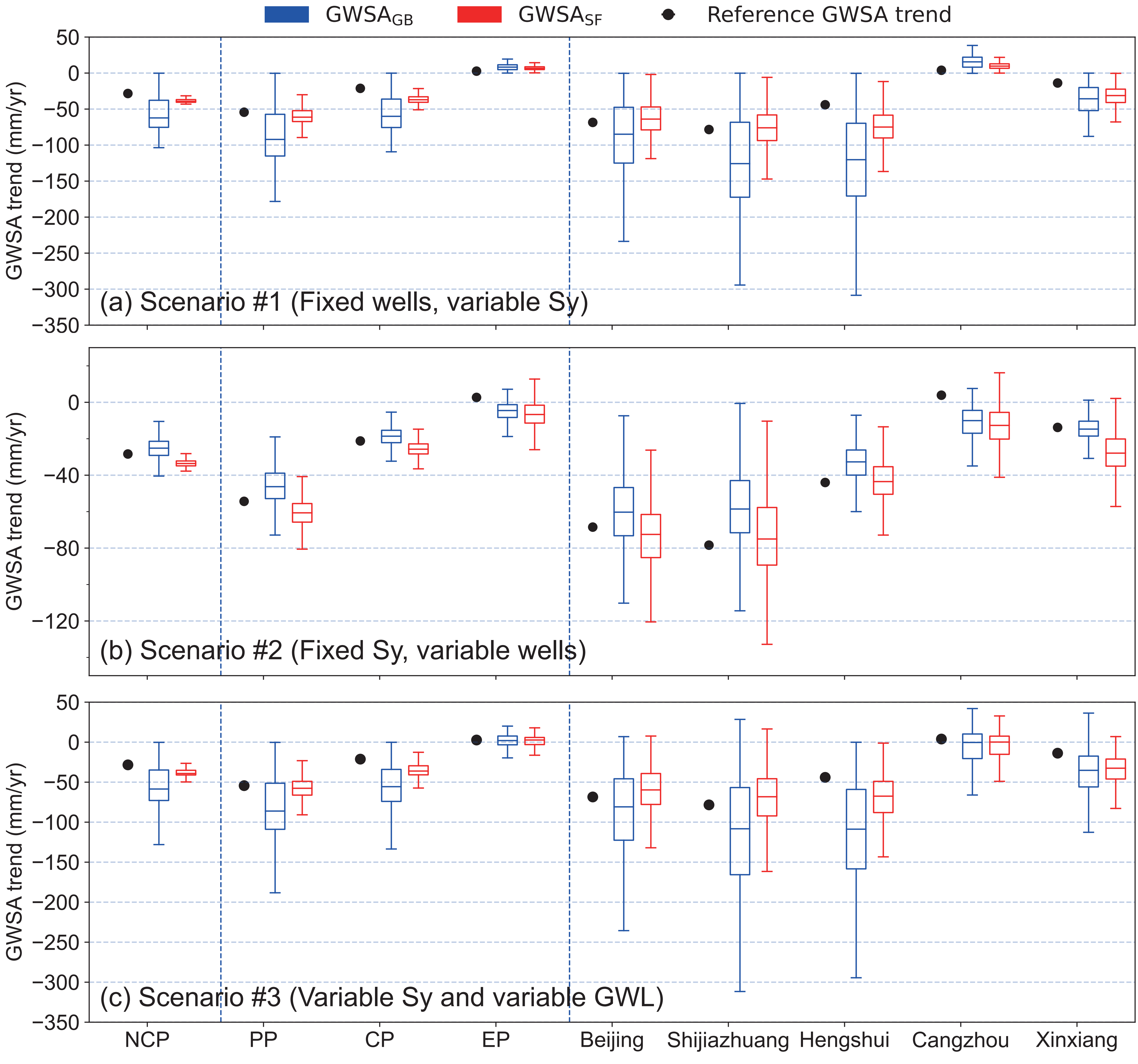

(a)
Beijing
Baoding
Cangzhou
Shijiazhuang
Hengshui
Xinxiang
N
S
W
E
PP (Piedmont Plain)
CP (Central Plain)
EP (Eastern Plain)
City boundary
0 50 100 200 km
mm/yr
-220 -180 -140 -120 -100 -80 -60 -40 -20 0 20 40
(b) NCP
GWSA$_{SF}$
P
-32.60±0.99 mm/yr
(c) PP
-66.76±2.03 mm/yr
(d) CP
-23.12±0.65 mm/yr
(e) EP
7.24±0.27 mm/yr
(f) Beijing
-84.33±3.03 mm/yr
(g) Shijiazhuang
-90.83±2.42 mm/yr
(h) Hengshui
-49.09±1.43 mm/yr
(i) Cangzhou
5.38±0.21 mm/yr
(j) Xinxiang
-22.74±0.59 mm/yr
GWSA (mm)
P(mm)
Year 2000+

Figure6.



GWSA$_{GB}$
GWSA$_{SF}$
Reference GWSA trend
GWSA trend (mm/yr)
(a) Scenario #1 (Fixed wells, variable Sy)
(b) Scenario #2 (Fixed Sy, variable wells)
(c) Scenario #3 (Variable Sy and variable GWL)
50
0
-50
-100
-150
-200
-250
-300
-350
-40
-80
-120
NCP
PP
CP
EP
Beijing
Shijiazhuang
Hengshui
Cangzhou
Xinxiang