Supporting Information for

Groundwater Depletion Estimated from GRACE: A Challenge of Sustainable Development in an Arid Region of Central Asia

Zengyun Hu [1,2,4], Qiming Zhou [3], Xi Chen [2,4], Deliang Chen [5], Jianfeng Li [3,*], Meiyu Guo [3,*], Gang Yin [6] and Zheng Duan [7,8]

[1] Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, 518055, China
[2] State Key Laboratory of desert and Oasis Ecology, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi 830011, China
[3] Department of Geography, Hong Kong Baptist University, Kowloon Tong, Kowloon, Hong Kong, China
[4] Research Center for Ecology and Environment of Central Asia, Chinese Academy of Sciences, Urumqi 830011, China
[5] Department of Earth Sciences, University of Gothenburg, Gothenburg, Sweden
[6] College of Resource and Environment Sciences,, Xinjiang University, Urumqi 830046, China
[7] Department of Physical Geography and Ecosystem Science, Lund University, S-223 62 Lund, Sweden
[8] Chair of Hydrology and River Basin Management, Technical University of Munich, Arcisstrasse 21, 80333 Munich, Germany
* Correspondence: jianfengli@hkbu.edu.hk (J.L.); meiyuguo@hkbu.edu.hk (M.G.)

Text S1 Procedure of PLSR

The procedure of PLSR analysis is introduced here.

For convenience, we assume that the dependent variables matrix is

$$X_{n\times m} = \begin{pmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nm} \end{pmatrix} = (x_1, x_2, \cdots, x_m)$$

with $(x_1, x_2, \cdots, x_m)$ as the dependent variables, m is the number of variables, and n is the length of each dependent variable. And the independent variables matrix is

$$Y_{n\times p} = \begin{pmatrix} y_{11} & \cdots & y_{1p} \\ \vdots & \ddots & \vdots \\ y_{n1} & \cdots & y_{np} \end{pmatrix} = (y_1, y_2, \cdots, y_p)$$

with $(y_1, y_1, \cdots, y_p)$ as the independent variables. $X_{n\times m}$ and $Y_{n\times p}$ are normalized as $A_{n\times m} = (a_1, a_2, \cdots, a_m)$ and $B_{n\times p} = (b_1, b_2, \cdots, b_p)$ by subtracting the mean values and dividing the standard deviation values.

Then, the first component of $X_{n\times m}$ and $Y_{n\times p}$ is extracted and noted as u_1 and v_1, respectively. They have the following linear combinations

$$u_1 = \alpha_{11}x_1 + \alpha_{12}x_2 + \cdots + \alpha_{1m}x_m = \rho^{(1)T}X \tag{1}$$

and

$$v_1 = \beta_{11}y_1 + \beta_{12}y_2 + \cdots \beta_{1p}y_p = \gamma^{(1)T}Y \tag{2}$$

It is requested that

(1) u_1 and $v_1$should extract much various information from $X_{n\times m}$ and $Y_{n\times p}$ as soon as possible that is $Var(u_1) \to max$ and $Var(v_1) \to max$;

(2) The correlation coefficient (Corr) between u_1 and v_1 should be maximum which is

$$Corr(u_1, v_1) \to max.$$

From the normalized matrixes $A_{n\times m}$ and $B_{n\times p}$, the score vectors of u_1 and v_1 have the following forms:

$$\widehat{u_1} = A\rho^{(1)} \text{ and } \widehat{v_1} = B\gamma^{(1)}.$$

The above two conditions are equivalent to the covariance between u_1 and v_1 satisfying

$$Cov(u_1, v_1) \to max,$$

which converts a conditional extreme value problem as

$$max(\widehat{u_1} \cdot \widehat{v_1}) = \left(A\rho^{(1)} \cdot B\gamma^{(1)}\right) = \rho^{(1)T}A^T \cdot B\gamma^{(1)}$$

$$s.t. \begin{cases} \rho^{(1)T}\rho^{(1)} = \left\|\rho^{(1)}\right\|^2 = 1, \\ \gamma^{(1)T}\gamma^{(1)} = \left\|\gamma^{(1)}\right\|^2 = 1. \end{cases}$$

Using the Lagrangian algorithm, $\rho^{(1)}$ is the corresponding eigenvector of the maximum eigenvalue of the matrix $A^T BB^T A$ and $\gamma^{(1)}$ is the corresponding eigenvector of the maximum eigenvalue of the matrix $B^T AA^T B$. And now, the equations (1) and (2) are computed.

According to the principal component analysis method, X and Y can be regressed on u_1. The regression model is

$$\begin{cases} A = \widehat{u_1}\sigma^{(1)T} + A_1, \\ B = \widehat{u_1}\tau^{(1)T} + B_1, \end{cases} \tag{3}$$

where A_1 and B_1 are the residual matrixes, and $\sigma^{(1)}$ and $\tau^{(1)}$ are the regression coefficients with the following form:

$$\begin{cases} \sigma^{(1)} = \dfrac{A^T\widehat{u_1}}{\|\widehat{u_1}\|^2}, \\ \tau^{(1)} = \dfrac{B^T\widehat{u_1}}{\|\widehat{u_1}\|^2}. \end{cases} \tag{4}$$

Repeating the above process by treating A_1 as A and B_1 as B, we have

$$\begin{cases} A_1 = \widehat{u_2}\sigma^{(2)T} + A_2, \\ B_1 = \widehat{u_2}\tau^{(2)T} + B_2, \end{cases} \tag{5}$$

where

$$\begin{cases} \sigma^{(2)} = \dfrac{{A_1}^T\widehat{u_2}}{\|\widehat{u_2}\|^2}, \\ \tau^{(2)} = \dfrac{{B_1}^T\widehat{u_2}}{\|\widehat{u_2}\|^2}. \end{cases} \tag{6}$$

Repeating this process r times, r is the rank of A with $r \le \min(n-1, m)$. Then, there exist r components $u_1, u_2, \cdots u_r$ such that

$$\begin{cases} A = \widehat{u_1}\sigma^{(1)T} + \widehat{u_2}\sigma^{(2)T} + \cdots + \widehat{u_r}\sigma^{(r)T} + A_r, \\ B = \widehat{u_1}\tau^{(1)T} + \widehat{u_2}\tau^{(2)T} + \cdots + \widehat{u_r}\tau^{(r)T} + B_r. \end{cases} \tag{7}$$

In the end, based on the relationships of the above variables, the PLSR equation is obtained as

$$Y = XC + F, \tag{8}$$

where C and F can be computed. The above processes are the major computing steps of PLSR in [1]. In this study, the monthly GWSA is the matrix Y, and the other five monthly hydroclimate variables (P, T, E, SM and SWE) are the dependent variables matrix X. We aim to develop PLSR models to estimate GWSA for Xinjiang and the five sub-regions, which are also used to reveal the major influencing variables of the changes in groundwater. Moreover, the differences of the major influencing variables in the five sub-regions are also discussed to understand the hydro-climatic mechanisms of groundwater changes in the study area.

Text S2 Discussion on the accuracy of the derivation of GWSA based on GRACE

The method used in this study to derive GWSA based on GRACE has be evaluated by in-situ observations from monitoring well stations in various data-rich regions across the world, demonstrating the reliability of the method [2]. However, the groundwater level observations are extremely limited in the study area, making a direct comparison of the estimated GWSA against groundwater observations difficult [3-5]. To the performance of the derivation method of GWSA in the study area, we here discuss the accuracy of each hydrological components used in this method, including P, T, E, TWSA, SM, and SWE (Table 1).

Accuracy of P and T

The P and T datasets used in this study are the gridded observations at the spatial resolution of 0.5°×0.5° from the China Meteorological Administration (http://data.cma.cn/site/index.html). These two gridded datasets were developed by interpolating observations from 2,472 meteorological stations over China based on the Thin Plate Spline (TPS) interpolation method. Strict quality control and generalized cross-validation test have been conducted for the datasets. The datasets can reasonably describe the spatiotemporal characteristics of precipitation and temperature and in Northwest China [6].

Accuracy of E, SM, and SWE

The dataset of E applied in this study is the Global Land Evaporation Amsterdam Model V3.1a (https://www.gleam.eu/). The GLEAM was developed based on a set of algorithms that separately estimate different components of terrestrial ET based on satellite observations, including transpiration, interception loss, bare-soil ET, snow sublimation and open-water ET [7]. Against the *in-situ measurements*, the GLEAM V3 dataset shows high accuracy with a correlation coefficient value larger than 0.8 [7].

The SM dataset we used is GLDAS Noah 2.1 (https://search.earthdata.nasa.gov/). The SM of GLDAS Noah 2.1 has been widely used in many previous studies in the field of hydrology [5,8,9]. The accuracy of Noah 2.1 SM data has been evaluated against the in-situ and remotely sensed observations over China and the globe [9,10]. Previous studies have shown that GLDAS Noah SM is in agreement with the observations in China, although there are no observed stations of SM in Xinjiang [9].

Due to the lake of the *in-situ* observations, the accuracy of SWE obtained from GLDAS Noah 2.1 is not evaluated in this study. However, the SWE data of GLDAS Noah 2.1 has been used in detecting the groundwater changes over other regions [2,5,11].

Although we do not have observed SM and SWE, we further conduct an uncertainty analysis among various SM and SWE outputs based on [12] in the revised manuscript. SM and SWE from five GLDAS models, namely CLM, Mosaic, VIC, GLDAS 1 Noah 2.7 (hereafter Noah 2.7) and GLDAS 2.1 Noah 3.3 (hereafter Noah 3.3) are used to calculate the standard deviations (STDs) of the trends from the five GLDAS models as trend uncertainties. The trend uncertainties reflect discrepancies of the slopes estimated from the five models.

The spatial distributions of the SM linear trends of the five GLDAS models are provided in Figure S3. Mosaic, VIC, Noah 2.7, and Noah 3.3 have positive linear trends over the same regions (Figures S3B-S3E), except the CLM model (Figure S3A). The spatial distributions of the multi-model average of the trends and the STDs of the five GLDAS models are displayed in Figure S4. Most of the STDs are smaller than 1 mm/month which indicates the uncertainties of SM from different GLDAS models are small (Figure S4B).

For SWE, CLM, Mosaic, VIC, and Noah 2.7 have similar spatial distributions of the linear trends (Figures S5B-S5E). Although there are some differences between the above four models and Noah 3.3, the linear trend values of the five models are very small, and most of them are insignificant (Figure S5). The STDs of the linear trends of the five models are smaller than 0.25 mm/month over most parts of Xinjiang (Figure S6).

Accuracy of TWSA and GWSA

The TWSA of GRACE data has been used to detect the water storage changes over Xinjiang successfully [3,4], and the GWSA data based on GRACE and GLDAS has been widely employed to detect the groundwater change in north China [5,13] and India [11]. Therefore, we have high confidence in employing this method to evaluate groundwater variations in Xinjiang.

To further evaluate the robustness of the results, we conduct an uncertainty analysis of GWSA derived from different GLDAS models: CLM, Mosaic, VIC, Noah 2.7, and Noah 3.3. To avoid the impacts of choice of GRACE data on GWSA, we use monthly EM-GRACE to estimate GWSA in the uncertainty analysis.

The spatial distributions of the GWSA from the five models are provided in Figure S7. All models indicate negative trends of the GWSA in JGB, and positive trends in the east KLM (Figure S7). Compared to CLM and Noah 2.7, the spatial distributions of Mosaic and VIC have better agreement with Noah 3.3 with negative linear trends over the southern part of Xinjiang (Figure S7). Figure S8 shows the mean trends and the STDs of the GWSA linear trends based on the five GLDAS models and EM-GRACE. The mean trends also show the negative linear trends across Xinjiang except the positive linear trends over KLM. These results indicate the groundwater depletion in a majority of our study area (Figure S8A). The mountainous areas have large STDs, which shows the larger uncertainties in mountainous areas than plain areas caused by the complex hydrological process (Figure S8B).

Here we also try to validate the GRACE-based GWSA based on very limited amount of observed data available, including annual regional-average observed groundwater recharge (OBS-GWR) and annual total water use in Xinjiang during 2003-2015 from the Xinjiang Water Resources Bulletin (http://www.xjslt.gov.cn/zwgk/slgb/index.html), as well as observed groundwater depth (GWD) of the monitoring well station (41.79N°, 81.62E°) in Kaidu-Konqi River basin during 2004-2010 from the Department of Water Resources of Xinjiang Uygur Autonomous Region (http://www.xjslt.gov.cn). Figure S9 shows a basin-scale comparison between the EM-GWSA and the observed GWD in the Kaidu-Kongqi basin in 2004-2009. The EM-GWSA data is extracted from the grid cell where the wells are located. There exists a perfect match between EM-GWSA and GWD with a CC value of 0.96 ($R^2 = 0.91, p < 0.01$). These comparisons with observations confirm that GRACE-based GWSA can reasonably capture the groundwater variations. GWSA is jointly affected by groundwater recharge, water withdrawal, and other factors. Figure S4A shows that the temporal variation of regional-average EM-GWSA resembles that of the groundwater recharge across Xinjiang in 2003-2015. In general, EM-GWSA decreases as groundwater recharge gets lower from 2003-2009, showing that EM-GWSA can reflect the decreases in groundwater recharge although the CC is only 0.25 because groundwater variations are also jointly affected by other factors. As shown in Figure S10B, the EM-GWSA is negatively correlated (CC=-0.85) with the total water use in Xinjiang during 2003-2015 ($R^2 = 0.73, p < 0.01$), showing the importance of water withdrawal in groundwater changes. From 2003-2010, the total water use increased from 49.44 billion m^3 to 53.51 billion m^3 accompanied with the decrease of the GWSA from 9.65mm (15.44 billion m^3) to -25.25mm (-40.4 billion m^3) over the whole Xinjiang. After a small decrease of the total water use in 2011 (i.e. 52.35 billion m^3), the averaged total water use jumped to 58.43 billion m^3 between 2012 and 2015. A persistent decrease in GWSA is detected in that period except an increase in 2014. Although the validation is not comprehensive given the constraints of data availability, it suggests the temporal variations of EM-GWSA match with the changes in measured groundwater depth as well as the affecting factors, i.e. groundwater recharge and total water use.

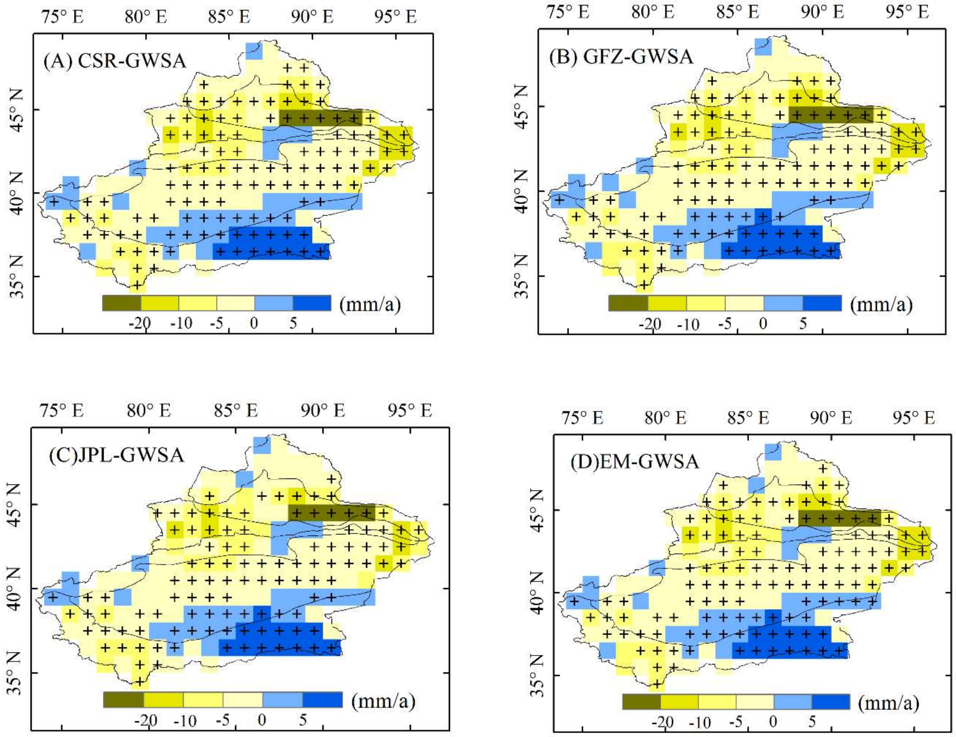

Figure S1. Spatial distributions of the linear trends (mm/month) of the JJA GWSA during 2003-2016. The cross signs denote the trends are significant at the 95% significance level.

.

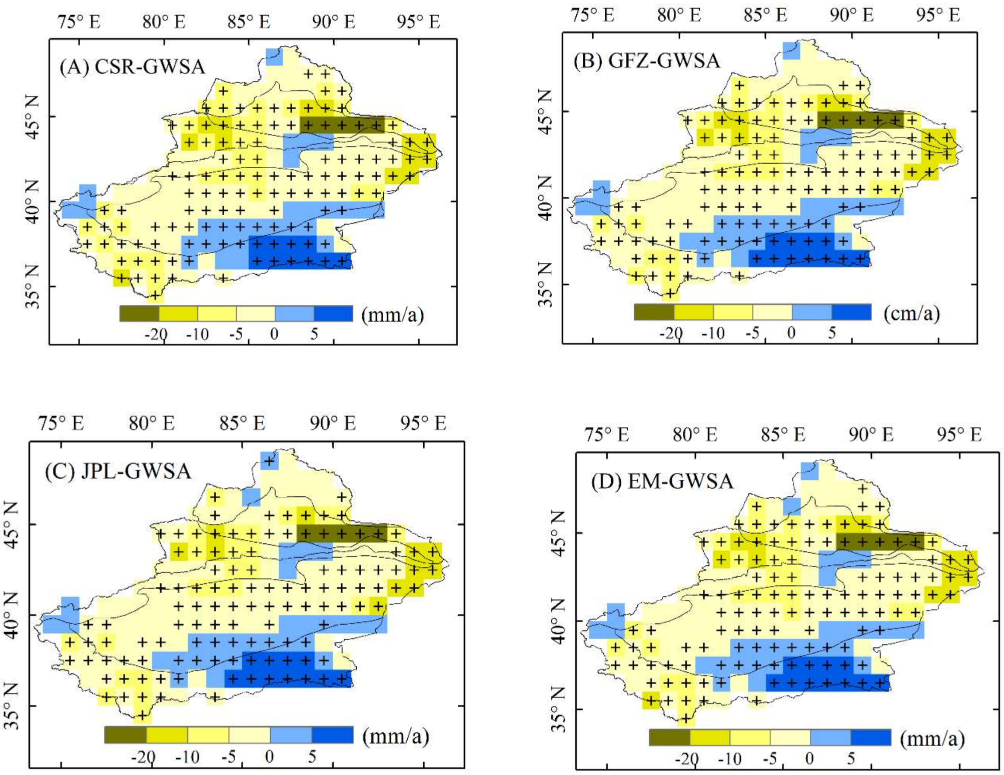

Figure S2. Same as Figure S1 but for annual GWSA.

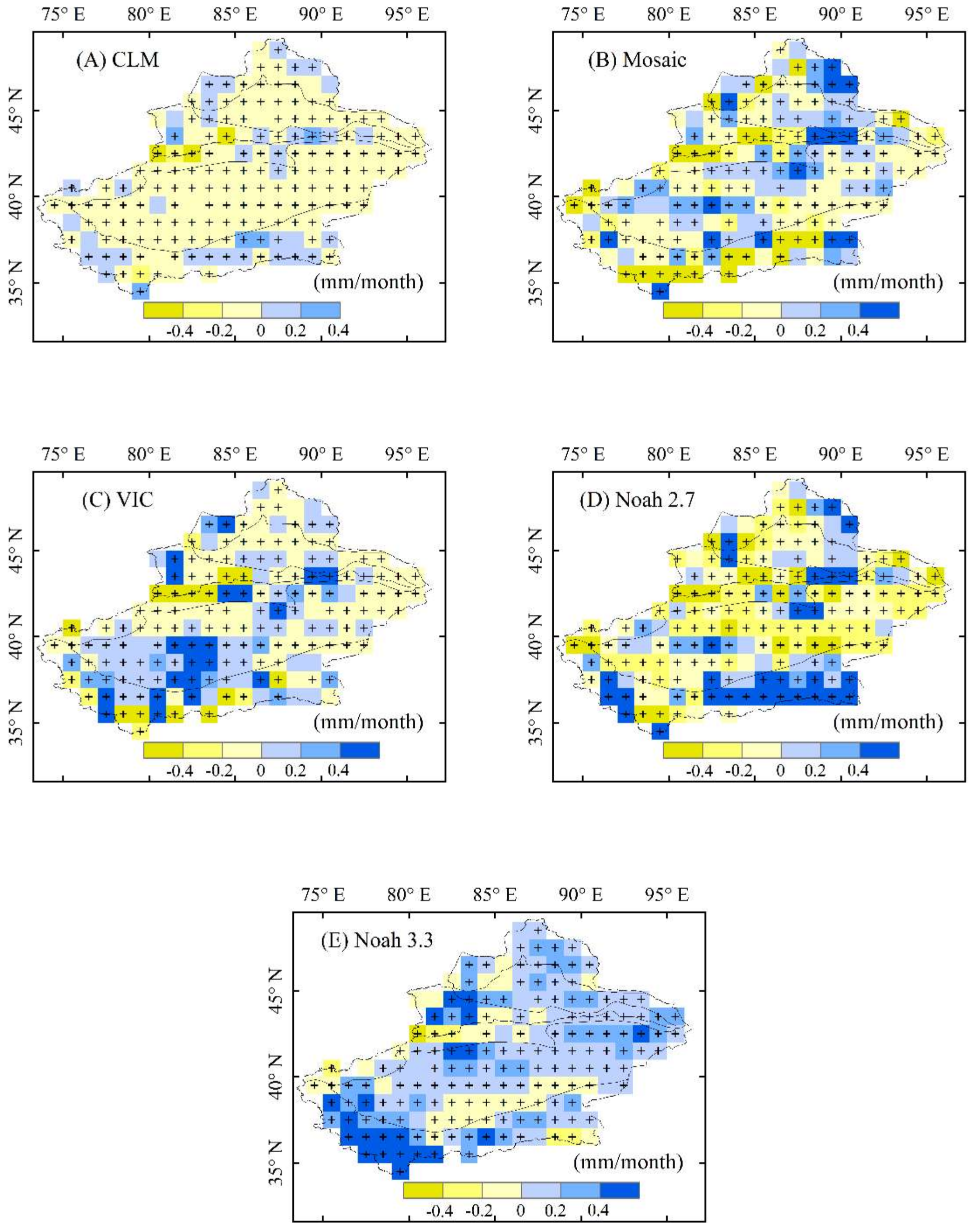

Figure S3: Spatial distribution of linear trends of monthly SM derived from CLM, Mosaic, VIC, Noah 2.7 and Noah 3.3 over Xinjiang during 2003-2016. The cross signs denote the trends are significant at the 95% significance level.

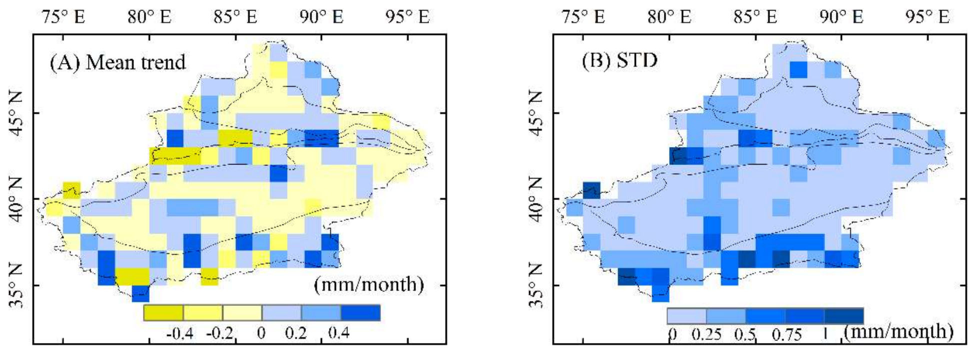

Figure S4: The multimodel average of the linear trends of monthly SM derived from CLM, Mosaic, VIC, Noah 2.7 and Noah 3.3 over Xinjiang during 2003-2016 (A), and the corresponding STD (standard deviation) indicating the uncertainty of SM (B).

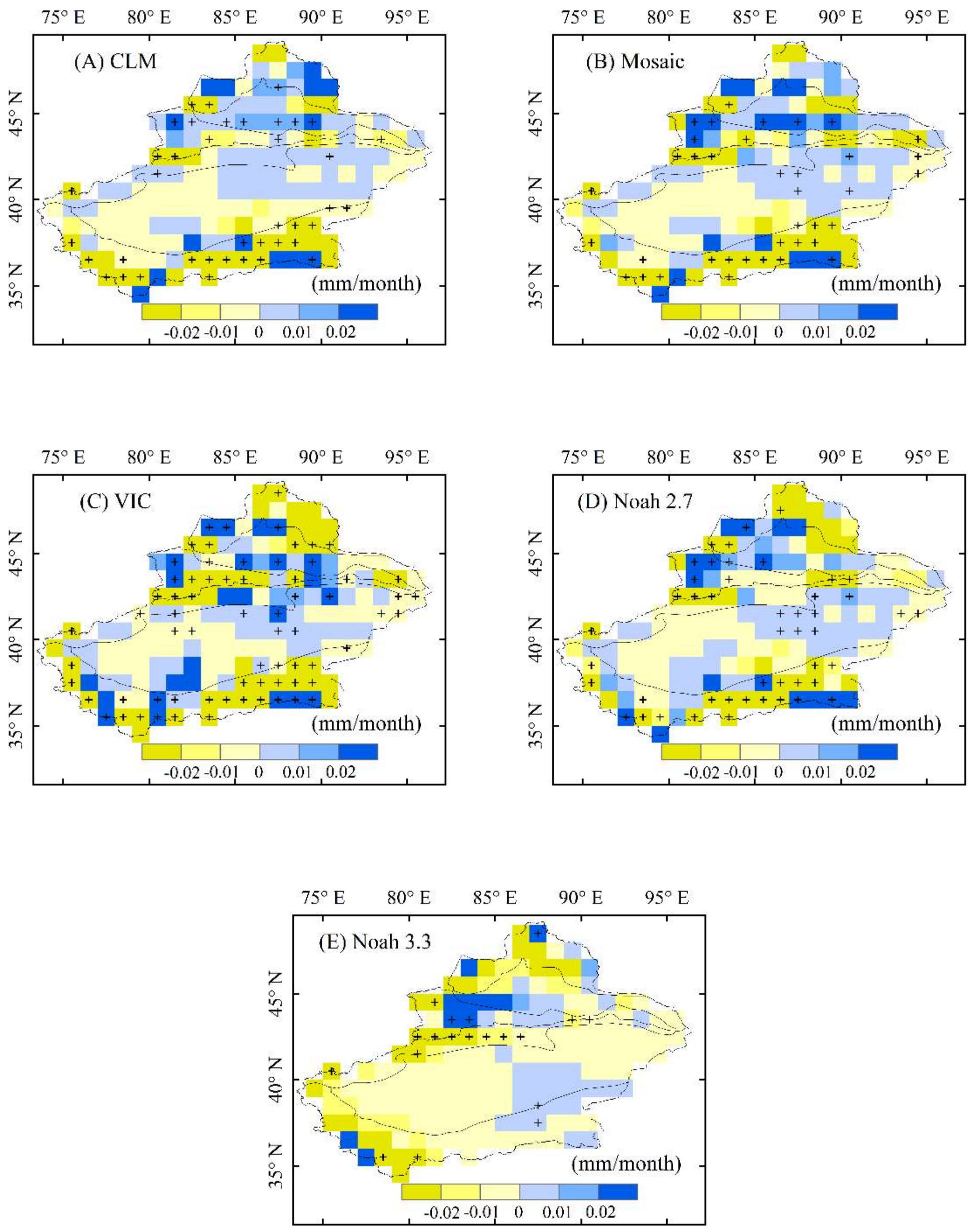

Figure S5: Spatial distribution of linear trends of monthly SWE derived from CLM, Mosaic, VIC, Noah 2.7 and Noah 3.3 over Xinjiang during 2003-2016. The cross signs denote the trends are significant at the 95% significance level.

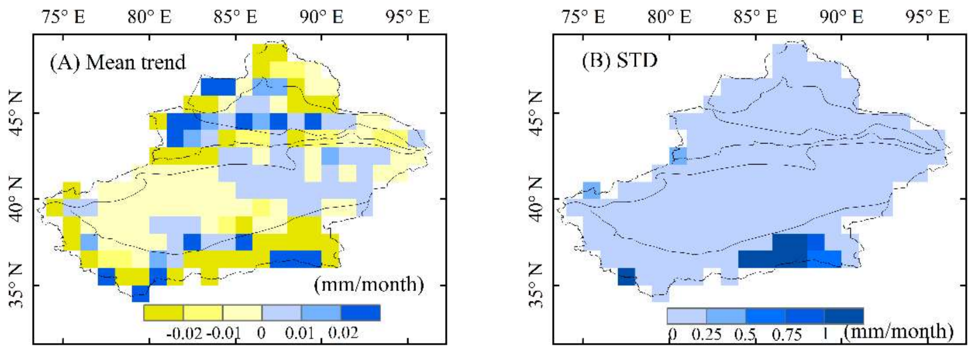

Figure S6: The multimodel average of the linear trends of monthly SWE derived from CLM, Mosaic, VIC, Noah 2.7 and Noah 3.3 over Xinjiang during 2003-2016 (A), and the corresponding STD (standard deviation) indicating the uncertainty of SM (B).

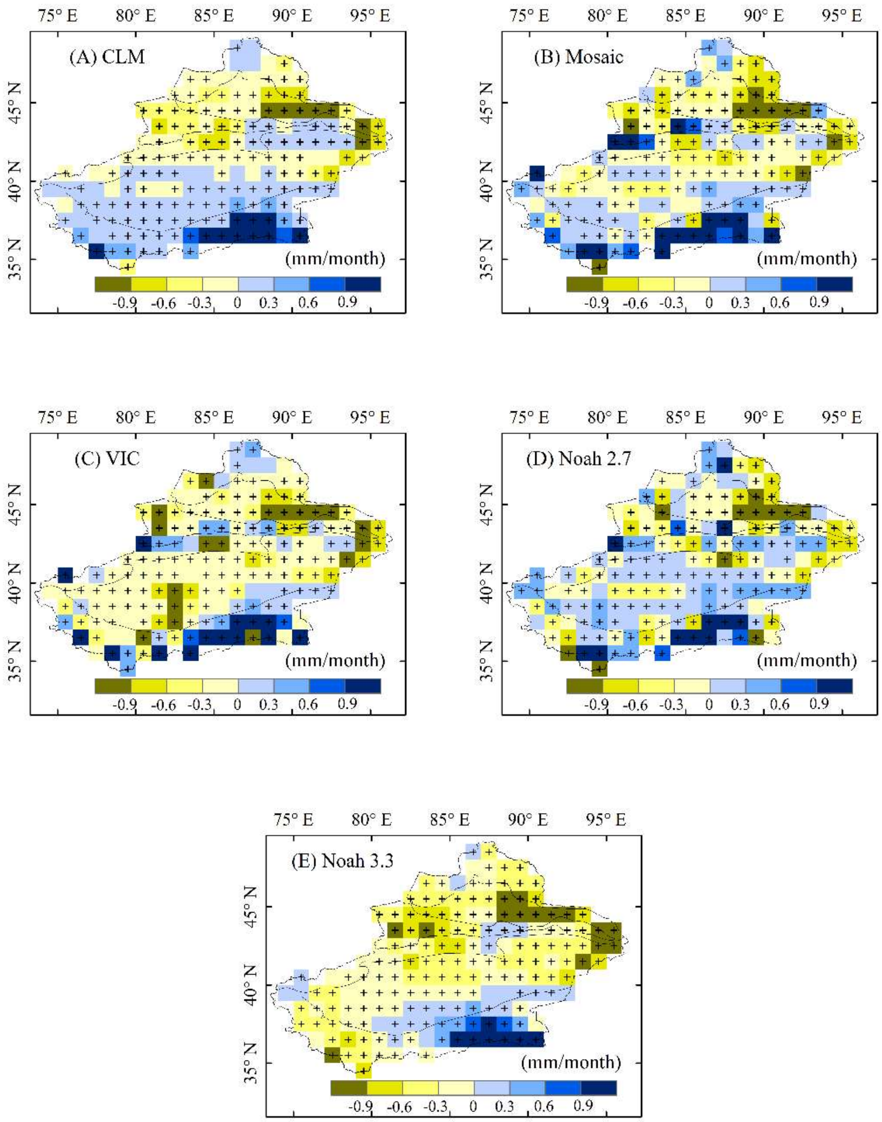

Figure S7: Spatial distribution of linear trends of monthly GWSA derived from CLM, Mosaic, VIC, Noah 2.7 and Noah 3.3 over Xinjiang during 2003-2016. The monthly EM-TWSA is used here. The cross signs denote the trends are significant at the 95% significance level.

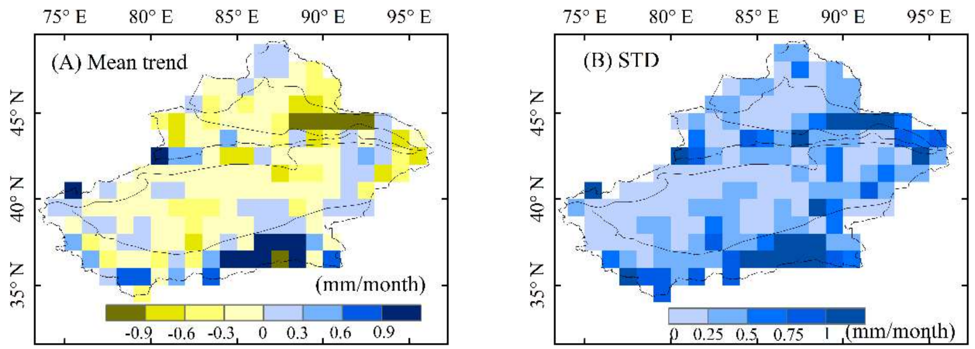

Figure S8: The multimodel average of the linear trends of monthly GWSA derived from CLM, Mosaic, VIC, Noah 2.7 and Noah 3.3 over Xinjiang during 2003-2016 (A), and the corresponding STD (standard deviation) indicating the uncertainty of GWSA (B). The monthly TWSA data is from EM-TWSA.

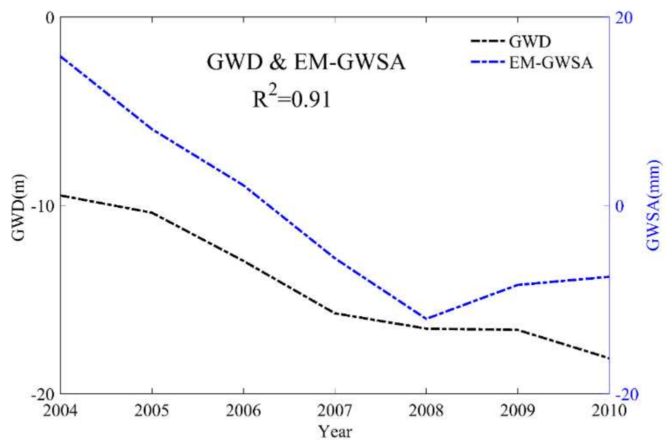

Figure S9. Comparison between observed groundwater depth (GWD) and EM-GWSA in Kaidu-Konqi River basin during 2004-2010.

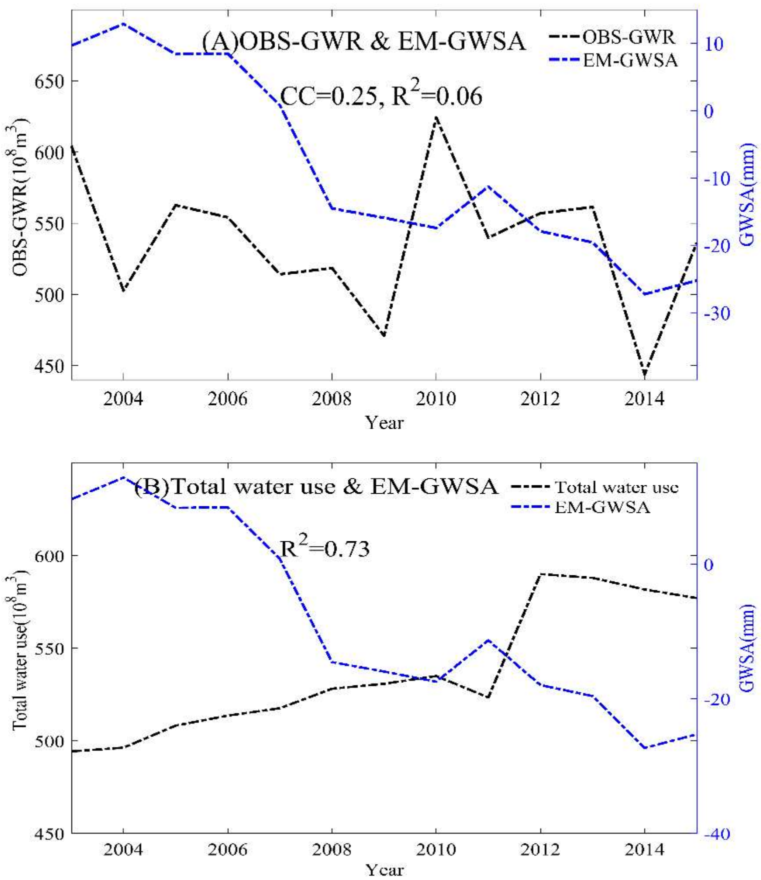

Figure S10. Comparisons between observed groundwater recharge (OBS-GWR) and EM-GWSA (A), and between total water use and EM-GWSA (B) in Xinjiang during 2003-2015.

Table S1. Linear trends of climate variables at monthly, seasonal and annual scales during 1961-2016. ** denotes the trend is significant at the 95% or 99% significance level. ± values are the 5% and 95% confidence intervals.

Dataset	Timescale	2003-2016	1980-2016	1961-2016
T (℃/month for monthly data, ℃/a for annual and seasonal data)	Monthly	0.005	0.004	0.002
	MAM	0.051	0.053**±0.026	0.025**±0.014
	JJA	0.046	0.043**±0.013	0.026**±0.007
	SON	-0.022	0.040**±0.022	0.032**±0.012
	DJF	0	0.008	0.03**±0.018
	Annual	0.022	0.038**±0.013	0.029**±0.007
P (mm/month for monthly data, mm/a for annual and seasonal data)	Monthly	0.018	0.007	0.006**±0.004
	MAM	0.08	0.16	0.13*±0.10
	JJA	1.48	0.49*±0.37	0.40**±0.18
	SON	0.77	0.20**±0.18	0.16**±0.09
	DJF	0.04	0.18**±0.11	0.12**±0.05
	Annual	2.30	1.01**±0.58	0.81**±0.28
E (unite same as P)	Monthly	0.009	0.004	-
	MAM	0.13	0.12	-
	JJA	1.15	0.39**±0.28	-
	SON	0.31	0.08	-
	DJF	-0.02	0.004	-
	Annual	1.55	0.60**±0.42	-
SM (unite same as P)	Monthly	0.15**±0.02	-	-
	MAM	1.43**±0.86	-	-
	JJA	1.86**±1.22	-	-
	SON	2.26**±1.35	-	-
	DJF	1.39**±1.87	-	-
	Annual	1.82**±0.93	-	-
SWE (unite same as P)	Monthly	-0.01	-	-
	MAM	-0.1	-	-
	JJA	0.01	-	-
	SON	0.05	-	-
	DJF	-0.04	-	-
	Annual	-0.01	-	-

Table S2. Change rates of TWSA derived from GRACE at different time scales (mm/month for monthly scale, mm/a for seasonal and annual scales) during 2003-2016. ** denotes the trend is significant at the 95% or 99% significance level. ± values are the 5% and 95% confidence intervals.

Timescale	**CSR**	**GFZ**	**JPL**	**EM**
Monthly	-0.21**±0.07	-0.17**±0.07	-0.16**±0.08	-0.18**±0.07
MAM	-2.31**±1.48	-2.96**±1.74	-2.68*±2.44	-2.20*±1.76
JJA	-1.43**±1.94	-3.26**±1.57	-1.26	-1.35
SON	-2.64*±2.13	-4.94**±2.01	-1.37	-2.21*±1.87
DJF	-3.96**±2.40	-3.61**±2.04	-2.45	-2.89*±2.29
Annual	-2.33**±1.61	-3.70**±1.14	-1.74	-1.99*±1.60

References:

1 Wold, S.; Sjostrom Mand Eriksson, L. PLS-regression: A basic tool of chemometrics. Chemom. Intel. Lab. Syst. 2001, 58, 109–130.

2 Feng, W.; Zhong, M. and Lemoine, J.; Evaluation of groundwater depletion in North China using the Gravity Recovery and Climate Experiment (GRACE) data and ground-based measurements. Water Resources Research, 2013, 49, 2110-2118.

3 Yang, T.; Wang, C. and Chen, Y., et al., Climate change and water storage variability over an arid endorheic Region. Journal of Hydrology, 2015, 529, 330-339.

4 Deng, H.; Chen, Y.; Influences of recent climate change and human activities on water storage variations in Central Asia. Journal of Hydrology, 2017, 544, 46-57.

5 Feng, W.; Shum, C. and Zhong, M.; Groundwater Storage Changes in China from Satellite Gravity: An Overview. Remote Sensing, 2018, 10, 674.

6 Zhao, Y.; Zhu, J. and Xu, Y.; Establishment and assessment of the grid precipitation datasets in China for recent 50 years. Journal of the Meteorological Sciences, 2014, 34, 414-420.

7 Martens, B.; Miralles, D. and Lievens, H., et al., GLEAM v3: satellite-based land evaporation and root-zone soil moisture. Geosci. Model Dev., 2017, 10, 1903-1925.

8 Mukherjee, A. and Ramachandran, P.; Prediction of GWL with the help of GRACE TWS for unevenly spaced time series data in India: Analysis of comparative performances of SVR, ANN and LRM. Journal of Hydrology, 2018, 558, 647-658.

9 Gu, X., Zhang, Q. and Li, J., et al., Attribution of global soil moisture drying to human activities: A quantitative viewpoint. Geophysical Research Letters, 2019, 46, 2573-2582.

10 Gu, X.; Li, J.; Chen, Y.; Kong, D.; Liu, J.; Consistency and discrepancy of global surface soil moisture changes from multiple model‐based data sets against satellite observations. Journal of Geophysical Research: Atmospheres, 2019. 124, 1474-1495.

11 Rodell, M.; Velicogna, I. and Famiglietti, J.; Satellite-based estimates of groundwater depletion in India. Nature, 2009, 460, 999-1002.

12 Scanlon, B.; Zhang, Z. and Save, H., et al., Global models underestimate large decadal declining and rising water storage trends relative to GRACE satellite data. PNAS, 2018, E1080-E1089.

13 Huang, Z.; Pan, Y. and Gong, H., et al., Subregional-scale groundwater depletion detected by GRACE for both shallow and deep aquifers in North China Plain. Geophys. Res. Lett. 2015, 42, 1791-1799.