



1 **Improving estimates of water resources in a semi-arid region by assimilating GRACE**
2 **data into the PCR-GLOBWB hydrological model**

3 N. Tangdamrongsub^{1,2}, S. C. Steele-Dunne³, B. C. Gunter^{1,4}, P. G. Ditmar¹, E. H.
4 Sutanudjaja⁵, Y. Sun¹, T. Xia⁶, and Z. Wang^{6,7}

5 [1] Department of Geoscience and Remote Sensing, Faculty of Civil Engineering and
6 Geosciences, Delft University of Technology, Delft, The Netherlands

7 [2] School of Engineering, Faculty of Engineering and Built Environment, The University of
8 Newcastle, Callaghan, New South Wales, Australia

9 [3] Department of Water Resources, Faculty of Civil Engineering and Geosciences, Delft
10 University of Technology, Delft, The Netherlands

11 [4] School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, The United
12 States of America

13 [5] Department of Physical Geography, Faculty of Geosciences, Utrecht University, Utrecht,
14 The Netherlands

15 [6] Department of Hydraulic Engineering, Tsinghua University, Beijing 100084, China

16 [7] State Key Lab of Hydrosience and Engineering, Tsinghua University, Beijing 100084,
17 China

18

19 Correspondence to: N. Tangdamrongsub (Natthachet.tangdamrongsub@newcastle.edu.au)

20

21 **Abstract**

22 An accurate estimation of water resources dynamics is crucial for proper management of both
23 agriculture and the local ecology, particularly in semi-arid regions. Imperfections in model
24 physics, uncertainties in model land parameters and meteorological data, as well as the human
25 impact on land changes often limit the accuracy of hydrological models in estimating water
26 storages. To mitigate this problem, this study investigated the assimilation of Terrestrial
27 Water Storage (TWS) estimates derived from the Gravity Recovery And Climate Experiment
28 (GRACE) data using an Ensemble Kalman Filter (EnKF) approach. The region considered
29 was the Hexi Corridor of Northern China. The hydrological model used for the analysis was
30 PCR-GLOBWB, driven by satellite-based forcing data from April 2002 to December 2010. In
31 this study, EnKF 3D scheme, which accounts for the GRACE spatially-correlated errors, was
32 used. The correlated errors were propagated from the full error variance-covariance matrices
33 provided as a part of the GRACE data product. The impact of the GRACE Data Assimilation
34 (DA) scheme was evaluated in terms of the TWS, as well as individual hydrological storage
35 estimates. The capability of GRACE DA to adjust the storage level was apparent not only for
36 the entire TWS but also for the groundwater component, which had annual amplitude, phase,
37 and long-term trend estimates closer to the GRACE observations. This study also assessed the
38 benefits of taking into account correlations of errors in GRACE-based estimates. The
39 assessment was carried out by comparing the EnKF results, with and without taking into
40 account error correlations, with the in situ groundwater data from 5 well sites and the in situ
41 streamflow data from two river gauges. On average, the experiments showed that GRACE
42 DA improved the accuracy of groundwater storage estimates by as much as 25%. The



43 inclusion of error correlations provided an equal or greater improvement in the estimates. No
44 significant benefits of GRACE DA were observed in terms of streamflow estimates, which
45 reflect a limited spatial and temporal resolution of GRACE observations. Results from the 9-
46 year long GRACE DA study were used to assess the status of water resources over the Hexi
47 Corridor. Areally-averaged values revealed that TWS, soil moisture, and groundwater
48 storages over the region decreased with an average rate of approximately 0.2, 0.1, and 0.1
49 cm/yr in terms of equivalent water heights, respectively. A substantial decline in TWS
50 (approximately -0.4 cm/yr) was seen over the Shiyang River Basin in particular, and the
51 reduction mostly occurred in the groundwater layer. An investigation of the relationship
52 between water resources and agriculture suggested that groundwater consumption required to
53 maintain the growing period in this specific basin was likely the cause of the groundwater
54 depletion.

55

56 **1. Introduction**

57 Semi-arid regions can be broadly classified as areas on the boundaries of larger deserts,
58 receiving just enough annual precipitation (300 mm or less) to sustain a limited amount of
59 agriculture. Inefficient use of the limited amount of surface water can often lead to overuse of
60 groundwater resources and salinization of the soil. This can result in desertification, which not
61 only reduces the amount of production but also may have long-term effects on the local
62 ecology. Improving the water resources management of such regions requires accurate
63 knowledge of the hydrological processes involved. For small areas, this can be partially
64 obtained through a network of in-situ measurement systems, such as meteorological stations,
65 river gauges, groundwater wells, evaporation trays, etc. (Dahlgren & Possling, 2007; Huo et
66 al., 2007; Kang et al., 2004; Ma et al., 2005; Du et al., 2014). However, as point
67 measurements, these observations are very local in scope. A sensor at a point several
68 kilometres away may record significantly different values. For large scales ($> 10,000$ km²),
69 such techniques are unlikely capable of delivering accurate results.

70 Two options for estimating the variations of the large-scale Terrestrial Water Storage (TWS)
71 of a particular region are: using observations from the Gravity Recovery And Climate
72 Experiment satellite mission (GRACE, Tapley et al., 2004) or utilizing a regional or global
73 hydrological model. A number of prior studies have reported on the potential of GRACE in
74 the estimation of snow water equivalent (Niu et al., 2007), groundwater (Döll et al., 2014),
75 and evapotranspiration (Long et al., 2014) in terms of temporal and spatial variability.
76 However, GRACE only provides the total column of the water storage at a monthly time scale
77 and large spatial scales (> 300 km). Furthermore, it is not possible to identify the contribution
78 of separate hydrological components to the TWS from GRACE data alone. On the other hand,
79 a hydrological model can be used to estimate the individual storage components at very high
80 spatial and temporal scale. The major drawback of the model is mainly the significant
81 uncertainties influenced by the quality of the model parameter calibration and the accuracy of
82 the meteorological input data.



83 Data Assimilation (DA) can be employed to combine the strengths of GRACE and
84 hydrological models while mitigating their respective weaknesses. A number of studies show
85 that GRACE DA can be used to improve the estimation of groundwater and streamflow
86 (Zaitchik et al., 2008; Tangdamrongsub et al., 2015), snow water equivalent (Forman et al.,
87 2012; Su et al., 2012), and as well as for evaluation of drought events (Houborg et al., 2012;
88 Li et al., 2012). Different temporal and spatial resolution of GRACE observations and
89 hydrological models require proper design of the DA scheme. Several DA schemes have been
90 developed to distribute GRACE observations into the model, which include assuming a
91 uniform observation value available every 10 days and updating the model every 10 days
92 (Zaitchik et al., 2008), using 5-day interpolated observations and updating the model every 5
93 days (Tangdamrongsub et al., 2015), using a monthly observation value and applying the
94 model update only at the end of the month (Eicker et al., 2014), and using a monthly
95 observation value and distributing the update as a daily increment (Forman et al., 2012).
96 Although all DA schemes are acceptable, the scheme proposed by Forman et al. (2012) is
97 advantageous because it does not require an interpolation of the observations and can reduce
98 the spurious jump caused by applying the update at the end of the month only. As such, a
99 scheme similar to (Forman et al., 2012) is used in this study. Spatial disaggregation is also
100 needed to reconcile the difference in horizontal resolution between the observations and the
101 model. Recent studies by Forman et al. (2013) and Eicker et al. (2014) suggested including
102 the GRACE variance-covariance error information in the spatial disaggregation step. Forman
103 et al. (2013) recommended using the finest observation resolution where observation error is
104 considered uncorrelated, while Eicker et al. (2014) proposed using 500-km GRACE spatial
105 resolution to mitigate the ill-posedness of the error covariance matrices in the spatial domain.
106 In line with Forman et al. (2013) and Eicker et al. (2014), the assimilation scheme in this
107 study accounts for spatially correlated errors by using full error variance-covariance matrices
108 of GRACE data. This study will show that considering the GRACE error correlations leads to
109 an improvement of the state estimates.

110 The semi-arid region investigated in this study, the Hexi Corridor, is located between the
111 Gansu province of China and Mongolia (Fig. 1). In the past 50 years, over-utilization of the
112 groundwater has significantly reduced the groundwater supply and quality there. This has
113 resulted in increased soil salinization, to the extent that desertification has become a genuine
114 threat to both agriculture and the natural ecology (Wang et al., 2003). Due to a small size of
115 the Hexi Corridor area, the Signal-to-Noise Ratio (SNR) of the TWS variations estimated
116 there on the basis of GRACE data is much lower than in river basins considered in the
117 previous studies, e.g., Mississippi (Zaitchik et al., 2008), Rhine (Tangdamrongsub et al.,
118 2015), and Mackenzie (Forman et al., 2012). GRACE observations are assimilated into the
119 PCRaster Global Water Balance (PCR-GLOBWB; Van Beek et al., 2011; Sutanudjaja et al.,
120 2014; Wada et al., 2014) hydrological model over the Hexi Corridor. TWS is computed from
121 PCR-GLOBWB as the sum of all the hydrological components (soil moisture, groundwater,
122 surface water, inundated water, interception, and snow). The previous studies showed very
123 strong correlations of PCR-GLOBWB based estimates with GRACE observations in several
124 river basins (Wada et al., 2014; Tangdamrongsub et al., 2016). However, the performance of
125 PCR-GLOBWB has not been evaluated over the Hexi Corridor yet. In addition, the model has



126 not been incorporated into any GRACE DA scheme, making this study the first attempt to do
127 so.

128 The results of the GRACE DA approach are validated with independent in-situ and remote
129 sensing data. The groundwater storage (GWS) and streamflow estimates are validated with
130 the well and river stream gauge measurements, respectively. The precipitation from the
131 Tropical Rainfall Measuring Mission (TRMM; Huffman et al., 2007) and the Moderate
132 Resolution Imaging Spectroradiometer (MODIS) derived Normalized Difference Vegetation
133 Index (NDVI; Huete et al., 2002) are used to investigate the connection between agricultural
134 activity and the groundwater consumption in the area.

135 The main objective of this study is to investigate the added value of GRACE DA in the Hexi
136 Corridor. Approximately 9 years of data – between April 2002 and December 2010 – are
137 considered. The performance of the GRACE DA scheme is evaluated in terms of its impact
138 on the total terrestrial water storage as well as on the individual hydrological storage
139 estimates. The impact of taking into account correlation in GRACE errors is also assessed.
140 This will be shown to improve the storage estimates, particularly groundwater. Finally, results
141 from this 9-year long GRACE DA study are used to assess the status of water resources over
142 the Hexi Corridor. The connections between the water storage, especially groundwater, and
143 agriculture in the area are also presented and discussed.

144

145 **2. Study region**

146 The Hexi Corridor is a long and narrow area between the Qilian Mountain range and southern
147 Mongolia (Fig. 1a). The basin's elevation ranges from 5,200 m in the southern upstream area
148 (Qilian Mountains) to 900 m in the northern downstream zone (Inner Mongolia) (Fig. 1b).
149 The region is comprised of four typical inland arid and semi-arid regions: the Shiyang River
150 Basin (41,600 km²), the Heihe River Basin (143,000 km²), the Shule River Basin (157,000
151 km²), and a Desert Region (152,445 km²) (Geng and Wardlaw, 2013; Zhu et al., 2015).
152 Located next to the Gobi Desert, most parts of the region have a cold desert climate, where
153 precipitation is relatively low to sustain vegetation or crops. Approximately 60 to 80 % of the
154 annual rainfall is concentrated during the timeframe from June to September. The inland
155 rivers mainly originate from the Qilian Mountains and disappear after entering the
156 midstream/downstream plains and oases. As such, the southern part of the region is more
157 favourable for agriculture.

158 The four basins have distinct characteristics. First, the smallest river basin, Shiyang, has 8
159 main river streams, including the Xida and Xiying Rivers (Fig. 1c). The annual rainfall and
160 the mean temperature are approximately 250 mm and 5 °C (Fig. 2a, b), respectively. The
161 Shiyang River Basin is considered the wettest basin compared to the others, with relatively
162 high mean total renewable annual water resources of approximately 1.66 billion m³ (Zheng et
163 al., 2013). However, a highly developed economy and population growth in the past decade
164 have resulted in a severe water resources overexploitation problem (Zheng et al., 2013). The
165 Heihe River Basin has a semi-arid climate and the mean daily temperature of ~6 °C (Fig. 2d).



166 The average annual rainfall is ~150 mm (Fig. 2c) with high heterogeneity both in temporal
167 and spatial distribution. The mean total annual available water resources are estimated at 3.7
168 billion m³ (Hu, 2015). Similar to the Shiyang River Basin, increased water exploitation,
169 increasing population, and changing climate have aggravated the damage to the downstream
170 ecology. The Shule River Basin has an arid climate, the mean temperature there is around 4
171 °C (Fig. 2f), and the average annual rainfall is only approximately 98 mm (Fig. 2e).
172 Compared to the Shiyang River Basin, the Shule River Basin is approximately four times as
173 large in terms of surface area, but has similar mean total annual water resources, ~1.6 billion
174 m³ (Hu, 2015). The district irrigation areas are mainly located in the middle of the Shule
175 River Basin. Agricultural water consumption accounts for more than 80% of the total water
176 use. Finally, the Desert Region has an extreme continental desert climate with an average
177 temperature of 8 °C, and the annual rainfall of ~130 mm. Extensive groundwater abstraction
178 was also observed over the region (Jiao et al., 2015).

179

180 3. Hydrology model

181 The global hydrological model PCR-GLOBWB (van Beek et al., 2011; Sutanudjaja et al.,
182 2016) simulates spatial and temporal continuous fields of fluxes and storages in various water
183 storage components (soil moisture, groundwater, surface water, inundated water, interception,
184 and snow). The model version used here (Sutanudjaja et al., 2016) has a spatial resolution of
185 30 arc minutes (approximately 50 km at the equator). It includes an interactive module
186 simulating water availability and water abstraction, including variations in irrigation and other
187 sectoral water demands (e.g. livestock, industrial and domestic water demands) and dynamic
188 allocation of available surface water and groundwater resources (see also, e.g., de Graaf et al.,
189 2014; Wada et al., 2014; Sutanudjaja et al., 2015). This feature determines how the demands
190 are allocated to the available water resources and where the return flows of unconsumed water
191 go. Modelling this feedback is essential, particularly in irrigated areas where water demand is
192 large. Details of the water demand calculation in PCR-GLOBWB can be found in Appendix
193 A.

194 Figure 3 illustrates the structure of PCR-GLOBWB model. The model includes 2 soil layers
195 (SM_{upp} , SM_{low}), an underlying hydrologically active and replenishable groundwater
196 (GWS_{active}) layer, a non-renewable groundwater (GWS_{fossil}) layer, as well as an interception,
197 surface water, and snow stores. The non-renewable groundwater is available for abstraction to
198 satisfy water demands once the overlying hydrologically active groundwater storage is
199 depleted. For soil, snow, inundated top water, and interception stores, an individual grid cell is
200 divided into sub-grids associated with different types of topography, vegetation phenology,
201 and soil properties, as well as land cover types. Specifically, there are 4 types of land covers
202 defined: short natural vegetation, tall natural vegetation, irrigated non-paddy field, and
203 irrigated paddy field. Soil components include the upper layer (SM_{upp} , 0 – 30 cm) and the
204 lower layer (SM_{low} , 30 – 150 cm). The snow component includes snow water equivalent
205 (SWE), as well as snow free water (SFW) representing the storage of melted snow. The water
206 stored in the stream channels and lakes is also included in the TWS estimate. Based on the



207 structure of PCR-GLOBWB, the total water storage (TWS) is computed as the sum of 27
208 different states: 2 groundwater, 8 soil moisture, 4 interceptions, 8 snow, 4 inundated top
209 water, and 1 surface water layer.

210 For each grid cell and for each daily time step, the model determines the water balance in two
211 vertically stacked soil layers and the groundwater store. The model also computes the vertical
212 water exchanges between the soil layers and between the inundated top water layer and the
213 atmosphere, i.e. rainfall and snowmelt, percolation and capillary rise, as well as evaporation
214 and transpiration fluxes. The active groundwater store underlies the soil, is fed by net
215 groundwater recharge, discharges to baseflow as a linear reservoir, and is exempt from the
216 direct influence of evaporation and transpiration fluxes. However, capillary rise from the
217 active groundwater store can occur depending on the simulated groundwater storage, the soil
218 moisture deficit, and the unsaturated hydraulic conductivity. Fluxes are simulated according
219 to the different land cover types. The model includes a physically-based scheme for
220 infiltration and runoff, resulting in the direct runoff, interflow, as well as groundwater
221 baseflow and recharge. River discharge is calculated by accumulating and routing the specific
222 runoff along the drainage network. For further details, including model parameterization, the
223 reader is referred to the technical reports and other relevant publications (van Beek and
224 Bierkens, 2009; van Beek, 2008; Sutanudjaja et al., 2011, 2014).

225

226 4. Data and data processing

227 4.1 GRACE data

228 The GRACE gravity product release 5 (RL05), generated by the University of Texas at
229 Austin's Center of Space Research (CSR, Bettadpur, 2012), was used as input. The product
230 consists of monthly sets of spherical harmonic coefficients (SHC) complete to the degree and
231 order 60. On this basis, TWS variations were obtained for the study period between April
232 2002 and December 2010. The GRACE data were further processed in this study as follows:

- 233 • SHCs of degree 1 provided by Swenson et al. (2008) were restored, and all 5
234 coefficients of degree 2 were replaced by the values estimated from satellite laser
235 ranging (Cheng and Tapley, 2004).
- 236 • SHC variations were computed by removing the long-term mean (computed between
237 April 2002 and December 2010) from each monthly solution.
- 238 • A destriping filter (Swenson and Wahr, 2006) was applied to the SHC variations. The
239 filter used a 5th degree polynomial (Savitsky-Golay) over a 5-point window to remove
240 the correlations; orders below 8 remained unchanged.
- 241 • An additional 250-km radius Gaussian smoothing (Jekeli, 1981) was applied to SHC
242 variations to suppress high-frequency noise, and the TWS variations ($\Delta\sigma$ [m]) were
243 then computed using (Wahr et al 1998)

$$\Delta\sigma(\theta, \phi) = \sum_{l=1}^{60} \sum_{m=-l}^l W_l \overbrace{\frac{a_e(2l+1)}{3(1+k_l)} \frac{\rho_e}{\rho_w}}^{S_l} \Delta\bar{C}_{lm} \hat{Y}_{lm}(\theta, \phi), \quad (1)$$



244

245 where θ , ϕ are co-latitude and longitude in spherical coordinates, $\Delta\bar{C}_{lm}$ is the SHC
246 variations of degree l and order m , \hat{Y}_{lm} is the normalized surface spherical harmonic,
247 W_l is the Gaussian smoothing function, S_l is a scaling factor used to convert
248 dimensionless coefficients to TWS in terms of Equivalent Water Heights (EWH), a_e
249 is the semi-major axis of the reference ellipsoid, k_l is the load love number of degree
250 l , ρ_e and ρ_w are the average density of the Earth and water, respectively. In this study,
251 the TWS variations were computed at every $0.5^\circ \times 0.5^\circ$ grid cell. This cell size was
252 selected through trial and error as a balance between performance and resolution.

253 In general, filters suppress not only noise but also the genuine TWS signal, and are a well-
254 known source of signal leakage. To address this, a signal restoration method (Chen et al.,
255 2014; Tangdamrongsub et al., 2016) was employed. The method iteratively determined the
256 possible signal reduction caused by the filter applied and added it back to the filtered signals.
257 The errors of the procedure grew with the number of iterations, requiring a proper selection of
258 the convergence criterion. In this study, the criterion was chosen empirically: the signal
259 restoration process was iteratively repeated until the increment in every grid cell inside the
260 Hexi Corridor became smaller than 0.5 cm. This value is 2-3 times smaller than the GRACE
261 uncertainty (Wahr et al., 2006; Klees et al., 2008; Dahle et al., 2014). Figure 4 demonstrates
262 the signal restoration for October 2002. The convergence criterion was met after
263 approximately 6 iterations. The signal over the mountain range and Inner Mongolia became
264 apparent after the signal restoration was applied (see Fig. 4f).

265 4.2 Forcing data

266 The forcing data required by PCR-GLOBWB are precipitation, air temperature, and potential
267 evapotranspiration. Tangdamrongsub et al. (2015) showed that the use of high-quality
268 precipitation data may lead to better estimates of hydrological fluxes (e.g., TWS and
269 streamflow). In principle, local precipitation and surface temperature measurements could be
270 obtained from the China Daily Ground Climate Dataset provided by the China Meteorological
271 Data Sharing Service System (<http://cdc.cma.gov.cn/home.do>). A total of 23 weather stations
272 were found over the Hexi Corridor (see Fig. 1b). However, the measurements were spatially
273 sparse and did not cover the entire region. Therefore, the global precipitation data were used
274 to achieve a better spatial coverage. Four global precipitation products were considered for
275 inclusion:

- 276 • The European Centre for Medium-range Weather Forecasts (ERA-Interim, spatial
277 resolution: $0.75^\circ \times 0.75^\circ$; Dee et al., 2011)
- 278 • The Tropical Rainfall Measuring Mission (TRMM 3B42, spatial resolution: 0.25°
279 $\times 0.25^\circ$; Huffman et al., 2007; Kummerow et al., 1998)
- 280 • The Climate Research Unit dataset (CRU, spatial resolution: $0.5^\circ \times 0.5^\circ$; Mitchell and
281 Jones, 2005; van Beek, 2008)
- 282 • The Princeton's Global Meteorological Forcing Dataset (Princeton, spatial resolution:
283 $0.5^\circ \times 0.5^\circ$; Sheffield et al., 2005)



284 To select the best product, the global precipitation values were firstly interpolated to the
 285 weather station locations and then the correlation coefficient, Nash-Sutcliffe (NS) coefficient,
 286 and RMS difference (RMSD) between the interpolated and observed ground data were
 287 calculated. The mean values of the statistical estimates are shown in Fig. 5a. Overall, TRMM
 288 provided the best data quality, with the highest correlation (~0.85) and NS coefficients
 289 (~0.46), and an RMSD approximately 2–3 mm lower than other products. The high spatial
 290 resolution of TRMM is probably the reason for its better performance. Therefore, this product
 291 was chosen as the precipitation input. The low NS coefficient in all 4 cases suggests that the
 292 coarse spatial resolution of the global precipitation datasets prevents them from capturing all
 293 the local precipitation events.

294 A similar procedure was used to compare the air temperature data from ERA-Interim, CRU,
 295 and Princeton. The statistical estimates are shown in Fig. 5b. Although the results from all
 296 products were very similar, CRU provided the highest data quality in terms of correlation and
 297 RMSD values, and therefore, it was used as the temperature input. As far as
 298 evapotranspiration is concerned, few data are available for this region, so the data from (van
 299 Beek, 2008) were used.

300 4.3 Validation data

301 4.3.1 Groundwater

302 Monthly groundwater well measurements at 5 locations (Fig. 1c) were obtained from the
 303 ground network maintained by the Shiyang River Basin Management Bureau, and Institute of
 304 Water Resources and Hydropower of Gansu Province. The in situ data were provided in the
 305 form of piezometric heads, which needed to be converted to units of storage. For such a task,
 306 several parameters, e.g., storage coefficient and specific yield are required, but they are not
 307 available over the Hexi Corridor. To solve that problem, a scale factor computed using the
 308 information from GRACE and the Global Land Data Assimilation System (GLDAS, Rodell et
 309 al., 2004) was used for the conversion using the approach outlined by Tangdamrongsub et al.
 310 (2015). In short, the procedure was as follows. First, GLDAS-based soil moisture storage
 311 variations (ΔSM) were removed from GRACE-derived TWS variations. Four variants of
 312 GLDAS model (NOAH, CLM, MOSAIC, and VIC; see Rodell et al., 2004) were considered
 313 and the average ΔSM value was calculated. Taking into account that ΔSM and groundwater
 314 storage variations (ΔGWS) are the major contributions to TWS variations, this resulted in
 315 $\Delta GWS_{(GRACE-\Delta SM)}$. Then, by conducting a regression analysis between the monthly
 316 time-series of piezometric head variation (Δh) and $\Delta GWS_{(GRACE-\Delta SM)}$, a scale factor (f) was
 317 estimated using the following relationship:

$$318 \Delta GWS_{(GRACE-\Delta SM)} + e = f \cdot \Delta h, \quad (2)$$

319 where e indicates the observation error. Finally, the in situ head measurements were
 320 multiplied with the estimated scale factor (\hat{f}) to obtain the converted groundwater storage
 321 variation ($\Delta GWS_{in situ}$) as:

$$322 \Delta GWS_{in situ} = \hat{f} \cdot \Delta h. \quad (3)$$



323

324 **4.3.2 Streamflow**

325 Monthly river gauge data were obtained from the same data centre as the groundwater
326 measurements. Due to the coarse spatial resolution of PCR-GLOBWB, it models only the
327 main river streams. Therefore, the gauge measurements of small river streams, as well as the
328 gauge measurements that contained many data gaps (e.g., more than 24 months), were
329 excluded. As a result, the measurements from only 2 gauges – at Xida and Xiying Rivers (see
330 Fig. 1c) – were used in this study.

331 **4.4.3 Normalized Difference Vegetation Index (NDVI)**

332 NDVI (Carlson and Ripley, 1997) is an indicator of vegetation health or “greenness”. In this
333 study, NDVI and GWS were analysed to determine if the growing season was being extended
334 beyond the limited rainy period through groundwater extraction for irrigation. NDVI was
335 computed from the MODIS 8-day, 500-m spatial resolution surface reflectance product
336 (Vermote et al., 2011) based on data from Aqua satellite (MYD09A1 product). Based on the
337 location of the in situ groundwater measurements, the MODIS tiles h25v05 and h26v05 were
338 selected. First, the data were quality controlled: pixels with cloud cover were flagged and
339 filled values were masked. The 8-day NDVI was then computed as (Huete et al., 2002)

$$340 \quad NDVI = \frac{\rho_{NIR} - \rho_R}{\rho_{NIR} + \rho_R}, \quad (4)$$

341 where ρ_{NIR} and ρ_R are the surface reflectance in the near-infrared and red portions of the
342 observed electromagnetic spectrum. The monthly-averaged NDVI was then computed based
343 on the derived 8-day NDVI values. Note that NDVI values range between -1 and 1 , and a
344 value over 0.2 generally indicates vegetation.

345

346 **5. Methodology and implementation**347 **5.1 Ensemble Kalman Filter (EnKF)**

348 The Ensemble Kalman Filter (EnKF; Evensen, 2003) is used to assimilate GRACE derived
349 TWS into the PCR-GLOBWB model. The EnKF works in two steps, a forecast step and
350 analysis (update) step. The forecast step involves propagating the states forward in time using
351 the model (PCR-GLOBWB). Identical to how the EnKF is implemented by Forman et al.
352 (2012), the state vector ($\boldsymbol{\psi}(t)$) in this study is an $nm \times 1$ vector, where $n = 27$ is the number of
353 TWS-related states from PCR-GLOBWB (see Sect. 3), and m is the number of model grid
354 cells. The model estimates are related to the GRACE observations by

$$355 \quad \mathbf{d}(t) = \mathbf{H}\boldsymbol{\psi}(t) + \boldsymbol{\epsilon}; \boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{R}), \quad (5)$$

356 where $\mathbf{d}(t)$ is an $m \times 1$ vector containing the GRACE observations for the month of interest,
357 and \mathbf{H} is a measurement operator which relates the PCR-GLOBWB state $\boldsymbol{\psi}(t)$ to the
358 observation vector $\mathbf{d}(t)$. Notice that the number of observations is equal to the number of grid



359 cells because the GRACE-based estimates are obtained for all the grid cells of the PCR-
 360 GLOBWB model (see Sect. 4.1). The uncertainties in the observations are given in the
 361 random error ϵ , which is assumed to have zero mean and covariance matrix $\mathbf{R}_{m \times m}$. As the
 362 sum of all state elements at a given cell is equal to TWS, the \mathbf{H} matrix is defined as:

$$363 \quad \mathbf{H} = \begin{bmatrix} (1 \ 1 \ 1 \ \dots \ 1)_{1 \times n} & 0 & \dots & 0 \\ 0 & (1 \ 1 \ 1 \ \dots \ 1)_{1 \times n} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & (1 \ 1 \ 1 \ \dots \ 1)_{1 \times n} \end{bmatrix}_{m \times nm} \quad (6)$$

364 If the ensembles of the states are stored in a matrix

365 $\mathbf{A}_{nm \times N} = (\boldsymbol{\psi}_1(t), \boldsymbol{\psi}_2(t), \boldsymbol{\psi}_3(t), \dots, \boldsymbol{\psi}_N(t))$, the ensemble perturbation matrix is defined as
 366 $\mathbf{A}' = \mathbf{A} - \bar{\mathbf{A}}$, where $\bar{\mathbf{A}}$ is the mean computed from all ensemble members. Similarly, the
 367 members of the GRACE observation vector are stored in the matrix
 368 $\mathbf{D}_{m \times N} = (\mathbf{d}_1(t), \mathbf{d}_2(t), \mathbf{d}_3(t), \dots, \mathbf{d}_N(t))$. Notice that $\mathbf{D}_{m \times N}$ contains N identical columns.
 369 The analysis equation can be expressed as (Evensen, 2003)

$$370 \quad \mathbf{A}^a = \mathbf{A} + \Delta\mathbf{A} = \mathbf{A} + \mathbf{K}(\mathbf{D} - \mathbf{H}\mathbf{A}) \quad (7)$$

371 with

$$372 \quad \mathbf{K} = \mathbf{P}_e \mathbf{H}^T (\mathbf{H} \mathbf{P}_e \mathbf{H}^T + \mathbf{R})^{-1}, \quad (8)$$

373 where $\mathbf{A}_{nm \times N}^a$ is the updated model state, $\Delta\mathbf{A}_{nm \times N}$ is the update from Kalman filter, and
 374 $\mathbf{K}_{nm \times m}$ is the Kalman gain matrix. The model error covariance matrix $(\mathbf{P}_e)_{nm \times nm}$ is
 375 computed as

$$376 \quad \mathbf{P}_e = \mathbf{A}' (\mathbf{A}')^T / (N - 1). \quad (9)$$

377 The computation of matrix \mathbf{R} , the error variance-covariance matrix of GRACE data in the
 378 spatial domain, is discussed in Sect. 5.2.2.

379 In the initialization phase, to obtain the initial states, the model was spun up between 1
 380 January 2000 and 31 December 2000 as a hot start. This time interval was sufficient to reach
 381 the dynamic equilibrium. The obtained initial state $\boldsymbol{\psi}(t)$ was perturbed and $N = 100$
 382 ensemble members $\boldsymbol{\psi}_i(t), i = 1, 2, 3, \dots, N$ were generated. The $N = 100$ ensemble runs
 383 between 1 January 2001 and 31 March 2002 were then conducted independently based on the
 384 perturbed initial states. This resulted in ensemble spread of the estimated states. The model
 385 was then propagated in time between 1 April 2002 and 31 December 2010 without
 386 assimilating any observation. This case is referred to hereafter as the Ensemble Open Loop
 387 (EnOL). For the EnKF, the model was also propagated beginning from 1 April 2002, but the
 388 observations (when available) were assimilated.

389 The processing diagram is shown in Fig. 6, and follows the methodology introduced by
 390 Forman et al. (2012). The state is first propagated in time from the first to the last day of the
 391 month without applying DA, and the monthly averaged states are calculated from the daily
 392 values. When the GRACE observation for that month is available, the DA routine is activated.



393 Otherwise, the model continues propagating to the next month without applying DA. The DA
394 routine updates the TWS-related states using Eq. (7). The daily increment (DINC) of the
395 update is then computed by dividing the monthly averaged update by the total numbers of
396 days in that month ($\text{numday}_{\text{month}}$). The model propagation is then restarted (second run), using
397 the last day of the previous month ($\text{month}-1$, $\text{numday}_{\text{month}-1}$) as the initial state. In this second
398 run, the DINC is added to the initial states every day up to the last day of the month. The DA
399 scheme is repeated to the end of the study period.

400 Spatial correlations of model and observation errors were also taken into account. De Lannoy
401 et al. (2009) proposed a so-called 3D-Fm (3-dimensional fine scale with multiple observation)
402 approach, which is called EnKF 3D in this paper. The approach only considered the spatial
403 correlations between the neighbouring grid cells. This reduced the computational cost, as only
404 a small subset of cells pairs was considered instead of all cells pairs. That approach was
405 applied not only to observation errors, but also to model errors in TWS and TWS-related
406 components in this study. The EnKF 3D scheme is illustrated in Fig. 7. For a particular grid
407 cell (centre grid cell), all TWS-related components of the neighbouring grid cells and the
408 centre grid cell are used to form the state ($\mathbf{A}_{np \times N}^s$) and observation ($\mathbf{D}_{p \times N}^s$) matrices, where p
409 is the number of the considered grid cells. The matrix notation with superscript s (e.g., \mathbf{A}^s) is
410 only used to emphasize the cell-dependent version, and it can be substituted into the original
411 matrix notation (e.g., \mathbf{A}) in Eqs. (5–9). It is emphasized here that EnKF 3D involves only p
412 grid cells instead of all m grid cells. As such, the measurement operator, model error
413 covariance matrix, and observation error covariance matrix becomes $\mathbf{H}_{p \times np}^s$, $(\mathbf{P}_e^s)_{np \times np}$, and
414 $\mathbf{R}_{p \times p}^s$, respectively. In this study, the neighbouring grid cells were assumed to be the ones
415 inside the Gaussian smoothing radius applied, i.e., 250 km. The EnKF was then applied and
416 the states of the centre grid cell (only) were updated. The procedure was repeated through all
417 grid cells. Note here that to avoid the edge effects, the grid cells at the edge were not updated
418 in the computation. To investigate the impact of including spatial correlations of errors, the
419 EnKF 1D was also considered. The EnKF 1D scheme is similar to EnKF 3D, but the spatial
420 correlations are omitted (i.e., the off-diagonal elements of the covariance matrices \mathbf{P}_e^s and \mathbf{R}^s
421 are set to zero).

422 Furthermore, sampling errors caused by finite ensemble size might lead to spurious
423 correlations in the estimated model error covariance matrices (Hamill et al., 2001). To reduce
424 such an effect, a distance-dependent localization function is applied to \mathbf{P}_e^s (pair-wise). In this
425 study, the Gaussian function ($c(\alpha)$) (Jekeli, 1981) was used:

$$426 \quad c(\alpha_{j_1, j_2}) = \frac{e^{-b[1 - \cos(\alpha_{j_1, j_2}/a_e)]}}{1 - e^{-2b}} \quad (10)$$

$$427 \quad \text{with } b = \frac{\ln(2)}{1 - \cos(L/a_e)}, \quad (11)$$

428 where α_{j_1, j_2} is the distance on the Earth surface between two grid cells (j_1 and j_2), and L is the
429 correlation distance. The variogram analysis was used to derive the TWS correlation distance
430 (L) of PCR-GLOBWB, assuming that it is similar to the correlation distance of model errors.
431 It was found to be approximately equal to 110 km over the Hexi Corridor. For GRACE



432 observations, to ensure that the spurious error correlations at distances greater than the
433 Gaussian smoothing distance, 250 km, was insignificant, the localization applied to \mathbf{R}^s was
434 based on $L = 250$ km.

435

436 **5.2 Errors of PCR-GLOBWB model and errors in GRACE observations**

437 **5.2.1 Model errors**

438 The two primary sources of considered errors in the PCR-GLOBWB model are the
439 meteorological forcing data and the model parameters. For forcing data, the precipitation
440 uncertainties were quantified as the RMS error provided by the TRMM product (Huffman,
441 1997). The uncertainties of temperature and potential evapotranspiration were not provided as
442 parts of the corresponding products, and therefore errors of 2°C, and 30% of the nominal
443 potential evapotranspiration value were assumed, respectively. The error levels were chosen
444 through trial-and-error, mainly to allow the ensemble to grow between updates. The
445 precipitation and potential evapotranspiration were perturbed with additive lognormal noise
446 while the temperature was perturbed with additive Gaussian noise. The forcing data
447 uncertainties were assumed to be spatially correlated, which was accounted for using an
448 exponential decay function. Based on a variogram analysis, the correlation distances of
449 precipitation, temperature and potential evapotranspiration were found to be approximately
450 150 km, 450 km, and 450 km, respectively.

451 As far as model parameters are concerned, a total of 15 TWS-related parameters (see Table 1)
452 were perturbed using additive Gaussian noise without spatial correlations. The standard
453 deviation of the perturbations of the parameters was set to 20% of the range of the nominal
454 values.

455 **5.2.2 GRACE observation errors**

456 Spatial correlations of GRACE observation errors were also taken into account in the DA
457 scheme. The uncertainties in the GRACE-derived TWS over the Hexi Corridor were
458 computed using the monthly calibrated error variance-covariance matrix of the SHCs ($\mathbf{\Sigma}$)
459 provided by the CSR. Recalling the replacement of the low degree SHCs (see Sect. 4.1), the
460 error (co-)variances of SHCs degree 2 were not provided by Cheng and Tapley (2004), and
461 therefore the values obtained from the CSR were used. As for SHCs of degree 1, the error (co-
462) variances were not available from (Swenson et al., 2008) either and were set to zero. Note
463 that $\mathbf{\Sigma}$ only reflects the error of the original GRACE data, i.e. before the GRACE processing
464 described in Sect. 4.1 was applied. To obtain the error variance-covariance matrix associated
465 with the post-processed GRACE data, an ensemble of SHC noise realizations \mathbf{Q}^c was first
466 generated based on $\mathbf{\Sigma}$ as follows:

$$467 \quad \mathbf{Q}^c = (\mathbf{\Sigma})^{\frac{1}{2}} \mathbf{Q}^w, \quad (12)$$

468 where $\mathbf{Q}^w = (q_1^w, q_2^w, q_3^w, \dots, q_N^w)$ contains a set of white noise realizations and has the
469 dimension of $s \times N$, where $s = 1891$ is the number of SHCs, and $N = 100$ is the number of



470 realizations. The matrix $\mathbf{Q}^c = (q_1^c, q_2^c, q_3^c, \dots, q_N^c)$ has the same dimension as \mathbf{Q}^w and contains
 471 an ensemble of correlated noise realizations in SHCs. Then, each noise realization (i.e.,
 472 column of \mathbf{Q}^c) was post-processed in the same way as the GRACE data (Sect. 4.1), which
 473 resulted in $\hat{\mathbf{Q}}^c = (\hat{q}_1^c, \hat{q}_2^c, \hat{q}_3^c, \dots, \hat{q}_N^c)$. The post-processing included applying the destriping and
 474 Gaussian smoothing filters, as well as the signal restoration using the same number of
 475 iterations as was used in the GRACE data post-processing. The error variance-covariance
 476 matrix $\hat{\Sigma}$ associated with the SHCs after post-processing was then computed as

$$477 \quad \hat{\Sigma} = [\hat{\mathbf{Q}}^c (\hat{\mathbf{Q}}^c)^T] / (N - 1). \quad (13)$$

478 Recalling Eq. (1), the TWS variations over the Hexi Corridor can be computed as

$$479 \quad \Delta\sigma = \mathbf{Y}\mathbf{S}\mathbf{x}, \quad (14)$$

480 where $\Delta\sigma$ is the vector composed of the computed TWS variations at grid cells, \mathbf{Y} is the
 481 matrix of spherical harmonic synthesis (cf. Eq. (1)), \mathbf{S} is the matrix containing the scaling
 482 factors S_l , and \mathbf{x} is the vector composed of the dimensionless SHC variations after GRACE
 483 data post-processing described in Sect. 4.1. Then, the error covariance matrix \mathbf{R} of the
 484 GRACE-based TWS variations over the Hexi Corridor was computed with the error
 485 propagation law as

$$486 \quad \mathbf{R} = \mathbf{Y}\mathbf{S}\hat{\Sigma}(\mathbf{Y}\mathbf{S})^T. \quad (15)$$

487 Some statistics of GRACE TWS errors over the Hexi Corridor are shown in Fig. 8. The error
 488 standard deviation in Oct. 2002 varied with location (Fig. 8a), whereas the error correlation
 489 showed a distance-decay pattern in all directions (Fig. 8b). The areally-averaged standard
 490 deviations over 4 basins stayed in most of the months at a similar level of approximately 1 cm
 491 (Fig. 8c). The large uncertainty in September 2004 was likely caused by the near-repeat orbit
 492 of GRACE satellites during that month.

493

494 **6. Results and discussion**

495 The structure of this section is as follows. First, the impact of assimilation using EnKF 3D on
 496 the total TWS is considered in Sect. 6.1. Then, the impact of the EnKF 3D on the estimates of
 497 the individual stores is investigated in Sect. 6.2. The performances of the EnKF 1D and EnKF
 498 3D schemes are compared in Sect. 6.3 in terms of total TWS and the individual stores.
 499 Finally, in Sect. 6.4 the assimilation results are used together with ancillary remote sensing
 500 data to study water resources in the Hexi Corridor.

501

502 **6.1 Performance of GRACE DA scheme**

503 To demonstrate the impact of DA, Fig. 9 shows the daily TWS estimates over the Shiyang
 504 River Basin between 1 April 2002 and 31 December 2003. Several features associated with



505 the EnKF can be observed. Firstly, when a GRACE observation is available, the EnKF moves
506 the estimated TWS towards it. As a result, the estimated TWS lies between the EnOL estimate
507 and the GRACE observation most of the time. Secondly, the standard deviation across the
508 EnKF ensemble of TWS values is smaller than that of the EnOL and smaller than the GRACE
509 observation error. Thirdly, at the first month (April 2002) the TWS estimates of the EnOL and
510 EnKF were similar at the forecast step (as the initial states were the same), but became
511 different when the daily increment was applied to the EnKF (see point (a) in Fig. 9). Finally,
512 discontinuities in the time-series before the update were observed at the end of the month e.g.,
513 in November and December 2002 (point (b) and (c)), and February 2003 (point (d)).
514 Applying the daily increment (see Sect. 5.3) served as a smoother, and these stepwise changes
515 were reduced. Similar features were also seen in the EnKF 1D TWS estimates (not shown).

516

517 **6.2 Impact of GRACE DA on individual stores**

518 The monthly-averaged values of the TWS and individual stores in each of the 4 basins are
519 presented in Fig. 10. Overall, TWS estimates over the Hexi Corridor mostly reflect variations
520 of SM and GWS components, while snow and surface water are minor contributors,
521 constituting less than 5% in most basins. Clear seasonal variations in TWS were seen in all
522 basins for GRACE, EnOL and GRACE DA (both EnKF 1D and EnKF 3D) (Fig. 10 a,b,c,d).
523 As observed in Fig. 9, the GRACE DA estimated TWSs are generally between the GRACE
524 observations and the EnOL estimates. As a result of assimilating GRACE data, both the EnKF
525 1D and EnKF 3D added water to all basins between 2002 and 2005 and reduced it from the
526 basins between 2006 and 2010. This is also reflected in the time-series of SM (Fig. 12 e,f,g,h)
527 and GWS (Fig. 12 i,j,k,l). Additionally, the annual amplitudes and phases of GRACE DA
528 estimated TWS were also found mostly in between the values computed from the GRACE
529 observations and the EnOL results (see Table 2). In particular, the GRACE-DA estimated
530 TWS's phase was always closer to the GRACE observation. The phase shifts of
531 approximately 1 month were seen in both GRACE DA estimated TWS and GRACE
532 observations compared to the EnOL results. Similar phase differences of approximately 1
533 month were also observed in SM and GWS components.

534 Differences in the long-term trends were also detected between the TWS estimates from the
535 model alone (EnOL) and the GRACE DA. The GRACE DA results showed similar
536 decreasing TWS trends to the GRACE data, while the EnOL showed increasing trends (Fig.
537 10 a,b,c,d, see also Table 7). This change in TWS trend was clearly a result of assimilating
538 GRACE observations. The negative trends were also observed after DA in the GWS
539 component in most basins (Fig. 10 i,j,l). This indicates the potential of GRACE DA in
540 adjusting GWS. In this way, one can reveal continued groundwater consumption to support
541 local agricultural activities (Li et al., 2013). Unlike over other basins, the negative trend of
542 GWS estimates was not clearly present over the Desert Region (Fig.10k). This could be due to
543 the small size of the groundwater store of this region (see also below), and most of the update
544 took place in the SM component. As a result, a relatively large negative trend was seen in SM



545 rather than GWS after GRACE DA (see also Table 7). Further discussions of the trends are
546 given in Sect. 6.4.

547 The impact of GRACE DA on different stores was influenced by both the model parameters
548 and the forcing data assigned. For example, comparing to other basins, the Shiyang River
549 Basin has smaller values of saturated hydraulic conductivities ($K_{\text{sat,up}}$, $K_{\text{sat,low}}$) and a higher
550 value of the recession coefficient (J) (see Table 3), which allows for larger updates to the
551 groundwater compartment. This explains the significant amount of update seen in GWS (Fig.
552 10i). Similar behaviour was also reported in (Tangdamrongsub et al., 2015). Conversely, the
553 groundwater compartment of the Desert Region received a tiny amount of water due to the
554 high value of K_{sat} and low value of J.

555 Forcing data also had an impact on which stores were updated. Consider, for instance, the SM
556 compartment. Although the Shiyang River Basin has the smallest K_{sat} compared to other
557 basins (Table 3), the annual amplitude of SM (~ 1 cm) there was not the smallest, but the
558 largest one (see Fig. 10e), as the basin received the greatest amount of rainfall (see Fig. 2a). In
559 contrast to the Shiyang River Basin, the Shule River Basin has similar K_{sat} but the smallest
560 SM annual amplitude (~ 0.3 cm) was observed (Fig. 10h). This reflects the limited amount of
561 rainfall received by the basin (see Fig. 2g).

562 Snow estimates (SWE plus SFW) were very small (less than 0.2 cm) over the Hexi Corridor
563 and therefore were only slightly updated by GRACE DA. The large amount of snow seen as
564 the sharp peaks, e.g., in January 2008 was caused by the precipitation and temperature
565 variability. In January 2008, the precipitation records were 159 % higher than the January
566 average value while the temperature was $2 - 3^{\circ}\text{C}$ lower. Such a condition resulted in a large
567 amount of snow. Finally, GRACE DA influences the surface water, but the amplitude is still
568 lower than that of the GRACE uncertainties. Validation of the surface water estimates in
569 terms of river streamflow is given in Sect. 6.3.2.

570

571 **6.3 Impact of taking spatial correlations of errors into account**

572 The impact of taking spatial correlation of errors into account was evaluated by comparing
573 estimates from EnKF-1D and EnKF-3D with the in situ measurements and independent
574 satellite observations. First of all, the impact of accounting for the error correlations was
575 clearly seen in the TWS estimates (Fig. 10 a,b,c,d). When the error correlations were ignored
576 (EnKF 1D), the TWS estimate received a larger update from GRACE, particularly between
577 2002 and 2005. Hence, the estimate was drawn significantly closer to the observation.
578 Presence of error correlations effectively reduces amount of information in GRACE data,
579 since spatial averaging of such data mitigates noise to a much less extent than averaging of
580 data with uncorrelated errors. Therefore, the impact of GRACE data in the EnKF 3D case is
581 reduced. As such, the EnKF 3D estimated TWS was always between the EnOL and EnKF 1D
582 results. Validating against the in situ groundwater and streamflow data will reveal which
583 results are closer to the truth (Sect. 6.3.1, 6.3.2).



584 Taking error correlations into account also has a clear impact on the SM and GWS
585 components. For SM, similarly to TWS, the EnKF 1D yielded a larger update between 2002
586 and 2005 compared to the EnKF 3D (Fig. 10 e,f,g,h). The difference between EnKF 1D and
587 3D results became smaller after 2005. This can be attributed to the fact that the ensemble
588 spread in the SM component becomes smaller after several years of updates. After 2005, the
589 ensemble spread of SM was lower than the GRACE uncertainty, and therefore taking the error
590 correlations into account did not have a significant impact on the SM estimates. For GWS, the
591 impact of taking error correlations into account was even clearer, especially in terms of the
592 long-term trend (Fig. 10, i,j,k,l). With the exception of the Desert Region, the EnKF 1D
593 showed a steeper decreasing trend in all basins. For snow and surface water, the impact of
594 considering error correlations was not significant due to the fact that the stores are small, as
595 compared to SM and GWS.

596 **6.3.1 Validation against groundwater estimates**

597 The GWSs estimated from GRACE DA were validated against the well measurements at 5
598 locations shown in Fig. 1c. The GWS estimate at each well location is shown in Fig. 11.
599 Compared to the EnOL results, GRACE DA results were visually closer to the well
600 measurements at all 5 locations. The EnKF 1D and EnKF 3D showed a noticeable difference
601 at each location. The updated GWS estimates were evaluated in terms of the correlation
602 coefficient, RMS difference (RMSD), and long-term trend (Table 4, 5). Overall, the EnOL
603 resulted in relatively poor correlation coefficients at most stations (except station W1), with
604 the average value of only 0.06. Clear improvements were seen after GRACE DA was applied.
605 The average correlation coefficient increased to approximately 0.6 – 0.7. Although the EnKF
606 1D introduced a greater update than the EnKF 3D, it only showed higher correlation
607 coefficients at stations W1 and W3. Applying the EnKF 3D led to correlation coefficients
608 greater than 0.45 in all stations, and on average it improved the correlation coefficient by
609 approximately 0.1 over EnKF 1D. In terms of RMSD, applying GRACE DA reduced the
610 difference by approximately 15 – 25% compared to the EnOL. Compared to EnKF 1D, the
611 EnKF 3D significantly improved the RMSD in most stations. The EnKF 1D only performed
612 better than EnKF 3D at station W1, where it reduced the RMSD by approximately 16 %
613 compared to 8% reduction by the EnKF 3D. The noticeably low GWS observed by the well
614 data at station W2 in the summers of 2007 and 2008 (Fig. 11b) was probably caused by
615 significant groundwater abstraction. These local features could not be reproduced due to a
616 limited spatial resolution of the model and GRACE observations. As a result, neither of the
617 EnKF algorithms could improve the GWS estimates at the W2 location during those periods.

618 The long-term trend estimated between 2007 and 2010 was also used to evaluate the impact of
619 taking the error correlations into account (Table 5). The EnOL trend estimates were
620 considered poor as they showed the largest RMS difference respected to the in situ data. In
621 fact, they were the least consistent with the in situ estimates at each individual station. Similar
622 to the results in terms of correlation coefficient and RMSD (see Table 4), the EnKF 3D led to
623 the largest improvement in the trend estimates (RMSD=0.54 compared to 0.93 after EnKF
624 1D). However, while the EnKF 3D showed closer long-term trends to the in situ



625 measurements at stations W2, W4, W5, the EnKF 1D produced better estimates at station W1
626 and W3.

627 Thus, both EnKF 1D and 3D led to the improvement of the GWS estimates in terms of all
628 metrics. In terms of the average results and at the majority of well locations, the EnKF 3D
629 provided more improvement than the EnKF 1D.

630 **6.3.2 Validation against streamflow estimates**

631 The streamflow estimates were validated against the river gauge measurements at locations
632 G1 and G2 (Fig 1c). Results are shown in Figure 12 and Table 6. Only modest improvements
633 in the streamflow estimates are observed in terms of the correlation coefficient, NS
634 coefficient, and RMS difference (RMSD). This behaviour is similar to what was observed
635 previously for Rhine River Basin when a different hydrology model and input data were used
636 (Tangdamrongsub et al., 2015). Figure 12 shows that taking error correlations into account
637 had little impact, i.e. similar streamflow estimates were seen for EnKF 1D and 3D results. At
638 location G1 (Fig. 12a), GRACE DA added more water to the stream channel between 2002
639 and 2006 and reduced it between 2008 and 2010. This behaviour is consistent with the TWS
640 estimates discussed in Sect. 6.2. GRACE DA increased the correlation coefficient from 0.82
641 to 0.84, increased the NS coefficient from 0.65 to 0.69, and reduced the RMSD by
642 approximately 5 % (Table 6). Similar improvements were also observed at G2.

643 Comparing to the gauge measurements, both the EnOL and GRACE DA overestimated the
644 streamflow in September 2007 and September 2008 at G2. The sudden surge in streamflow
645 results from heavy rainfall while the soil is saturated (Fig. 13). For example, in September
646 2007, the second highest amount of SM storage in the record (~19.5 cm) was observed when
647 the third largest amount of rainfall (~3 mm) occurred. Similarly, in September 2008, large SM
648 storage (~20 cm) and the heaviest rainfall (~3.4 mm) forced PCR-GLOBWB to generate a
649 large amount of streamflow, which significantly exceeded the actual one observed at G2.
650 Inaccurate precipitation data and model calibration led to the discrepancy here. GRACE DA
651 was unable to reduce these spurious peaks due to the limited spatial (~250 km) and temporal
652 (1 month) resolution.

653

654 **6.4 Declining water storages in the Hexi Corridor**

655 The water resources situation over the Hexi Corridor was assessed using long-term trends
656 estimated from 9-year EnKF 3D results. This DA variant is primarily discussed here as it
657 provided better agreement with in-situ observations than the EnKF 1D (see Sect. 6.3.1). For
658 completeness, however, the values estimated from GRACE, EnOL, EnKF 1D, and
659 precipitation are also provided. The trends in the TWS, SM and GWS variations for the 4
660 basins, as well as the areally-averaged values for the entire Hexi Corridor, are given in Table
661 7. The average EnKF 3D trends are all negative: approximately -0.2, -0.1, and -0.1 cm/yr for
662 TWS, SM, and GWS, respectively. This reduction in the water storages is observed despite
663 the increased amount of rainfall, which shows a positive trend of about 0.4 (mm/month)/yr.



664 The water storage reductions can likely be attributed to the extraction of groundwater to meet
665 irrigation demands. In Sect. 6.5, it will be shown that groundwater extractions are essential for
666 that purpose in the Hexi Corridor.

667 Focusing on individual river basins provides additional insight into the water storage issue, as
668 the influence of the large desert area is removed. The water storage losses in the individual
669 basins are even more pronounced, particularly in the Shiyang River Basin. This basin had the
670 greatest TWS loss (approximately 0.4 cm/yr), which was entirely caused by the reduction of
671 GWS. This can be explained by groundwater abstraction to meet the irrigation demand in the
672 region. The Heihe and Shule River Basins also experienced a TWS loss of ~0.2 cm/yr, which
673 came from a reduction of both SM and GWS storages. Again, the negative GWS trend was
674 likely caused by significant pumping of groundwater to maintain crop production. This is
675 consistent with the extreme water stress over the Heihe River basin between 2001 and 2010,
676 which was documented in Table 11.7 of the study by Chen et al. (2014). Finally, the
677 decreasing TWS trend of 0.1 cm/yr detected in the Desert Region was dominated by loss of
678 SM storage.

679

680 **6.5 Connection to agriculture activity**

681 Figure 14 shows the monthly averaged groundwater head measurements at wells W1 to W5 in
682 the Shiyang River Basin (Fig. 1c). Monthly averaged precipitation and NDVI values are
683 shown as well. Since extracted water can be used to support agriculture not only at the well
684 location but also in the nearby area, precipitation and NDVI are reported as the average values
685 within a circular area of the 10-km radius. These data will be used to ascertain if groundwater
686 extractions to support agriculture might be the source of the negative GWS trends observed in
687 Fig. 11 and Table 6. From Fig. 14, it is noticed that the growing period is approximately
688 between May and October, where the amount of rainfall is higher than 15 mm/month and the
689 NDVI is typically greater than 0.2. By observing well measurements, precipitation, and NDVI
690 together, some groundwater extraction signatures can be explained by the extension of the
691 growing period over the dry season. For example, at station W1, the groundwater in 2010 was
692 lower than the average, showing a gradual decrease in summer (Fig. 14a). One may attribute
693 this to the shortage of rainfall in July and August 2010, which was lower than the average
694 (Fig. 14b). However, the NDVI value was higher than the average during summer 2010 (Fig.
695 14c), which implies that water from other sources than precipitation was probably used to
696 maintain the growing period. This additional water was likely extracted from the ground, and
697 such an activity led to a decreased groundwater table during summer 2010. A similar
698 explanation can be applied to station W2, where low groundwater head, low rainfall, and high
699 NDVI were observed in summer 2007 and summer 2008 (Fig. 14 d,e,f). At station W3, the
700 behaviour is similar to station W1: the extension of the growing period was observed in
701 summer 2010, where the GWS and precipitation were lower than the average, while NDVI
702 was significantly higher (Fig. 14 g,h,i). Groundwater pumping signatures were not present at
703 stations W4 and W5.

704



705 **7. Conclusions**

706 This study improved the estimation of water resources dynamics in the Hexi Corridor by
707 assimilating GRACE-derived TWS variations into the PCR-GLOBWB hydrological model. It
708 was found that including the spatially-correlated errors into the DA scheme led to the
709 improvement of the state estimates. Furthermore, GRACE DA estimates revealed the
710 reduction of water storages between 2002 and 2010. The Shiyang River Basin suffered the
711 most from the water loss, which was likely caused by the overuse of the groundwater for
712 irrigation. Due to inaccurate groundwater abstraction information, PCR-GLOBWB alone
713 could not properly capture the downward trend of water storages. This highlights the value of
714 the GRACE DA in this situation.

715 Furthermore, we demonstrate how the error covariance matrix of GRACE TWS can be
716 derived from the error covariance matrix of GRACE SHCs (which is currently provided
717 together with the SHCs themselves). This study shows that it is necessary to consider the error
718 correlations in the DA scheme. To illustrate, the assimilation schemes considered 2 variants of
719 the error variance-covariance matrix, excluding and including error correlations. Validating
720 against ground data found that both DA schemes lead to noticeable improvement in the state
721 estimates in terms of correlation, RMS difference, and long-term trend. However, ignoring
722 error correlations in DA tended to over-fit results to the observations, and in many cases led to
723 less accurate state estimates. On average, greater improvement was seen when the error
724 correlations were taken into account. At the same time, we admit that the derivation of
725 GRACE TWS error variance-covariance matrices is very computationally demanding. Still,
726 we believe that this is a reasonable price to pay as full (and only full) error covariance
727 matrices reflect the real uncertainty of GRACE observations. Of course, the performance of
728 EnKF 3D needs to be further investigated in other geographical locations and with different
729 hydrological models to confirm the benefits of this scheme.

730 GRACE DA strengthened the PCR-GLOBWB model in capturing the signature of
731 groundwater abstraction. It should be emphasized that GRACE does not fix a technical
732 problem of the hydrological model, but it rather provides information which is not available
733 otherwise. Note that, in principle, the model may predict any long-term behaviour of water
734 storage, but that information should be brought in "by hand" (e.g., via the groundwater
735 abstraction parameter). As soon as that information is not available, reliable long-term
736 predictions on the basis of hydrological modelling alone are conceptually impossible.
737 GRACE DA acts as a provider of a missing puzzle piece here.

738 Substantial decreasing of water storage was observed in the Hexi Corridor. The region
739 received increased precipitation between 2002 and 2010, but the water storage was found to
740 be declining, particularly over the Shiyang River Basin. The amount of water from rainfall
741 was obviously insufficient to support irrigation water requirements. Irrigation water demands
742 increased significantly to maintain the crop production and, as a result, the region was under
743 extreme water stress. Water consumption from all available sources was essential for bridging
744 the deficit, including an enormous amount of groundwater extraction. This study illustrates



745 the connection between groundwater pumping and agriculture activity, which is clearly seen
746 from the ground observation and remote sensing data.

747 To further improve the DA performance, an extended or an alternative DA framework can be
748 considered. One of the points of attention is only a minor improvement in streamflow
749 estimates, which is caused by an insufficient temporal and spatial resolution of GRACE data.
750 A promising way to go is to improve the runoff scheme at a conceptual level, e.g., by
751 extending GRACE DA with a simultaneous parameter calibration. To that end, the state
752 vector should be extended to include also selected model parameters (Eicker et al., 2014;
753 Wanders et al., 2014). This allows for the adjustment of the storage size and might lead to a
754 more accurate estimate of model states, including streamflow (Wanders et al., 2014).
755 Alternative ensemble-based DA approaches, such as particle filters (Weerts and El Serafy,
756 2006), can also be considered. Particle filters estimate a sample from the true posteriori
757 distribution, which is not necessarily Gaussian, like in the EnKF. The approach has been
758 shown very effective for the parameter calibration (Dong et al., 2015). Finally, the uses of
759 improved gravity solutions from the GRACE Follow-on (Flechtner et al., 2014) will probably
760 further increase the accuracy of the GRACE DA estimates.

761

762 **Acknowledgement**

763 This research was funded by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek
764 (Netherlands Organisation for Scientific Research, NWO; project number 842.00.006) and
765 Ministry of Science and Technology of China (MoST, Project Number 2010DFA21750)
766 under the Samenwerking China - Joint Scientific Thematic Research Programme (JSTP). The
767 research was also sponsored by the NWO Exacte Wetenschappen, EW (NWO Physical
768 Sciences Division) for the use of supercomputer facilities, with financial support from NWO.
769 The research was also co-funded by National Natural Science Foundation of China (NSFC,
770 project number 51279076).

771

772 **Appendix A: Water demand calculation in PCR-GLOBWB**

773 PCR-GLOBWB includes an online and integrated scheme to simulate irrigation water
774 requirement. The scheme separately parameterizes two different irrigated crop groups: paddy
775 and non-paddy, aggregated from 26 crop classes from the MIRCA2000 dataset, which
776 accounts for various growing season lengths. The calculated irrigation water requirement is
777 applied according to crop specific calendars, which ensure optimal crop growth. Principally,
778 this irrigation water demand scheme aims to maintain certain soil moisture levels in order to
779 provide optimal crop transpiration but still takes into account soil water availability,
780 interception, bare soil evaporation, as well as open water evaporation over inundated paddy
781 fields. Over daily time steps, irrigation water demand is calculated by considering the deficit
782 of readily available water in the soil moisture layers (thickness ≤ 1.2 m) to their total storage
783 capacities. To represent flooding irrigation over paddy fields, a certain surface water depth



784 (e.g., 50–100 mm) is maintained until the late crop development stage before the harvest. The
785 dynamic irrigation scheme in PCR-GLOBWB also considers the historical growth of irrigated
786 areas based on FAOSTAT.

787 Other sectoral water demands, including those from livestock, industry, and household, are
788 compiled from several sources by considering many factors, including past change in
789 population, socio-economic and technological development.

- 790 • Livestock water demand is calculated by multiplying the number of livestock in a grid
791 cell with its corresponding drinking water requirement, which is a function of air
792 temperature. The gridded global livestock densities of cattle, buffalo, sheep, goats,
793 pigs and poultry and their corresponding drinking water requirements are obtained
794 from FAO (Steinfeld et al., 2006), and FAOSTAT (<http://faostat.fao.org/>).
- 795 • Historical and gridded industrial demand data are obtained from several sources
796 (Shiklomanov, 1997; WRI, 1998; Vörösmarty et al., 2005). The algorithm to calculate
797 this demand includes country-specific economic development based on four
798 socioeconomic variables: gross domestic product (GDP), electricity production,
799 energy consumption, and household consumption. Associated technological
800 development per country is then approximated by energy consumption per unit
801 electricity production, which accounts for industrial restructuring or improved water
802 use efficiency.
- 803 • Household or domestic water demand is estimated by multiplying the number of
804 population in a cell with the country-specific per capita domestic withdrawals. The
805 country domestic withdrawals are taken from the FAO AQUASTAT
806 (<http://www.fao.org/nr/water/aquastat/main/index.stm>) and (Gleick et al., 2009).
807 Economic and technological developments are taken into account. Seasonality of
808 household/domestic water demand is also considered, using air temperature as a
809 proxy. Here, available gridded global population maps per decade (Klein Goldewijk
810 and van Drecht, 2006) are used to downscale the country-scale map to produce the
811 gridded water demand data.

812 The allocation of various water sources, i.e. surface water, groundwater and desalinated water,
813 is principally based on the simulation of their availabilities. However, the model also
814 incorporates some inventory data and studies about local preferences of certain regions. This
815 means that areas with extensive surface water irrigation supplies and networks prioritize
816 surface water use. Moreover, cities with poor drinking water facilities and water distribution
817 networks use groundwater as their main source (although they may be close to the rivers with
818 abundant surface water).

819

820 8. References

821 Bettadpur, S.: Insights into the Earth System mass variability from CSR-RL05 GRACE
822 gravity fields, EGU Meeting, abstract #EGU2012-6409, Vienna, Austria, 2012.



- 823 Carlson, T. N. and Ripley, D. A.: On the relation between NDVI, fractional vegetation cover,
824 and leaf area index, *Remote Sens. Environ.*, 62(3), 241–252, doi:10.1016/S0034-
825 4257(97)00104-1, 1997.
- 826 Chen, J., Li, J., Zhang, Z., and Ni, S.: Long-term groundwater variations in Northwest India
827 from satellite gravity measurements. *Global Planet. Change*, 116, 130–138.
828 doi:10.1016/j.gloplacha.2014.02.007, 2014.
- 829 Chen, Y.: *Water resources research in northwest China*, Springer Netherlands,
830 doi:10.1007/978-94-017-8017-9, 2014.
- 831 Cheng, M. and Tapley, B.: Variations in the Earth's oblateness during the past 28 years, *J.*
832 *Geophys. Res.*, 109(B09402), doi:10.1029/2004JB003028, 2004.
- 833 Dahle, C., Flechtner, F., Gruber, C., König, D., König, R., Michalak, G., and Neumayer, K.-
834 H.: GFZ RL05: An Improved Time-Series of Monthly GRACE Gravity Field Solutions, In
835 Flechtner, F., Sneeuw, N., Schuh, W.-D. (Eds.), *Observation of the System Earth from Space*
836 - CHAMP, GRACE, GOCE and future missions, (GEOTECHNOLOGIEN Science Report;
837 20; Advanced Technologies in Earth Sciences), Berlin, Springer, 29–39,
838 http://doi.org/10.1007/978-3-642-32135-1_4, 2014.
- 839 Dahlgren, S. and Possling, B.: *Soil Water Modelling In Arid/Semiarid Regions of Northern*
840 *China Using Land Information System (LIS): A Minor Field Study in Shiyang River Basin*,
841 Dept. of Water Resources Engineering, Lund University, ISBN LUTVDG/TVVR-07/5019,
842 ISSN-1101-9824, 2007.
- 843 De Graaf, I. E. M., van Beek, L. P. H., Wada, Y., and Bierkens, M. F. P.: Dynamic attribution
844 of global water demand to surface water and groundwater resources: Effects of abstractions
845 and return flows on river discharges, *Adv. Water Resour.*, 64, 21–33,
846 doi:10.1016/j.advwatres.2013.12.002, 2014.
- 847 De Lannoy, G. J. M., Reichle, R. H., Houser, P. R., Arsenault, K. R., Verhoest, N. E. C., and
848 Pauwels, R. N.: Satellite-Scale Snow Water Equivalent Assimilation into a High-Resolution
849 Land Surface Model, *J. Hydrometeor.*, 11, 352–369,
850 doi:<http://dx.doi.org/10.1175/2009JHM1192.1>, 2009.
- 851 Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U.,
852 Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L.,
853 Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haiberg, L.,
854 Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M.,
855 McNally, A. P., Monge-Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey, C., de Rosnay, P.,
856 Tavolato, C., Thépaut, J. N., and Vitart, F. (2011). The ERA-Interim reanalysis: configuration
857 and performance of the data assimilation system. *Quarterly Journal of the Royal*
858 *Meteorological Society*, 137, 553–597, doi:10.1002/qj.828.



- 859 Dong, J., Steele-Dunne, S. C., Judge, J., and van de Giesen, N.: A particle batch smoother for
860 soil moisture estimation using soil temperature observations, *Adv. Water Resour.*, 83, 111 –
861 122, <http://dx.doi.org/10.1016/j.advwatres.2015.05.017>, 2015.
- 862 Döll, P., Schmied, H. M., Schuh, C., Portmann, F. T., and Eicker, A.: Global-scale assessment
863 of groundwater depletion and related groundwater abstractions: Combining hydrological
864 modeling with information from well observations and GRACE satellites, *Water Resour.*
865 *Res.*, 50, 5698–5720, doi:10.1002/2014WR015595, 2014.
- 866 Du, T., Kang, S., Zhang, Z., and Zhang, J.: China's food security is threatened by the
867 unsustainable use of water resources in North and Northwest China, *Food and Energy*
868 *Security*, 3(1), 7–18, doi: 10.1002/fes3.40, 2014.
- 869 Eicker, A., Schumacher, M., Kusche, J., Döll, P., and Müller Schmied, H.: Calibration data
870 assimilation approach for integrating GRACE data into the WaterGAP Global Hydrology
871 Model (WGHM) using an Ensemble Kalman Filter: First Results, *Surv. Geophys.*, 35(6),
872 1285-1309, doi:10.1007/s10712-014-9309-8, 2014.
- 873 Evensen, G.: The ensemble Kalman filter: Theoretical formulation and practical
874 implementation, *Ocean Dyn.*, 53(4), 343–367, doi:10.1007/S10236-003-0036-9, 2003.
- 875 Flechtner, F., Morton, P., Watkins, M., and Webb, F.: Status of the GRACE follow-on
876 mission, in IAG symposium gravity, geoid, and height systems, 141, Venice, Italy, Springer,
877 117–121, 2014.
- 878 Forman, B. A. and Reichle, R.: The spatial scale of model errors and assimilated retrievals in
879 a terrestrial water storage assimilation system, *Water Resour. Res.*, 49, 7457–7468,
880 doi:10.1002/2012WR012885, 2013.
- 881 Forman, B. A., Reichle, R. H., and Rodell, M.: Assimilation of terrestrial water storage from
882 GRACE in a snow-dominated basin, *Water Resour. Res.*, 48, W01507,
883 doi:10.1029/2011WR011239, 2012.
- 884 Geng, G. T. and Wardlaw, R.: Application of Multi-Criterion Decision Making Analysis to
885 Integrated Water Resources Management, *Water Resour. Manage.*, 27, 3191–3207, 2013.
- 886 Gleick, P. H., Cooley, H., Cohen, M. J., Morikawa, M., Morrison, J., and Palaniappan, M.:
887 *The Worlds Water 2008–2009, The Biennial Report on Freshwater Resources*, Island Press,
888 Washington, DC, USA, 2009.
- 889 Hamill, T. M., Whitaker, J. S., and Snyder, C.: Distance-Dependent Filtering of Background
890 Error Covariance Estimates in an Ensemble Kalman Filter, *Mon. Weather Rev.*, 129, 2776–
891 2790, 2001.
- 892 Houborg, R., Rodell, M., Li, B., Reichle, R., and Zaitchik, B. F.: Drought indicators based on
893 model-assimilated Gravity Recovery and Climate Experiment (GRACE) terrestrial water
894 storage observations, *Water Resour. Res.*, 48, W07525, doi:10.1029/2011WR011291, 2012.



- 895 Hu, Z. D.: Study on Evolution Pattern of Water Resources, Ecology and Oasis
896 Migration in the Hexi Corridor, Doctoral Dissertation, Tsinghua University, Beijing, China,
897 2015.
- 898 Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., and Ferreira, L. G.: Overview of
899 the radiometric and biophysical performance of the MODIS vegetation indices, *Remote*, 83,
900 195–213, doi:10.1016/S0034-4257(02)00096-2, 2002.
- 901 Huffman, G.J.: Estimates of Root-Mean-Square Random Error for Finite Samples of
902 Estimated Precipitation, *J. Appl. Meteor.*, 1191–1201, 1997.
- 903 Huffman, G. J., Adler, R. F., Bolvin, D. T., Gu, G., Nelkin, E. J., Bowman, K. P., Hong, Y.,
904 Stocker, E. F., and Wolf, D. B.: The TRMM multisatellite precipitation analysis (TMPA):
905 Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales, *J.*
906 *Hydrometeor.*, 8, 38–55, doi:10.1175/JHM560.1, 2007.
- 907 Huo, Z., Feng, S., Kang, s., Dai, x., Li, w., and Chen, S.: The Response of Water-Land
908 Environment to Human Activities in Arid Minqin Oasis, Northwest China, *Arid Land Res.*
909 *Manag.*, 21(1), 21–36, 2007.
- 910 Jekeli, C.: Alternative methods to smooth the Earth's gravity field, Rep., 327, Dept. of Geod.
911 Sci. and Surv., Ohio State Univ., Columbus, 1981.
- 912 Jiao, J. J., Zhang, X., and Wang, X.: Satellite-based estimates of groundwater depletion in the
913 Badain Jaran Desert, China, *Nature Sci. Rep.*, 5, 8960, doi:10.1038/srep08960, 2015.
- 914 Kang, S., Su, X., Tong, L., Shi, P., Yang, X., Abe, Y., Du, T., Shen, Q., and Zhang, J.: The
915 impacts of human activities on the water-land environment of the Shiyang River basin, an arid
916 region in northwest China, *Hydrolog. Sci. J.*, 49(3), 2014.
- 917 Klein Goldewijk, K. and van Drecht, G.: Integrated modelling of global environmental
918 change: An overview of IMAGE 2.4, chap. HYDE 3: Current and historical population and
919 land cover, MNP – Netherlands Environmental Assessment Agency, Bilthoven, the
920 Netherlands, 93–112, 2006.
- 921 Kummerow, C., Barnes, W., Kozu, T., Shiue, J., and Simpson, J.: The Tropical Rainfall
922 Measuring Mission (TRMM) sensor package. *J. Atmos. Ocean. Tech.*, 15, 809–817, 1998.
- 923 Li, B., Rodell, M., Zaitchik, B. F., Reichle, R. H., Koster, R. D., and van Dam, T. M.:
924 Assimilation of GRACE terrestrial water storage into a land surface model: Evaluation and
925 potential value for drought monitoring in western and central Europe, *J. Hydrol.*, 446–447,
926 2012.
- 927 Li, F., Zhu, G., and Guo, C.: Shiyang River ecosystem problems and countermeasures, *Agr.*
928 *Sci.*, 4, 72–78, doi: 10.4236/as.2013.42012, 2013.



- 929 Long, D., Longuevergne, L., and Scanlon, B. R.: Uncertainty in evapotranspiration from land
930 surface modeling, remote sensing, and GRACE satellites, *Water Resour. Res.*, 50, 1131–1151,
931 doi:10.1002/2013WR014581, 2014.
- 932 Ma, J.Z., Wang, X.S., and Edmunds, W.M.: The characteristics of ground-water resources and
933 their changes under the impacts of human activity in the arid Northwest China—a case study
934 of the Shiyang River Basin, *J. Arid Environ.*, 61, 277–295, 2005.
- 935 Mitchell, T. D. and Jones, P. D.: An improved method of constructing a database of monthly
936 climate observations and associated high-resolution grids. *Int. J. Climatol.*, 25(6), 693–712,
937 2005.
- 938 Niu, G.-Y., Seo, K.-W., Yang, Z.-L., Wilson, C., Su, H., Chen, J., and Rodell, M.: Retrieving
939 snow mass from GRACE terrestrial water storage change with a land surface model,
940 *Geophys. Res. Lett.*, 34, L15704, doi:10.1029/2007GL030413, 2007.
- 941 Reichle, R. H. and Koster, R. D.: Bias reduction in short records of satellite soil moisture,
942 *Geophys. Res. Lett.*, 31, L19501, doi:10.1029/2004GL020938, 2004.
- 943 Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C. J., Arsenault,
944 K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D.,
945 and Toll, D.: The global land data assimilation system, *Bull. Amer. Meteor. Soc.*, 85(3), 381–
946 394, 2004.
- 947 Sheffield, J., Goteti, G., and Wood, E. F.: Development of a 50-yr high-resolution global
948 dataset of meteorological forcings for land surface modeling, *J. Climate*, 19 (13), 3088–3111,
949 2005.
- 950 Shiklomanov, I. A.: Assessment of water resources and water availability in the world,
951 Comprehensive assessment of the freshwater resources of the world, World Meteorological
952 Organization and the Stockholm Environment Institute, Stockholm, Sweden, 1997.
- 953 Steinfeld, H., Gerber, P., Wassenaar, T., Castel, V., Rosales, M., and de Haan, C.: Livestocks
954 long shadow: Environmental issues and options, FAO, Rome, Italy, ISBN 978-92-5-105571-
955 7, 2006.
- 956 Su, H., Yang, Z. L., Dickinson, R. E., Wilson, C. R., and Niu, G. Y.: Multisensor snow data
957 assimilation at the continental scale: The value of Gravity Recovery and Climate Experiment
958 terrestrial water storage information, *J. Geophys. Res.*, 115, D10104,
959 doi:10.1029/2009JD013035, 2010.
- 960 Sutanudjaja, E. H., de Jong, S. M., van Geer, F. C., and Bierkens, M. F. P.: Using ERS
961 spaceborne microwave soil moisture observations to predict groundwater head in space and
962 time, *Remote Sens. Environ.*, 138, 172–188, 2013.
- 963 Sutanudjaja, E. H., van Beek, L. P. H., de Jong, S. M., van Geer, F. C., and Bierkens, M. F.
964 P.: Large-scale groundwater modeling using global datasets: a test case for the {Rhine-



- 965 Meuse} basin, Hydrol. Earth Syst. Sci., 15(9), 2913–2935, doi:10.5194/hess-15-2913-2011,
966 2011.
- 967 Sutanudjaja, E. H., van Beek, L. P. H., de Jong, S. M., van Geer, F. C., and Bierkens, M. F.
968 P.: Calibrating a large-extent high-resolution coupled groundwater-land surface model using
969 soil moisture and discharge data. *Water Resour. Res.*, 50, 687–705.
970 doi:10.1002/2013WR013807, 2014.
- 971 Sutanudjaja, E. H., van Beek, L. P. H., Drost, N., de Graaf, I. E. M., de Jong, K., Peßenteiner,
972 S., Straatsma, M. W., Wada, Y., Wanders, N., Wissler, D. and Bierkens, M. F. P.: PCR-
973 GLOBWB 2.0: a 5 arc-minute global hydrological and water resources model, *Geosci. Model*
974 *Dev. Diss.*, 2016 (in prep).
- 975 Swenson, S. and Wahr, J.: Post-processing removal of correlated errors in GRACE data,
976 *Geophys. Res. Lett.*, 33(L08402), doi:10.1029/2005GL025285, 2006.
- 977 Swenson, S., Chambers, D., and Wahr, J.: Estimating geocenter variations from a
978 combination of GRACE and ocean model output, *J. Geophys. Res.*, 113(B08410),
979 doi:10.1029/2007JB005338, 2008.
- 980 Tangdamrongsub, N., Steele-Dunne, S. C., Gunter, B. C., Ditmar, P. G., and Weerts, A. H.:
981 Data assimilation of GRACE terrestrial water storage estimates into a regional hydrological
982 model of the Rhine River basin, *Hydrol. Earth Syst. Sci.*, 19, 2079–2100, doi:10.5194/hess-
983 19-2079-2015, 2015.
- 984 Tangdamrongsub, N., Ditmar, P. G., Steele-Dunne, S. C., Gunter, B. C., and Sutanudjaja, E.
985 H.: Exploring irregular flood events over Tonlé Sap basin in Cambodia using GRACE and
986 MODIS satellite observations combined with altimetry observation and hydrological models,
987 *Remote Sens. Environ.*, 181, 162 – 173, <http://dx.doi.org/10.1016/j.rse.2016.03.030>, 2016.
- 988 Tapley, B. D., Bettadpur, S., Ries, J. C., Thompson, P. F., and Watkins, M.: GRACE
989 Measurements of Mass Variability in the Earth System, *Science*, 305 (5683), 503–505, 2004.
- 990 van Beek, L. P. H.: Forcing PCR-GLOBWB with CRU data, Technical Report, Department of
991 Physical Geography, Utrecht University, Utrecht, The Netherlands,
992 <http://vanbeek.geo.uu.nl/suppinfo/vanbeek2008.pdf>, 2008.
- 993 Van Beek, L. P. H. and Bierkens, M. F. P.: The Global Hydrological Model PCR-GLOBWB:
994 Conceptualization, Parameterization and Verification, Technical Report, Department of
995 Physical Geography, Utrecht University, Utrecht, The Netherlands,
996 <http://vanbeek.geo.uu.nl/suppinfo/vanbeekbierkens2009.pdf>, 2009.
- 997 van Beek, L. P. H., Wada, Y., and Bierkens, M. F. P. (2011). Global monthly water stress: 1.
998 Water balance and water availability. *Water Resour. Res.*, 47, W07517.
999 doi:10.1029/2010WR009791.
- 1000 Vermote, E. F., Kotchenova, S. Y., and Ray, J. P.: MODIS surface reflectance user's guide,
1001 version 1.3, http://modis-sr.ltdri.org/guide/MOD09_UserGuide_v1_3.pdf, 2011.



- 1002 Vörösmarty, C. J., Leveque, C., and Revenga, C.: Millennium Ecosystem Assessment
1003 Volume 1: Conditions and Trends, chap. 7: Freshwater ecosystems, Island Press, Washington
1004 DC, USA, 165–207, 2005.
- 1005 Wada, Y., Wisser, D., and Bierkens, M. F. P.: Global modeling of withdrawal, allocation and
1006 consumptive use of surface water and groundwater resources. *Earth System Dynamics*, 5, 15–
1007 40. doi:10.5194/esd-5-15-2014, 2014.
- 1008 Wahr, J., Molenaar, M., and Bryan, F.: Time variability of the Earth’s gravity field:
1009 Hydrological and oceanic effects and their possible detection using GRACE, *J. Geophys.*
1010 *Res.*, 103(B12), 30205–30229, 1998.
- 1011 Wahr, J., Swenson, S., and Velicogna, I.: Accuracy of GRACE mass estimates, *Geophys. Res.*
1012 *Let.*, 33, L06401, doi:10.1029/2005GL025305, 2006.
- 1013 Wanders, N., Bierkens, M. F. P., Jong, S. M., Roo, A., and Karssenber, D.: The benefits of
1014 using remotely sensed soil moisture in parameter identification of large-scale hydrological
1015 models, *Water Resour. Res.*, 50, 6874–6891, doi:10.1002/2013WR014639, 2014.
- 1016 Weerts, A. H. and El Serafy G. Y. H.: Particle filtering and ensemble Kalman filtering for
1017 state updating with hydrological conceptual rainfall-runoff models, *Water Resour. Res.*, 42,
1018 W09403, doi:10.1029/2005WR004093, 2006.
- 1019 World Resources Institute (WRI): *World Resources: A Guide to the Global Environment*
1020 1998–99, World Resources Institute, Washington DC, USA, 1998.
- 1021 Zaitchik, B. F., Rodell, M., and Reichle, E. H.: Assimilation of GRACE terrestrial water
1022 storage data into a land surface model: Results for the Mississippi basin, *Amer. Meteor. Soc.*,
1023 *J. Hydrometeor.*, 9, 535–548, doi:10.1175/2007JHM951.1, 2008.
- 1024 Zheng, H., Wang, Z. J., Hu, S. Y., Malano, H., and ASCE, A. M.: Seasonal Water Allocation:
1025 Dealing with Hydrologic Variability in the Context of a Water Rights System, *J. of Water*
1026 *Resour. Plann. Manage.*, 139, 76–85, 2013.
- 1027 Zhu, J. F., Winter, C. L., and Wang Z. J.: Nonlinear effects of locally heterogeneous hydraulic
1028 conductivity fields on regional stream–aquifer exchanges, *Hydrol. Earth Syst. Sci. Discuss.*,
1029 12, 7727–7764, 2015.
- 1030
- 1031
- 1032
- 1033
- 1034
- 1035



1036

1037

1038

1039

1040

1041

1042

1043 **Table 1.** PCR-GLOBWB model parameters related to the TWS estimate. Parameters are
 1044 functions of spatial coordinates, except DDF which is a constant.

Parameter	Description	unit
$K_{\text{sat,up}}$	Saturated hydraulic conductivity of the upper soil storage	m/day
$K_{\text{sat,low}}$	Saturated hydraulic conductivity of the lower soil storage	m/day
SC_{up}	Storage capacity of the upper soil	m
SC_{low}	Storage capacity of the lower soil	m
$f_g^{\text{min}}, f_f^{\text{min}}, f_p^{\text{min}}, f_{np}^{\text{min}}$	Minimum soil depth fraction of grassland (g), forest (f), paddy irrigation (p), non-paddy irrigation (np)	-
$f_g^{\text{max}}, f_f^{\text{max}}, f_p^{\text{max}}, f_{np}^{\text{max}}$	Maximum soil depth fraction of grassland (g), forest (f), paddy irrigation (p), non-paddy irrigation (np)	-
J	Groundwater recession coefficient	1/day
DDF	Degree-day factor in the snow pack	°Cm/day
KC^{min}	Minimum crop coefficient	-

1045

1046 **Table 2.** TWS, SM and GWS estimated annual amplitude (A, cm) and phase (P, month) in 4
 1047 different basins computed between April 2002 and December 2010. Areally averaged values
 1048 for the entire Hexi Corridor are also given.

			Shiyang	Heihe	Desert	Shule	Areally-average
TWS	GRACE	A	2.05 ± 0.31	1.49 ± 0.21	1.79 ± 0.23	1.21 ± 0.27	1.43 ± 0.18
		P	6.97 ± 0.29	6.80 ± 0.27	6.49 ± 0.24	8.61 ± 0.42	7.05 ± 0.24
	EnOL	A	1.35 ± 0.16	0.90 ± 0.07	0.66 ± 0.07	0.37 ± 0.06	0.70 ± 0.06
		P	6.35 ± 0.23	5.61 ± 0.14	5.80 ± 0.19	5.40 ± 0.31	5.74 ± 0.16
EnKF 1D	A	1.61 ± 0.16	0.87 ± 0.10	1.05 ± 0.11	0.40 ± 0.11	0.80 ± 0.09	
	P	6.96 ± 0.19	6.80 ± 0.22	6.47 ± 0.19	8.35 ± 0.51	6.92 ± 0.23	
EnKF 3D	A	1.49 ± 0.13	0.80 ± 0.08	0.72 ± 0.07	0.26 ± 0.09	0.72 ± 0.07	
	P	6.42 ± 0.17	6.12 ± 0.19	6.40 ± 0.20	8.48 ± 1.02	6.44 ± 0.22	
SM	EnOL	A	1.03 ± 0.11	0.70 ± 0.06	0.62 ± 0.07	0.31 ± 0.05	0.59 ± 0.06
		P	5.77 ± 0.20	5.60 ± 0.16	5.82 ± 0.21	5.03 ± 0.32	5.62 ± 0.18
	EnKF 1D	A	0.88 ± 0.09	0.75 ± 0.09	0.99 ± 0.11	0.36 ± 0.10	0.67 ± 0.08
		P	6.55 ± 0.21	7.01 ± 0.22	7.08 ± 0.21	8.47 ± 0.54	7.26 ± 0.24
EnKF 3D	A	1.30 ± 0.10	0.66 ± 0.07	0.71 ± 0.08	0.12 ± 0.08	0.55 ± 0.07	
	P	5.59 ± 0.15	6.25 ± 0.20	6.44 ± 0.20	8.19 ± 0.37	6.32 ± 0.22	
GWS	EnOL	A	0.50 ± 0.08	0.19 ± 0.03	0.02 ± 0.004	0.09 ± 0.01	0.12 ± 0.01
		P	7.84 ± 0.29	7.13 ± 0.26	5.43 ± 0.34	6.91 ± 0.29	7.22 ± 0.21
	EnKF 1D	A	0.65 ± 0.05	0.12 ± 0.03	0.01 ± 0.01	0.05 ± 0.01	0.10 ± 0.01
		P	8.69 ± 0.16	7.82 ± 0.40	7.91 ± 1.90	8.49 ± 0.29	8.32 ± 0.25



1049

	EnKF 3D	A	0.70 ± 0.06	0.11 ± 0.02	0.02 ± 0.01	0.05 ± 0.01	0.10 ± 0.01
		P	8.52 ± 0.16	7.50 ± 0.31	7.76 ± 1.00	8.66 ± 1.33	8.26 ± 0.23

1050 **Table 3.** Averaged values and standard deviations of precipitation and model parameters for 4
1051 different basins.

1052

	Shiyang	Heihe	Desert	Shule
Precipitation (mm/month)	21 ± 12	13 ± 12	11 ± 2	8 ± 6
$K_{sat,up}$ (m/day)	0.42 ± 0.24	0.71 ± 0.69	1.16 ± 0.89	0.42 ± 0.15
$K_{sat,low}$ (m/day)	0.24 ± 0.15	0.61 ± 0.50	0.93 ± 0.74	0.24 ± 0.05
J (1/day)	0.0057 ± 0.0088	0.0024 ± 0.0049	0.0018 ± 0.0017	0.0013 ± 0.0017

1053

1054 **Table 4.** Statistical values of the GWS computed from the in situ well measurement and
1055 GRACE DA estimates between January 2007 and December 2010. The average values are
1056 computed by averaging the estimated statistical values from all well locations.

1057

		W1	W2	W3	W4	W5	Average value
Correlation coefficient [-]	EnOL	0.74	0.17	-0.04	-0.05	-0.53	0.06
	EnKF 1D	0.84	0.32	0.90	0.45	0.64	0.63
	EnKF 3D	0.82	0.49	0.85	0.51	0.83	0.70
RMS difference [cm]	EnOL	0.69	1.67	0.77	3.34	3.81	2.06
	EnKF 1D	0.58	1.63	0.40	2.56	2.58	1.55
	EnKF 3D	0.63	1.43	0.38	2.24	1.27	1.19

1058 **Table 5.** Long-term trends and standard deviations of the in situ data and the DA estimates.
1059 The RMS difference (RMSD) between the in situ data and the DA trend estimates are also
1060 provided.

1061

	W1	W2	W3	W4	W5	RMSD
In situ	-0.49 ± 0.03	0.01 ± 0.06	-0.60 ± 0.004	0.56 ± 0.12	-1.40 ± 0.03	0
EnOL	-0.57 ± 0.01	-0.64 ± 0.002	-0.01 ± 0.01	-1.69 ± 0.01	1.29 ± 0.02	1.62
EnKF 1D	-0.52 ± 0.02	-0.58 ± 0.04	-0.74 ± 0.02	-1.33 ± 0.08	-1.99 ± 0.13	0.93
EnKF 3D	-0.83 ± 0.02	-0.51 ± 0.03	-0.38 ± 0.01	-0.44 ± 0.08	-1.18 ± 0.06	0.54

1062 **Table 6.** Statistical values of the streamflow computed from the river stream gauge
1063 measurement and GRACE DA estimates between April 2002 and December 2010. The
1064 average values are calculated by averaging the estimated statistical values from both gauge
1065 locations.

		G1	G2	Average value
Correlation coefficient [-]	EnOL	0.82	0.76	0.79
	EnKF 1D	0.84	0.77	0.81
	EnKF 3D	0.84	0.78	0.81



NS coefficient [-]	EnOL	0.65	0.56	0.61
	EnKF 1D	0.69	0.57	0.63
	EnKF 3D	0.69	0.57	0.63
RMS difference [cm]	EnOL	5.49	3.09	4.29
	EnKF 1D	5.18	3.08	4.14
	EnKF 3D	5.23	3.04	4.14

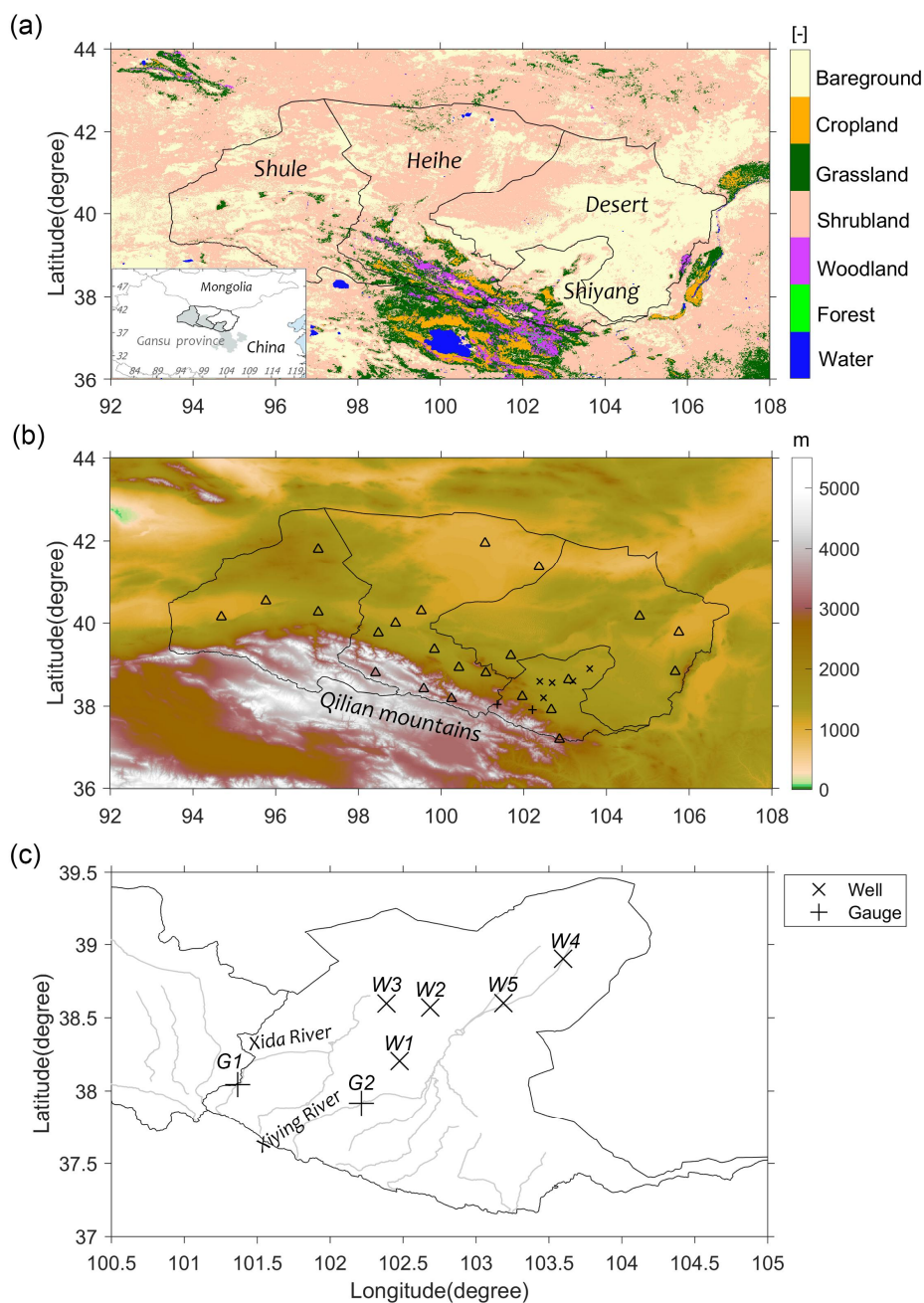
1066

1067

1068 **Table 7.** TWS, SM, GWS, and precipitation estimated long-term trends in 4 different basins
1069 computed between April 2002 and December 2010. Areally averaged values for the entire
1070 Hexi Corridor are also given.

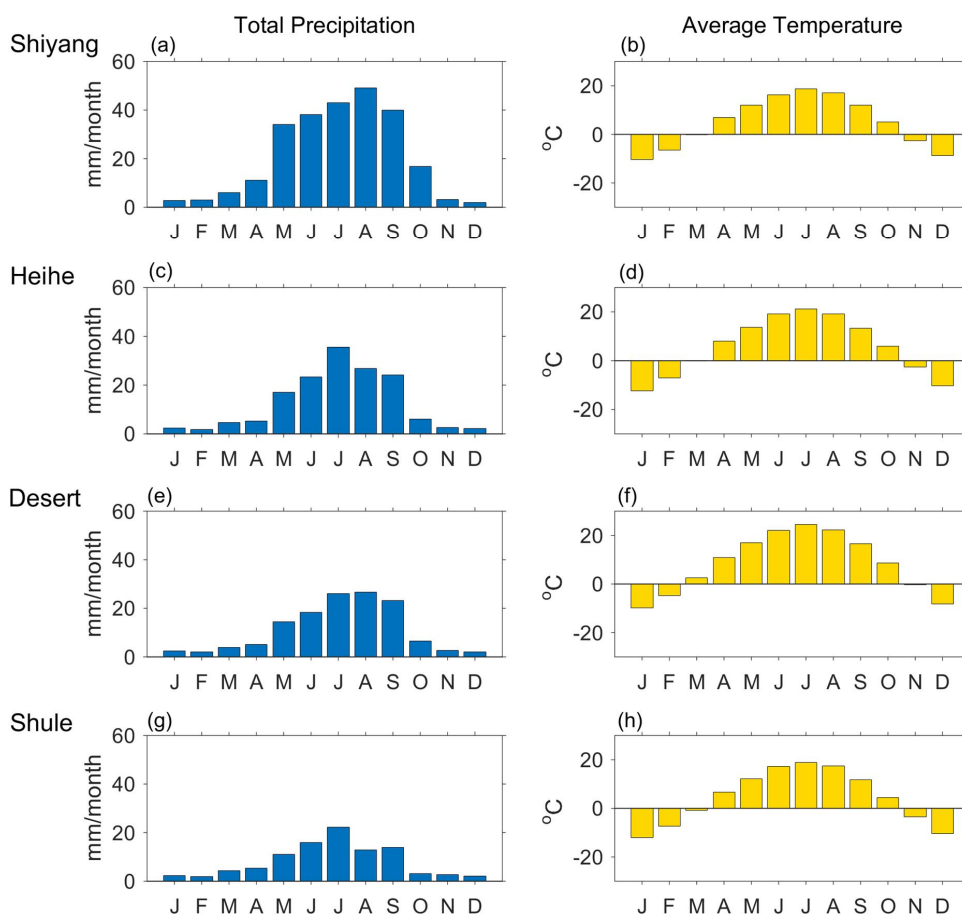
		Shiyang	Heihe	Desert	Shule	Areally-average
TWS (cm/yr)	GRACE	-0.73 ± 0.04	-0.64 ± 0.03	-0.72 ± 0.03	-0.34 ± 0.04	-0.59 ± 0.03
	EnOL	0.30 ± 0.02	0.24 ± 0.01	0.20 ± 0.01	0.18 ± 0.01	0.22 ± 0.01
	EnKF 1D	-0.72 ± 0.02	-0.41 ± 0.01	-0.33 ± 0.02	-0.34 ± 0.02	-0.39 ± 0.01
	EnKF 3D	-0.36 ± 0.02	-0.21 ± 0.01	-0.11 ± 0.01	-0.25 ± 0.01	-0.20 ± 0.01
SM (cm/yr)	EnOL	0.38 ± 0.02	0.21 ± 0.01	0.17 ± 0.01	0.14 ± 0.01	0.19 ± 0.01
	EnKF 1D	-0.11 ± 0.01	-0.20 ± 0.01	-0.29 ± 0.02	-0.22 ± 0.01	-0.23 ± 0.01
	EnKF 3D	0.10 ± 0.01	-0.12 ± 0.01	-0.12 ± 0.01	-0.14 ± 0.01	-0.11 ± 0.01
GWS (cm/yr)	EnOL	-0.08 ± 0.01	0.03 ± 0.004	0.02 ± 0.001	0.04 ± 0.002	0.02 ± 0.002
	EnKF 1D	-0.61 ± 0.01	-0.16 ± 0.004	-0.01 ± 0.002	-0.12 ± 0.001	-0.16 ± 0.002
	EnKF 3D	-0.39 ± 0.01	-0.09 ± 0.003	0.01 ± 0.001	-0.11 ± 0.001	-0.11 ± 0.002
Precipitation ((cm/month)/yr)		0.04 ± 0.01	0.04 ± 0.01	0.05 ± 0.01	0.02 ± 0.01	0.04 ± 0.01

1071



1072

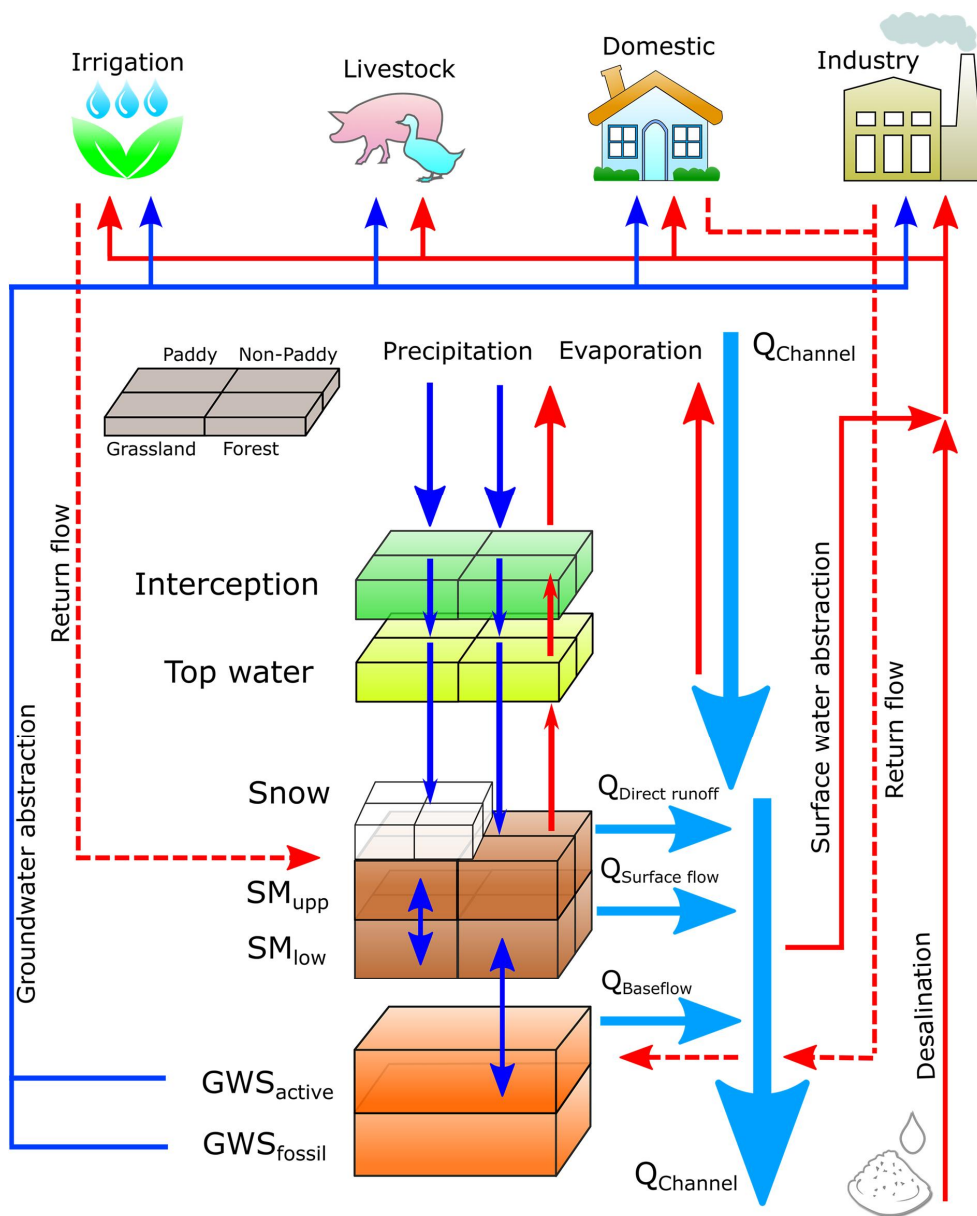
1073 **Figure 1.** Geography of the Hexi Corridor. (a) Land cover and division into individual regions
 1074 (Shiyang River Basin, Heihe River Basin, Shule River Basin, and a Desert), (b) Topography
 1075 and locations of the local meteorological stations (triangles), (c) Zoom-in on the Shiyang
 1076 River Basin, showing the locations of considered groundwater wells and river stream gauges.



1077

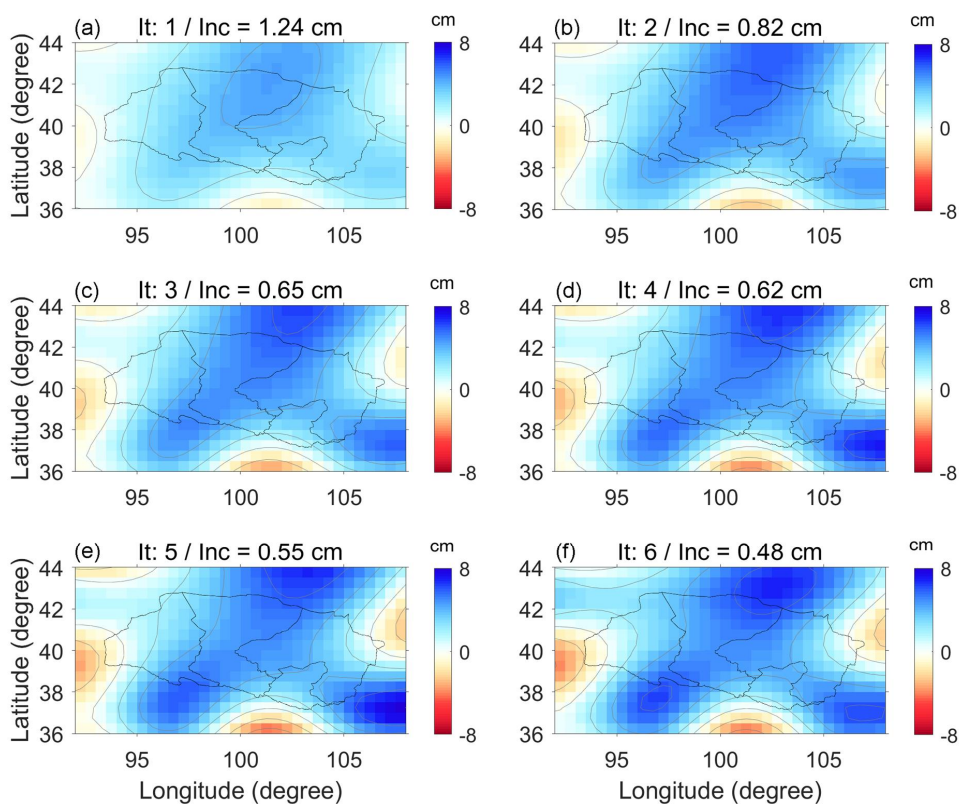
1078 **Figure 2.** Monthly total precipitation and averaged temperature over 4 regions of the Hexi
 1079 Corridor.

1080



1081

1082 **Figure 3.** The structure of PCR-GLOBWB hydrological model.



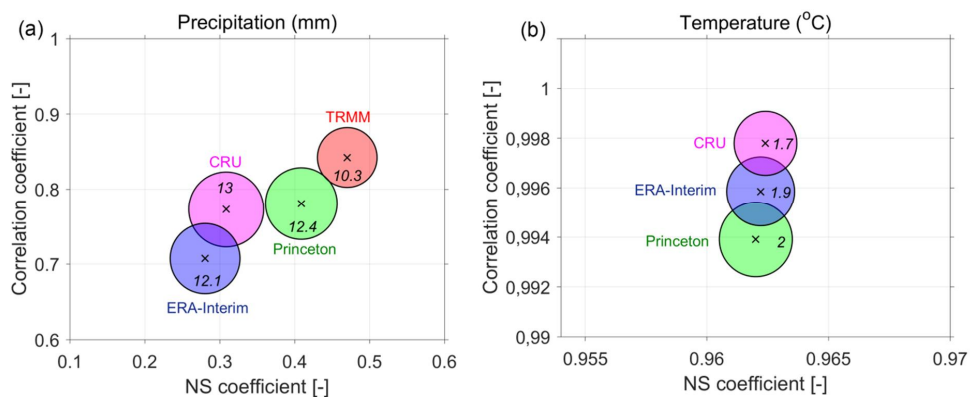
1083

1084 **Figure 4.** GRACE-derived TWS variation of October 2002. The signal restoration was
1085 applied to restore the signal mitigated by the applied spatial filter. After each iteration (It), the
1086 increment (Inc) was computed. The procedure was stopped after 6 iterations when the
1087 increment was lower than 0.5 cm (f).

1088

1089

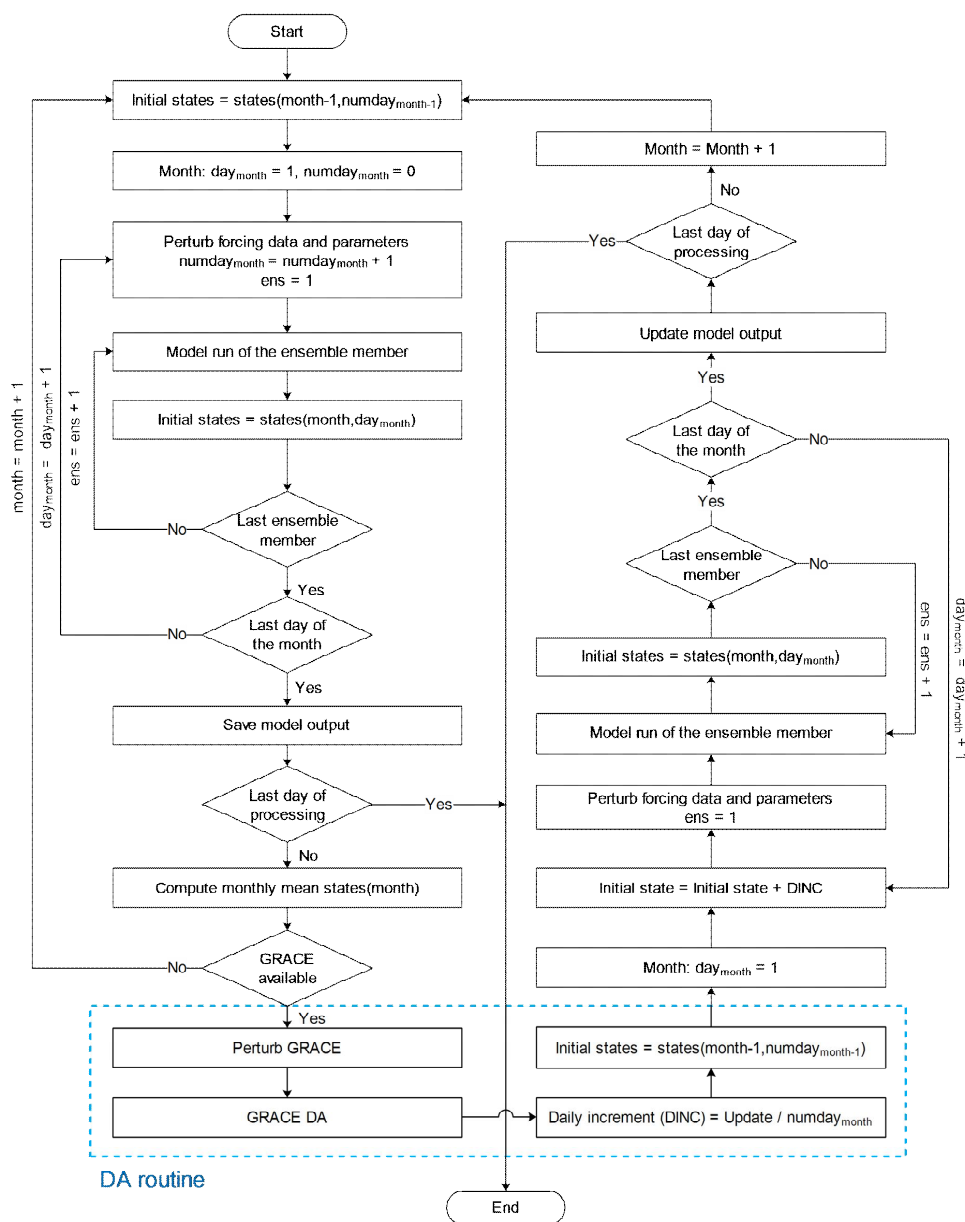
1090



1091

1092 **Figure 5.** The correlation coefficient, NS coefficient, and RMS difference computed between
1093 the local and different global forcing data. The RMS difference is shown as the radius of the
1094 circle (also explicitly provided as the number).

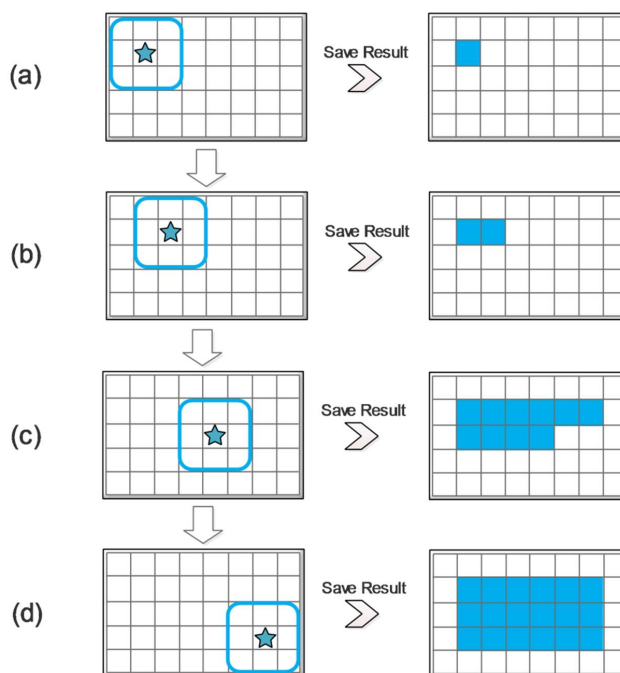
1095



1096

1097 **Figure 6.** DA diagram representing the disaggregation of monthly averaged TWS from
 1098 GRACE into the daily PCR-GLOBWB state estimates.

1099



1100

1101 **Figure 7.** Demonstration of EnKF 3D scheme, accounting for the spatially-correlated errors.
1102 For a centre grid cell, the state and observation matrices contain all TWS-related components
1103 of the neighbouring grid cells and the centre grid cell (left). The covariance matrices \mathbf{P}_e and \mathbf{R}
1104 are computed based on the data from these grid cells. Then, the EnKF is applied and the states
1105 of the centre grid cell are updated (right). The procedure is repeated through all grid cells.

1106

1107

1108

1109

1110

1111

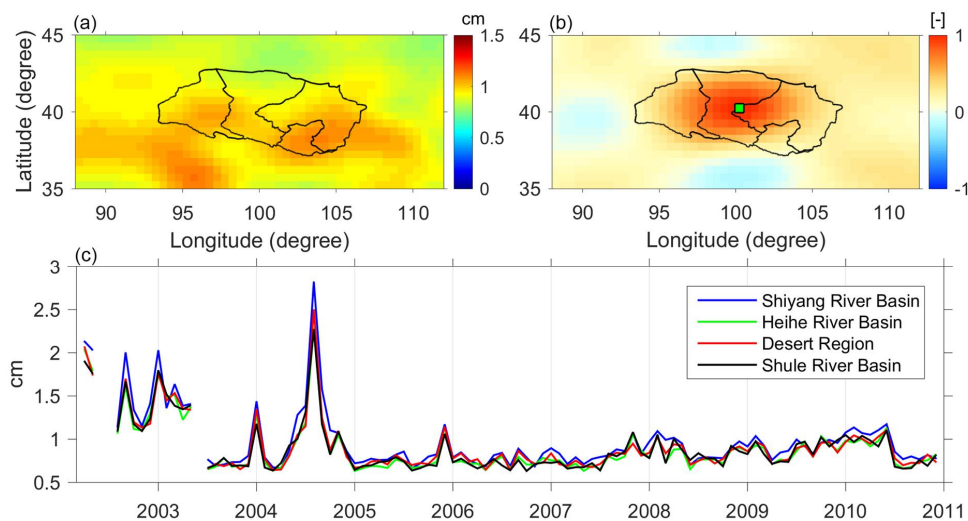
1112

1113

1114

1115

1116



1117

1118 **Figure 8.** Some statistics of errors in GRACE-derived TWS variation over the Hexi Corridor.
1119 The standard deviation (a) and the correlation coefficient with respect to the green point (b)
1120 for a sample month, October 2002, are shown in the top. The time-series of averaged standard
1121 deviation computed over 4 different basins are shown in the bottom plot (c).

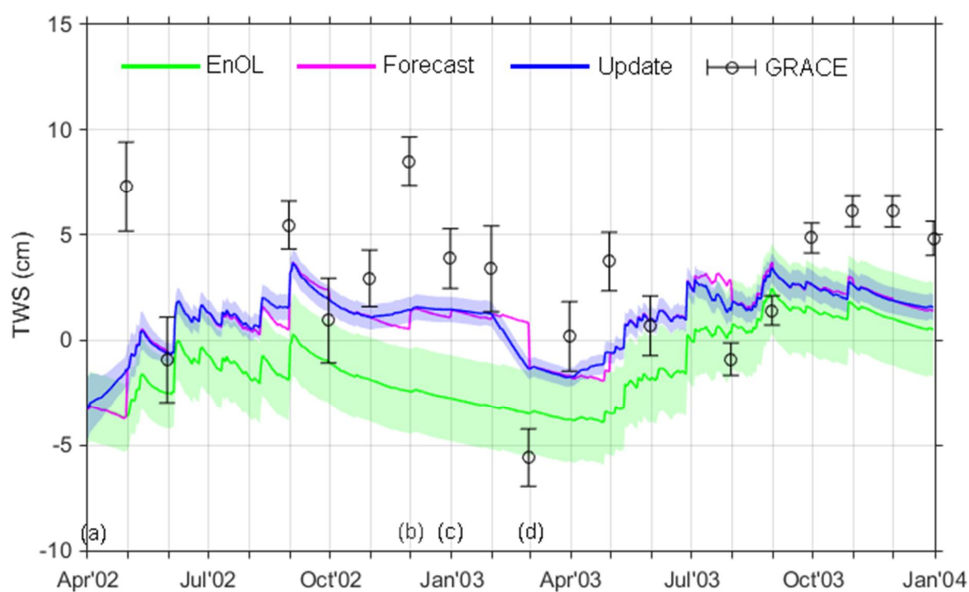
1122

1123

1124

1125

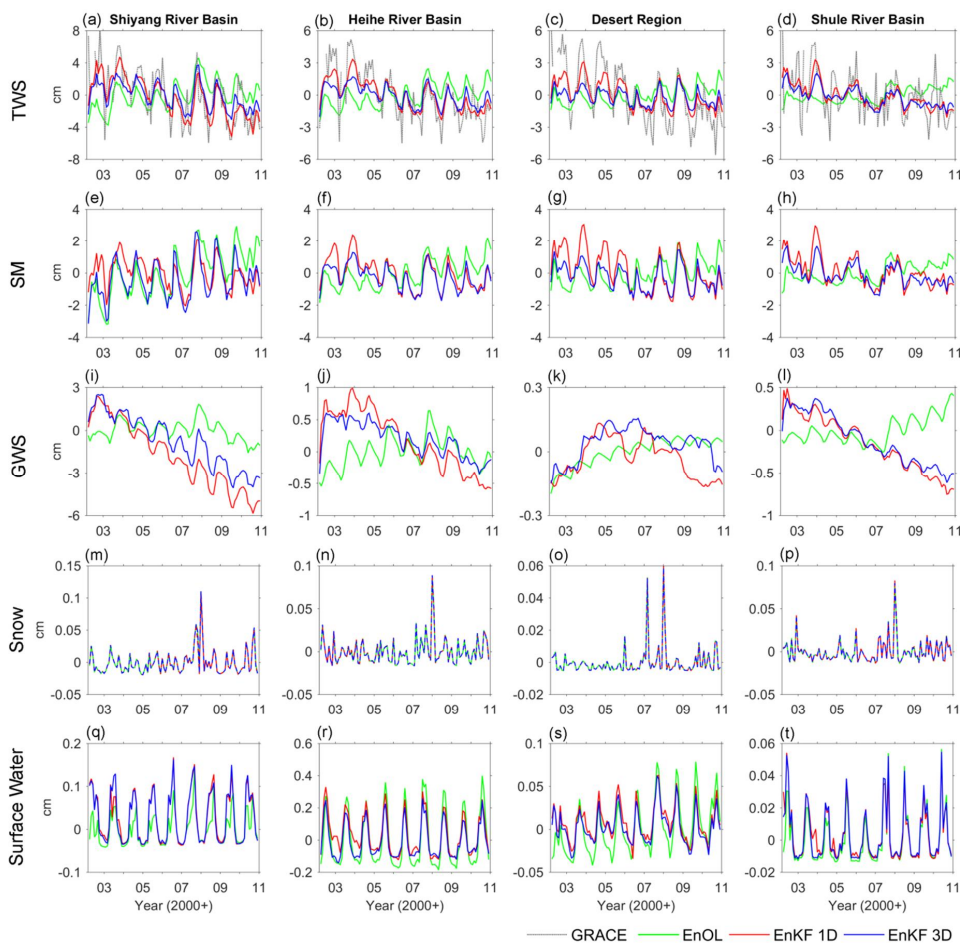
1126



1127

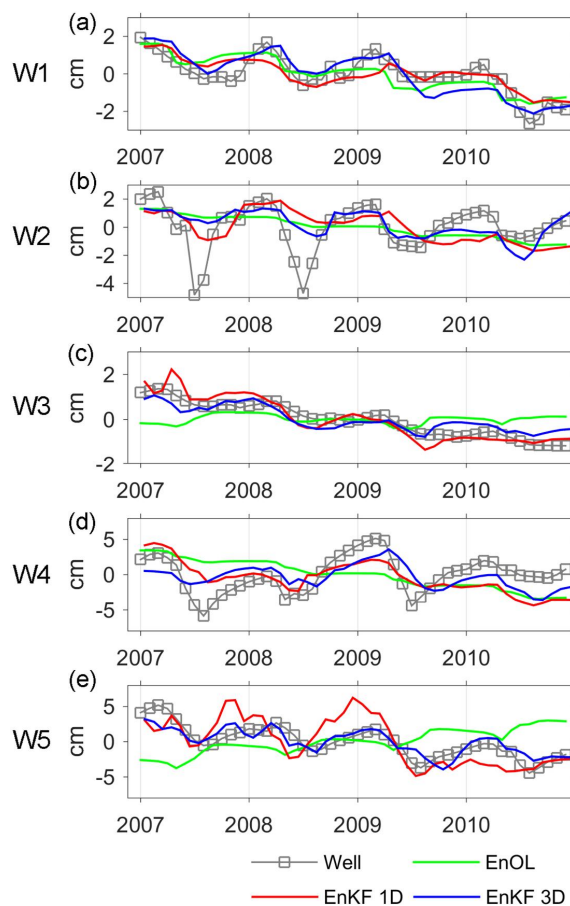
1128 **Figure 9.** Daily TWS variations estimated between 1 April 2002 and 31 December 2003,
1129 averaged over Shiyang River Basin. The mean value of the ensemble is given as the solid line,
1130 and the standard deviation is shown as the shaded envelope. The TWS estimates from model
1131 only (EnOL), GRACE DA forecast (EnKF before the update), GRACE DA update (EnKF
1132 after update), and GRACE observations are shown. The x-axis labels represent the first day of
1133 the month. Some features of the DA scheme regarding the identical TWS estimate seen at the
1134 beginning of the update (point a) and the observed spurious jumps (point b,c,d) are also
1135 shown.

1136



1137

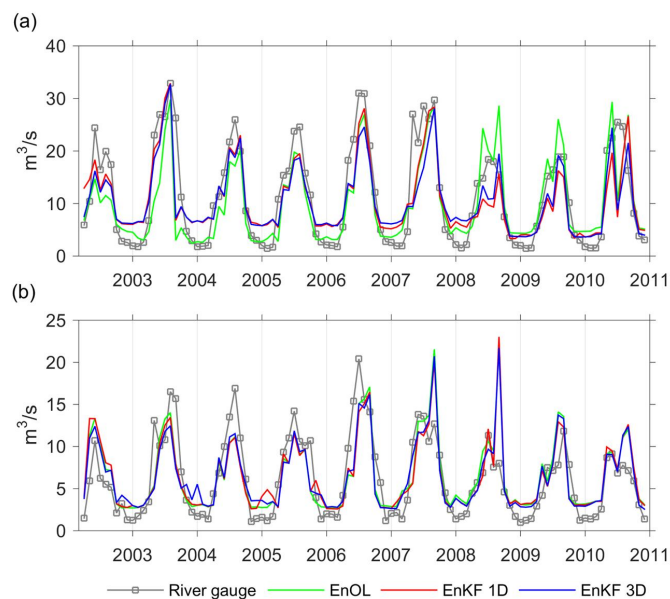
1138 **Figure 10.** Monthly TWS, soil moisture (SM), groundwater storage (GWS), snow, and
 1139 surface water variation estimated between April 2002 and December 2010 from the EnOL,
 1140 EnKF 1D, EnKF 3D, and GRACE observations over 4 basins.



1141

1142 **Figure 11.** Monthly GWS variation estimates from the in situ well measurements, as well as
1143 EnOL, EnKF 1D, and EnKF 3D results, between January 2007 and December 2010 at 5
1144 groundwater well locations. The chosen period is based on the availability of the in situ data.

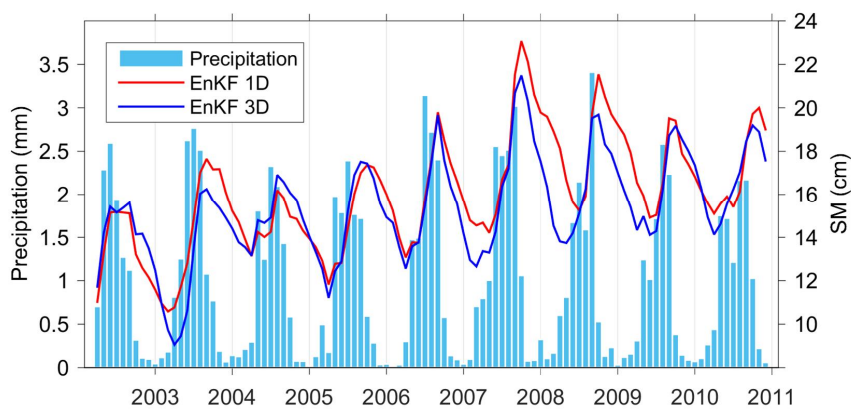
1145



1146

1147 **Figure 12.** Monthly streamflow estimates from the in situ river gauge measurements, as well
1148 as EnOL, EnKF 1D, and EnKF 3D results, between April 2002 and December 2010 at 2 river
1149 gauge locations, G1 (a) and G2 (b).

1150



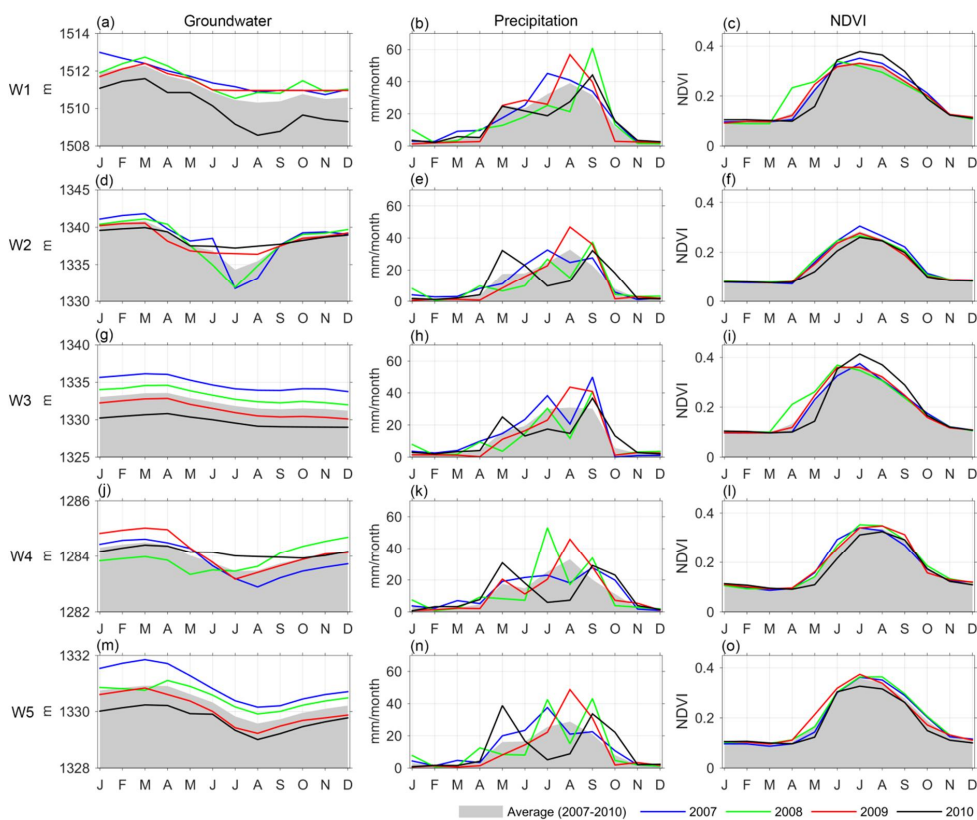
1151

1152 **Figure 13.** Monthly total precipitation (mm) and SM estimates (cm) from EnKF 1D and
1153 EnKF 3D results at river gauge G2 location.

1154

1155

1156



1157

1158 **Figure 14.** The monthly averaged groundwater head measurement (left), total precipitation
 1159 (middle) and NDVI (right) for 5 groundwater well locations. Precipitation and NDVI data are
 1160 reported as the average values within the circular areas of the 10-km radius. The long-term
 1161 average values between January 2007 and December 2010 are shown in the grey shed, and the
 1162 values in 2007, 2008, 2009, and 2010 are shown as blue, green, red, and black lines,
 1163 respectively. The period is chosen based on the availability of the well data.

1164