Delft University of Technology

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

Tangdamrongsub, Natthachet; Steele-Dunne, Susan C.; Gunter, Brian C.; Ditmar, Pavel G.; Sutanudjaja, Edwin H.; Sun, Yu; Xia, Ting; Wang, Zhongjing

Hydrology and
Earth System
Sciences

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

Natthachet Tangdamrongsub[1,2], Susan C. Steele-Dunne[3], Brian C. Gunter[1,4], Pavel G. Ditmar[1],
Edwin H. Sutanudjaja[5], Yu Sun[1], Ting Xia[6], and Zhongjing Wang[6,7]

[1]Department of Geoscience and Remote Sensing, Faculty of Civil Engineering and Geosciences,
Delft University of Technology, Delft, the Netherlands
[2]School of Engineering, Faculty of Engineering and Built Environment, The University of Newcastle,
Callaghan, New South Wales, Australia
[3]Department of Water Resources, Faculty of Civil Engineering and Geosciences, Delft University
of Technology, Delft, the Netherlands
[4]School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA, USA
[5]Department of Physical Geography, Faculty of Geosciences, Utrecht University,
Utrecht, the Netherlands
[6]Department of Hydraulic Engineering, Tsinghua University, Beijing, China
[7]State Key Lab of Hydroscience and Engineering, Tsinghua University, Beijing, China

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

Abstract. An accurate estimation of water resources dynamics is crucial for proper management of both agriculture and the local ecology, particularly in semi-arid regions. Imperfections in model physics, uncertainties in model land parameters and meteorological data, as well as the human impact on land changes often limit the accuracy of hydrological models in estimating water storages. To mitigate this problem, this study investigated the assimilation of terrestrial water storage variation (TWSV) estimates derived from the Gravity Recovery And Climate Experiment (GRACE) data using an ensemble Kalman filter (EnKF) approach. The region considered was the Hexi Corridor in northern China. The hydrological model used for the analysis was PCR-GLOBWB, driven by satellite-based forcing data from April 2002 to December 2010. The impact of the GRACE data assimilation (DA) scheme was evaluated in terms of the TWSV, as well as the variation of individual hydrological storage estimates. The capability of GRACE DA to adjust the storage level was apparent not only for the entire TWSV but also for the groundwater component. In this study, spatially cor-

related errors in GRACE data were taken into account, utilizing the full error variance–covariance matrices provided as a part of the GRACE data product. The benefits of this approach were demonstrated by comparing the EnKF results obtained with and without taking into account error correlations. The results were validated against in situ groundwater data from five well sites. On average, the experiments showed that GRACE DA improved the accuracy of groundwater storage estimates by as much as 25 %. The inclusion of error correlations provided an equal or greater improvement in the estimates. In contrast, a validation against in situ streamflow data from two river gauges showed no significant benefits of GRACE DA. This is likely due to the limited spatial and temporal resolution of GRACE observations. Finally, results of the GRACE DA study were used to assess the status of water resources over the Hexi Corridor over the considered 9-year time interval. Areally averaged values revealed that TWS, soil moisture, and groundwater storages over the region decreased with an average rate of approximately 0.2, 0.1, and 0.1 cm yr^{-1} in terms of equivalent water heights,

respectively. A particularly rapid decline in TWS (approximately $-0.4\,\mathrm{cm\,yr^{-1}}$) was seen over the Shiyang River basin located in the southeastern part of Hexi Corridor. The reduction mostly occurred in the groundwater layer. An investigation of the relationship between water resources and agricultural activities suggested that groundwater consumption required to maintain crop yield in the growing season for this specific basin was likely the cause of the groundwater depletion.

1 Introduction

The focus of this study is the Hexi Corridor. It is a semi-arid region located between the Gansu province of China and Mongolia (Fig. 1). A semi-arid region can be broadly classified as an area on the boundary of a larger desert, receiving just enough annual precipitation (300 mm or less) to sustain a limited amount of agriculture (Gong et al., 2004; Zhu et al., 2015). Inefficient use of the limited amount of surface water can often lead to overuse of groundwater resources and salinization of the soil (Cui and Shao, 2005). This can result in desertification, which not only reduces the amount of production but also may have long-term effects on the local ecology. All of this holds true for the Hexi Corridor (Wang et al., 2003).

Improving the water resources management of semi-arid regions requires accurate knowledge of the hydrological processes involved. For small areas, this can be partially obtained through a network of in situ measurement systems, such as meteorological stations, river gauges, groundwater wells, evaporation trays, etc. (Dahlgren and Possling, 2007; Huo et al., 2007; Kang et al., 2014; Ma et al., 2005; Du et al., 2014). While streamflow gauges provide integrated information for large catchment areas, point observations of hydrometeorological variables and even groundwater levels can be very local in scope. A sensor at a point several kilometres away may record significantly different values. For large scales ($> 10\,000\,\mathrm{km^2}$), such techniques are unlikely capable of delivering accurate results.

Two options for estimating the large-scale terrestrial water storage variation (TWSV) of a particular region are using observations from the Gravity Recovery And Climate Experiment satellite mission (GRACE; Tapley et al., 2004) or utilizing a regional or global hydrological model. A number of prior studies have reported on the potential of GRACE in the estimation of snow water equivalent (Niu et al., 2007), groundwater (Döll et al., 2014), and evapotranspiration (Long et al., 2014) in terms of temporal and spatial variability. However, GRACE only provides the total column of the water storage at a monthly timescale and large spatial scales ($> 300\,\mathrm{km}$). It is not possible to identify the contribution of separate hydrological components to the TWSV from GRACE data alone. On the other hand, a hydrological model
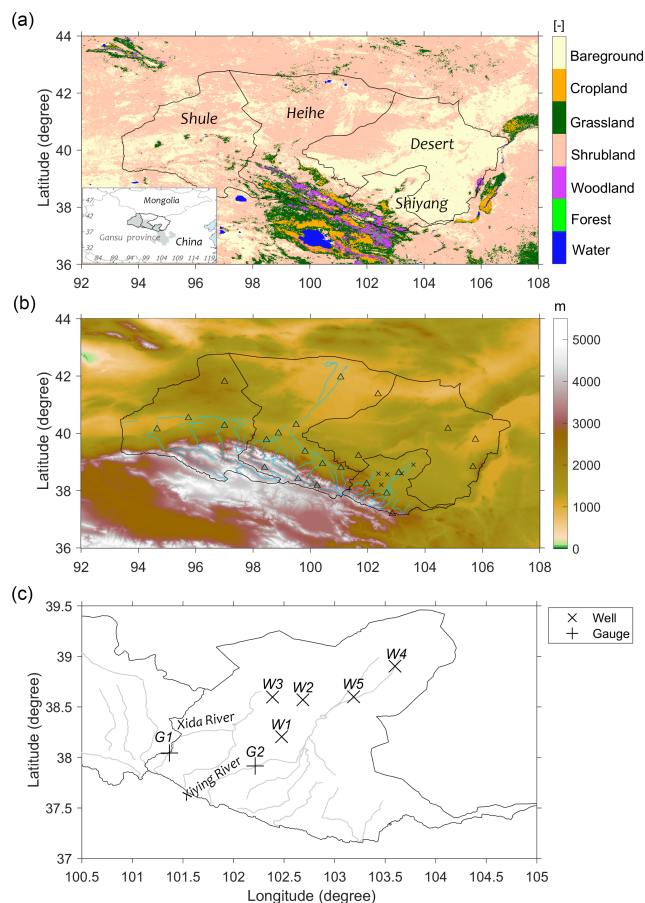

Figure 1. Geography of the Hexi Corridor. **(a)** Land cover and division into individual regions (Shiyang River basin, Heihe River basin, Shule River basin, and a desert), **(b)** topography and locations of the local meteorological stations (triangles), **(c)** zoom-in on the Shiyang River basin, showing the locations of considered groundwater wells ($\times$) and river stream gauges ($+$).

can be used to estimate the individual storage components at very high spatial and temporal scales. The major drawback of the model approach is mainly the significant uncertainties influenced by the quality of the model parameter calibration and the accuracy of the meteorological input data. In addition, hydrological models may suffer from inadequate process representations (model structure errors).

Data assimilation (DA) can be employed to combine the strengths of GRACE and hydrological models while mitigating their respective weaknesses. A number of studies have shown that GRACE DA can be used to improve the estimation of groundwater and streamflow (Zaitchik et al., 2008; Tangdamrongsub et al., 2015), snow water equivalent (Forman et al., 2012; Su et al., 2010), and as well as for evaluation of drought events (Houborg et al., 2012; Li et al., 2012). Different temporal and spatial resolution of GRACE observations and hydrological models require proper design of the DA scheme. Several DA schemes have been developed

to distribute GRACE observations into the model, which include using 5-day interpolated observations and updating the model every 5 days (Tangdamrongsub et al., 2015), using a monthly observation value and applying the model update only at the end of the month (Eicker et al., 2014), and using a monthly value and distributing the update as daily increments (Zaitchik et al., 2008; Forman et al., 2012; Girotto et al. 2016). Although all DA schemes are acceptable, the scheme proposed by Forman et al. (2012) is advantageous because it does not require an interpolation of the observations and can reduce the spurious jump of the water storage estimates caused by applying the update at the end of the month only. The only price to pay is the additional computational cost of running the model twice for the same month. A scheme similar to Forman et al. (2012) is used in this study. Spatial disaggregation is also needed to reconcile the difference in horizontal resolution between the observations and the model. Recent studies by Eicker et al. (2014) and Schumacher et al. (2016) suggested including the GRACE variance–covariance error information in the spatial disaggregation step. Both studies proposed using 500 km GRACE spatial resolution to mitigate the ill-posedness of the error covariance matrices in the spatial domain. In line with Eicker et al. (2014) and Schumacher et al. (2016), the assimilation scheme in this study accounts for spatially correlated errors by using full error variance–covariance matrices of GRACE data. This study will show that considering the GRACE error correlations leads to an improvement of the state estimates. Particularly, the signal-to-noise ratio (SNR) of the TWSV is much lower than in the river basins considered in the previous studies, e.g. the Mississippi (Zaitchik et al., 2008), Rhine (Tangdamrongsub et al., 2015), and Mackenzie (Forman et al., 2012).

Approximately 9 years of GRACE data – between April 2002 and December 2010 – are considered in this study. GRACE observations are assimilated into the PCRaster Global Water Balance (PCR-GLOBWB; van Beek et al., 2011; Sutanudjaja et al., 2014; Wada et al., 2014) hydrological model over the Hexi Corridor. TWS is computed from PCR-GLOBWB as the sum of all the hydrological components (soil moisture, groundwater, surface water, inundated water, interception, and snow). The previous studies showed very good agreement of PCR-GLOBWB based estimates with GRACE observations in several river basins (Wada et al., 2014; Tangdamrongsub et al., 2016). However, the performance of PCR-GLOBWB has not yet been evaluated over the Hexi Corridor. In addition, to date, the model has not been incorporated into any GRACE DA scheme, making this study the first attempt to do so. Investigating the added value of GRACE DA in the Hexi Corridor is the main objective of this study.

First of all, the impact of GRACE DA and the effect of taking correlations in GRACE errors into account are assessed. Both the total terrestrial water storage and the individual hydrological storage compartments are considered.

Next, the results of the GRACE DA are validated with independent in situ data. The agreement is analysed in terms of the correlation coefficient, Nash–Sutcliffe coefficient, and root mean square difference (RMSD). The groundwater storage variation (GWSV) and streamflow estimates after GRACE DA are validated with the well and river stream gauge measurements, respectively.

Finally, results from this GRACE DA study are used to assess the status of water resources over the Hexi Corridor. The connections between the water storage (including groundwater consumption) and agriculture in the area are also presented and discussed. At that stage, we use precipitation data from the Tropical Rainfall Measuring Mission (TRMM; Huffman et al., 2007) and the Moderate Resolution Imaging Spectroradiometer (MODIS) derived Normalized Difference Vegetation Index (NDVI; Huete et al., 2002).

2 Study region

The Hexi Corridor is a long and narrow area between the Qilian Mountains and southern Mongolia (Fig. 1a). The region's elevation ranges from 900 m in the northern downstream zone (Inner Mongolia) to 5200 m in the southern upstream area (Qilian Mountains) (Fig. 1b). The region is comprised of four typical inland arid and semi-arid regions (Zhu et al., 2015): the Shiyang River basin ($41\,600\,km^2$), the Heihe River basin ($143\,000\,km^2$), the Shule River basin ($157\,000\,km^2$), and a desert region ($152\,445\,km^2$) (Geng and Wardlaw, 2013; Zhu et al., 2015). Located next to the Gobi Desert, most parts of the region have a cold desert climate (Peel et al., 2007), where precipitation is relatively low to sustain vegetation or crops. Approximately 60 to 80 % of the annual rainfall is concentrated during the time frame from June to September. The inland rivers mainly originate from the Qilian Mountains and disappear after entering the midstream/downstream plains and oases. As such, the southern part of the region is more favourable for agriculture.

The four basins have distinct characteristics. First, the smallest river basin, Shiyang, has eight main river streams, including the Xida and Xiying rivers (Fig. 1c). The annual rainfall and the mean temperature are approximately 250 mm and 5 °C (Fig. 2a, b), respectively. The Shiyang River basin is considered the wettest basin compared to the others, with relatively high mean total renewable annual water resources of approximately 1.66 billion m^3 (Zheng et al., 2013). However, a highly developed economy and population growth in the past decade have resulted in a severe water resources overexploitation problem (Zheng et al., 2013). The Heihe River basin has a semi-arid climate and the mean daily temperature of ~ 6 °C (Fig. 2d). The average annual rainfall is ~ 150 mm (Fig. 2c) with high heterogeneity both in temporal and spatial distribution. The mean total annual available water resources are estimated at 3.7 billion m^3 (Hu, 2015). Similar to the Shiyang River basin, increased water exploita-
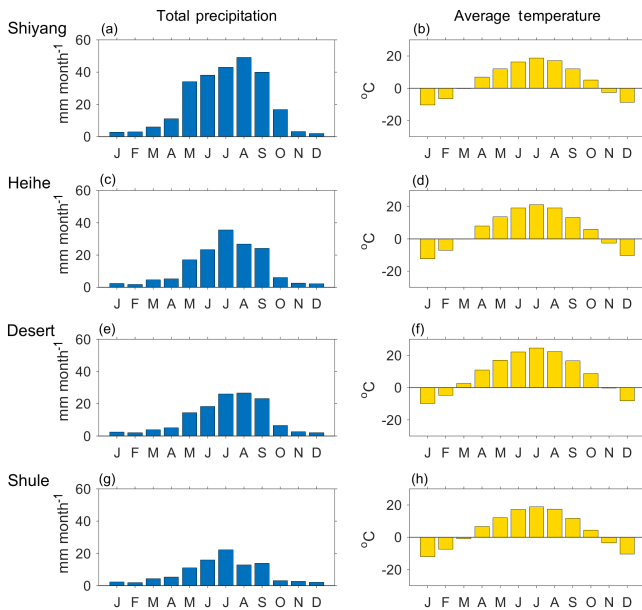
Figure 2. Monthly total precipitation and averaged temperature over four regions of the Hexi Corridor.

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

tion, increasing population, and changing climate have aggravated the damage to the downstream ecology. The Shule River basin has an arid climate, the mean temperature there is around 4 °C (Fig. 2f), and the average annual rainfall is only approximately 98 mm (Fig. 2e). Compared to the Shiyang River basin, the Shule River basin is approximately 4 times as large in terms of surface area, but has similar mean total annual water resources (~ 1.6 billion m^3; Hu, 2015). The district irrigation areas are mainly located in the middle of the Shule River basin. Agricultural water consumption accounts for more than 80 % of the total water use. Finally, the desert region has an extreme continental desert climate with an average temperature of 8 °C and the annual rainfall of ~ 130 mm. Extensive groundwater abstraction was also observed over the region (Jiao et al., 2015).

3   Hydrology model

The global distributed hydrological model PCR-GLOBWB (van Beek et al., 2011; Sutanudjaja et al., 2017) simulates spatial and temporal continuous fields of fluxes and storages in various water storage components (soil moisture, groundwater, surface water, inundated water, interception, and snow). The model version used here (Sutanudjaja et al., 2017) has a spatial resolution of 30 arcmin (approximately 50 km at the Equator), and a temporal resolution of 1 day. Figure 3 illustrates the structure of PCR-GLOBWB model. The model includes two soil layers (SM$_{upp}$, SM$_{low}$), an underlying hydrologically active and replenishable groundwater (GWS$_{active}$) layer, a non-renewable groundwater (GWS$_{fossil}$) layer, as well as interception, sur-
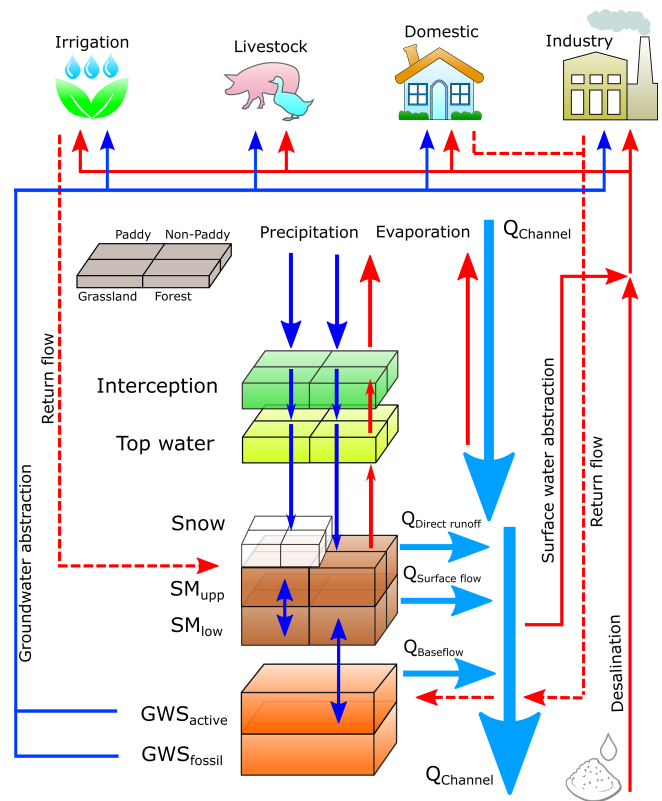
face water, and snow stores. The non-renewable groundwater is available for abstraction to satisfy water demands once the overlying hydrologically active groundwater storage is depleted. For soil, snow, inundated top water, and interception stores, an individual grid cell is divided into subgrids associated with different types of topography, vegetation phenology, and soil properties, as well as land cover types. Specifically, there are four types of land covers defined: short natural vegetation, tall natural vegetation, irrigated non-paddy field, and irrigated paddy field. Soil components include the upper layer (SM$_{upp}$, 0–30 cm) and the lower layer (SM$_{low}$, 30–150 cm). The snow component includes snow water equivalent (SWE), as well as snow-free water (SFW) representing the storage of melted snow. The water stored in the stream channels and lakes is also included in the TWS estimate. Based on the structure of PCR-GLOBWB, the total water storage (TWS) is computed as the sum of 27 different water storage components: 8 soil moisture layers, 2 groundwater layers, 4 interception layers, 8 snow layers, 4 inundated top water layers, and 1 surface water layer.

For each grid cell and for each daily time step, the model determines the water balance in two vertically stacked soil layers and the groundwater store. The model also computes the vertical water exchanges between the soil layers and between the inundated top water layer and the atmosphere, i.e. rainfall and snowmelt, percolation, and capillary rise, as well

as evaporation and transpiration fluxes. The active groundwater store underlies the soil, is fed by net groundwater recharge, discharges to baseflow as a linear reservoir, and is exempt from the direct influence of evaporation and transpiration fluxes. However, capillary rise from the active groundwater store can occur depending on the simulated groundwater storage, the soil moisture deficit, and the unsaturated hydraulic conductivity. Fluxes are simulated according to the different land cover types. The model includes a physically based scheme for infiltration and runoff, resulting in the direct runoff, interflow, as well as groundwater baseflow and recharge. River discharge is calculated by accumulating and routing the specific runoff along the drainage network. For further details, including model parameterization, the reader is referred to the technical reports and other relevant publications (van Beek and Bierkens, 2009; van Beek, 2008; Sutanudjaja et al., 2011, 2014).

4 Data and data processing

4.1 GRACE data

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

– SHCs of degree 1 provided by Swenson et al. (2008) were restored, and all five coefficients of degree 2 were replaced by the values estimated from satellite laser ranging (Cheng and Tapley, 2004).

– SHC variations were computed by removing the long-term mean (computed between April 2002 and December 2010) from each monthly solution.

– A destriping filter (Swenson and Wahr, 2006) was applied to the SHC variations. The filter used a fifth-degree polynomial (Savitzky–Golay) over a five-point window to remove the correlations; orders below 8 remained unchanged.

– An additional 250 km radius Gaussian smoothing (Jekeli, 1981) was applied to SHC variations to suppress high-frequency noise, and the TWS variations ($\Delta\sigma$ (m)) were then computed using (Wahr et al., 1998)

$$\Delta\sigma(\theta,\phi) = \sum_{l=1}^{60}\sum_{m=-l}^{l} W_l \Delta\overline{C}_{lm}\hat{Y}_{lm}(\theta,\phi),\qquad(1)$$

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

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

4.2 Forcing data

The forcing data required by PCR-GLOBWB are precipitation, air temperature, and potential evapotranspiration. Tangdamrongsub et al. (2015) showed that the use of high-quality precipitation data might lead to better estimates of hydrological fluxes (e.g. TWSV and streamflow). In principle, local precipitation and surface temperature measurements could be obtained from the China Daily Ground Climate Dataset provided by the China Meteorological Data Sharing Service System (http://www.cma.gov.cn/en2014/services/ProductsService). A total of 23 weather stations were found over the Hexi Corridor (see Fig. 1b). However, the measurements were spatially sparse and did not cover the entire region. Therefore, the global precipitation data were used to achieve a better spatial coverage. Four global precipitation products were considered for inclusion:

– the European Centre for Medium-range Weather Forecasts (ERA-Interim, spatial resolution: $0.75° \times 0.75°$; Dee et al., 2011),
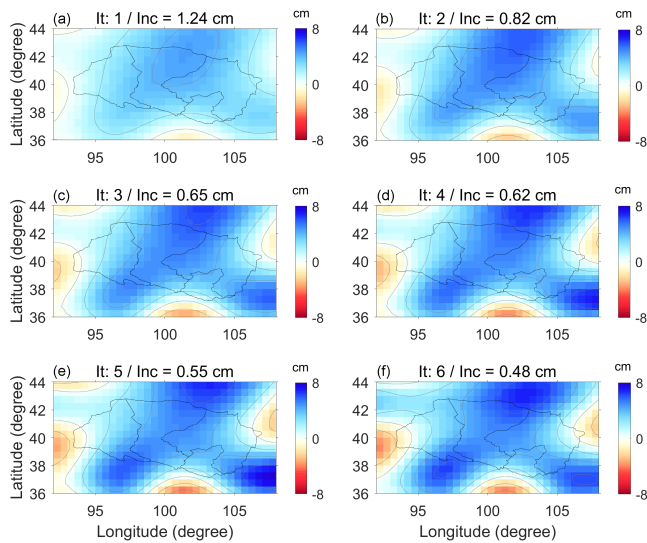
Figure 4. GRACE-derived TWS variation of October 2002. The signal restoration was applied to restore the signal mitigated by the applied Gaussian filter. After each iteration (It), the increment in each cell was computed. The procedure was stopped after six iterations, when the maximum increment (Inc) was lower than 0.5 cm (**f**).

- the Tropical Rainfall Measuring Mission (TRMM 3B42, spatial resolution: $0.25° × 0.25°$; Huffman et al., 2007; Kummerow et al., 1998),

- the Climate Research Unit dataset (CRU, spatial resolution: $0.5° × 0.5°$; Mitchell and Jones, 2005; van Beek, 2008), and

- Princeton's Global Meteorological Forcing Dataset (Princeton, spatial resolution: $0.5° × 0.5°$; Sheffield et al., 2005).

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

A similar procedure was used to compare the air temperature data from ERA-Interim, CRU, and Princeton. The statistical estimates are shown in Fig. 5b. Although the results from all products were very similar, CRU provided the high-
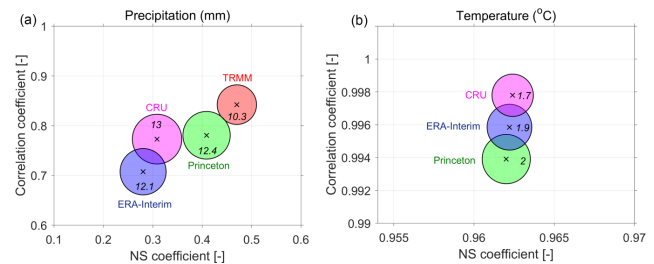

Figure 5. The correlation coefficient, NS coefficient, and RMSD (root mean square difference) computed between the local and different global forcing data. The rms difference is shown as the radius of the circle (also explicitly provided as the number).

est data quality in terms of correlation and RMSD values, and therefore it was used as the temperature input. As far as potential evapotranspiration is concerned, few data are available for this region, so the data from van Beek (2008) were used.

4.3 Validation data

4.3.1 Groundwater

Monthly groundwater well measurements at five locations (Fig. 1c) were obtained from the ground network maintained by the Shiyang River Basin Management Bureau, and Institute of Water Resources and Hydropower of Gansu Province. The in situ data were provided in the form of piezometric heads (relative to the mean sea level), which needed to be converted to units of storage. For such a task, several parameters, e.g. storage coefficient and specific yield, are required, but they are not available over the Hexi Corridor. To solve that problem, a scale factor computed using the information from GRACE and soil moisture (SM) from the Global Land Data Assimilation System (GLDAS; Rodell et al., 2004) was used for the conversion using the approach outlined by Tangdamrongsub et al. (2015). As discussed in Tangdamrongsub et al. (2015), it is ideally preferred to use the in situ soil moisture data to represent the SM term, but they are not available at the well locations. The soil moisture estimated from remote sensing was also not appropriate due to the limitation of the penetration depth. The use of SM from PCR-GLOBWB is avoided to reduce the bias when compared the adjusted well measurements to the final DA result. Therefore, the GLDAS-derived SM was used.

The adjustment procedure was as follows. First, GLDAS-based soil moisture storage variations (SMSVs) were removed from GRACE-derived TWSV. Four variants of GLDAS model (NOAH, CLM, MOSAIC, and VIC; see Rodell et al., 2004) were considered and the average SMSV value was calculated. Taking into account that SMSVs and GWSVs are the major contributions to TWS variations, this resulted in GWSV (GWSV$_{\text{(GRACE−SMSV)}}$). Then, by conducting a regression analysis between the

monthly time series of piezometric head variation (Δh) and $\Delta \mathrm{GWS}_{(\mathrm{GRACE-SMSV})}$ at each individual location, a bias (b) and a scale factor (f) were estimated using the following linear relationship:

$$\Delta \mathrm{GWS}_{(\mathrm{GRACE-SMSV})} + e = b + f \cdot \Delta h, \tag{2}$$

where e indicates the observation error. Finally, the estimated bias ($\hat{b}$) and scale factor ($\hat{f}$) were used to convert the in situ head measurements into groundwater storage variation ($\mathrm{GWSV}_{\mathrm{in\ situ}}$) as

$$\mathrm{GWSV}_{\mathrm{in\ situ}} = \hat{b} + \hat{f} \cdot \Delta h. \tag{3}$$

4.3.2 Streamflow

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

4.3.3 Normalized Difference Vegetation Index (NDVI)

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

$$\mathrm{NDVI} = \frac{\rho_{\mathrm{NIR}} - \rho_{\mathrm{R}}}{\rho_{\mathrm{NIR}} + \rho_{\mathrm{R}}}, \tag{4}$$

where ρ_{NIR} and ρ_{R} are the observed surface reflectances in the near-infrared and red portions of the electromagnetic spectrum, respectively. The monthly averaged NDVI was then computed based on the derived 8-day NDVI values.

5 Methodology and implementation

5.1 Ensemble Kalman filter (EnKF)

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

$$d = \mathbf{H}\psi + \epsilon; \epsilon \sim N(0, \mathbf{R}), \tag{5}$$

where d is an $m \times 1$ vector containing the GRACE observations for the month of interest, and $\mathbf{H}$ is a measurement operator which relates the PCR-GLOBWB state ψ to the observation vector d. Notice that the number of observations is equal to the number of grid cells because the GRACE-based estimates are obtained for all the grid cells of the PCR-GLOBWB model (see Sect. 4.1). The uncertainties in the observations are given in the random error ϵ, which is assumed to have zero mean and covariance matrix $\mathbf{R}_{m \times m}$. As the sum of all state elements at a given cell is equal to TWSV, the $\mathbf{H}$ matrix is defined as

$$\mathbf{H} = \tag{6}$$
$$\begin{bmatrix} \begin{pmatrix} (1\,1\,1\ldots1)_{1 \times n} & 0 & \cdots & 0 \\ 0 & (1\,1\,1\ldots1)_{1 \times n} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & (1\,1\,1\ldots1)_{1 \times n} \end{pmatrix} \end{bmatrix}_{m \times nm.}$$

Let the ensemble of the states be stored in a matrix $\mathbf{A}_{nm \times N} = (\psi_1, \psi_2, \psi_3, \ldots, \psi_N)$, where N is the number of ensemble members. Then, the ensemble perturbation matrix is defined as $\mathbf{A}' = \mathbf{A} - \overline{\mathbf{A}}$, where the matrix $\overline{\mathbf{A}}$ is of the same size as $\mathbf{A}$ and filled in with the mean values computed from all ensemble members. Similarly, the GRACE observation vector is stored in the matrix $\mathbf{D}_{m \times N} = (d_1, d_2, d_3, \ldots, d_N)$, in which each column is a replicate of the observation but perturbed with random noise $\sim N(0, \mathbf{R})$. The analysis equation can be expressed as (Evensen, 2003)

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

with

$$\mathbf{K} = \mathbf{P}_{\mathrm{e}}\mathbf{H}^T \left(\mathbf{H}\mathbf{P}_{\mathrm{e}}\mathbf{H}^T + \mathbf{R}\right)^{-1}, \tag{8}$$

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

$$\mathbf{P}_{\mathrm{e}} = \mathbf{A}'(\mathbf{A}')^T / (N - 1). \tag{9}$$

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

In the initialization phase, which was needed to obtain the initial states, the model was spun up between 1 January and

31 December 2000 as a hot start. This time interval was sufficient to reach the dynamic equilibrium. The initial state ψ for 31 December 2000 obtained this way was perturbed to yield $N = 100$ ensemble members $\psi_i, i = 1, 2, 3, \ldots, N$. The N ensemble runs between 1 January 2001 and 31 March 2002 were then conducted independently based on the perturbed initial states. This resulted in an ensemble spread of the estimated states. The model was then propagated in time between 1 April 2002 and 31 December 2010 without assimilating any observation. This case is referred to hereafter as the Ensemble Open Loop (EnOL). For the EnKF, the model was also propagated beginning from 1 April 2002, but the observations (when available) were assimilated.

The processing diagram is shown in Fig. 6 and follows the methodology introduced by Forman et al. (2012). The state is first propagated in time from the first to the last day of the month without applying DA, and the monthly averaged states are calculated from the daily values. When the GRACE observation for that month is available, the DA routine is activated (otherwise, the model continues propagating to the next month without applying DA). The DA routine computes the monthly averaged update $\Delta \mathbf{A}$ of the TWS-related states (see Eq. 6). The daily increment (DINC) of the update is then computed by dividing the monthly averaged update by the total numbers of days in that month ($\text{numday}_{\text{month}}$). The model propagation is then restarted (second run), using the last day of the previous month (month-1, $\text{numday}_{\text{month}-1}$) as the initial state. In this second run, the DINC is added to the current states every day up to the last day of the month. The DA scheme is repeated for each month up to the end of the study period.

Spatial correlations of model errors and observation errors were also taken into account in view of the fact that the latter are highly correlated at neighbouring $0.5° \times 0.5°$ grid cells. De Lannoy et al. (2009) proposed a so-called 3D-Fm (three-dimensional fine scale with multiple observations) approach, which is called "EnKF 3D" in this paper. The approach only considers the spatial correlations between the neighbouring grid cells. In this study, the neighbouring grid cells were assumed to be the ones inside the Gaussian smoothing radius applied, i.e. 250 km. This reduced the computational cost, as only a small subset of cells pairs was considered instead of all cells pairs. That approach was applied not only to observation errors but also to model errors in TWSV and TWS-related components in this study. The EnKF 3D scheme is illustrated in Fig. 7. For a particular grid cell (centre grid cell), all TWS-related components of the neighbouring grid cells and the 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 is the number of the considered grid cells. The matrix notation with superscript s (e.g. $\mathbf{A}^s$) is only used to emphasize the cell-dependent version, and it can be substituted into the original matrix notation (e.g. $\mathbf{A}$) in Eqs. (5)–(9). It is emphasized here that EnKF 3D involves only p grid cells instead of all m grid

cells. As such, the measurement operator, model error covariance matrix, and observation error covariance matrix become $\mathbf{H}_{p \times np}^s$, $\left(\mathbf{P}_{e}^s\right)_{np \times np}$, and $\mathbf{R}_{p \times p}^s$, respectively. The EnKF was then applied and the states of the centre grid cell (only) were updated. The procedure was repeated through all grid cells. To investigate the impact of including spatial correlations of errors, the "EnKF 1D" was also considered. The EnKF 1D scheme is similar to EnKF 3D, but the spatial correlations are omitted (i.e. the off-diagonal elements of the covariance matrices $\mathbf{P}_{e}^s$ and $\mathbf{R}^s$ are set to 0).

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

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

with

$$b = \frac{\ln(2)}{1 - \cos(L/a_e)}, \tag{11}$$

where α_{j_1, j_2} is the distance on the Earth's surface between two grid cells (j_1 and j_2), and L is the correlation distance. The variogram analysis was used to derive the TWSV correlation distance (L) of PCR-GLOBWB, assuming that it is similar to the correlation distance of model errors. It was found to be approximately equal to 110 km over the Hexi Corridor. For GRACE observations, to ensure that the spurious error correlations at distances greater than the Gaussian smoothing distance, 250 km, are insignificant, the localization applied to $\mathbf{R}^s$ was based on $L = 250$ km. The localization also makes the correlations at short distances slightly weaker. As a result, the condition number of the error covariance matrix is increased. In this study, for instance, the condition number increased from $\sim 10^{14}$ to $\sim 10^2$. Thus, this matrix had a full rank after localization (see Sect. 5.2.2 for a further discussion).

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

5.2.1 Model errors

The two primary sources of considered errors in the PCR-GLOBWB model are the meteorological forcing data and the model parameters. For forcing data, the precipitation uncertainties were quantified as the rms error provided by the TRMM product (Huffman, 1997). The uncertainties of temperature and potential evapotranspiration were not provided as parts of the corresponding products, and therefore errors of 2 °C and 30 % of the nominal potential evapotranspiration value were assumed, respectively. The error levels were chosen through trial and error, mainly to allow the ensemble to
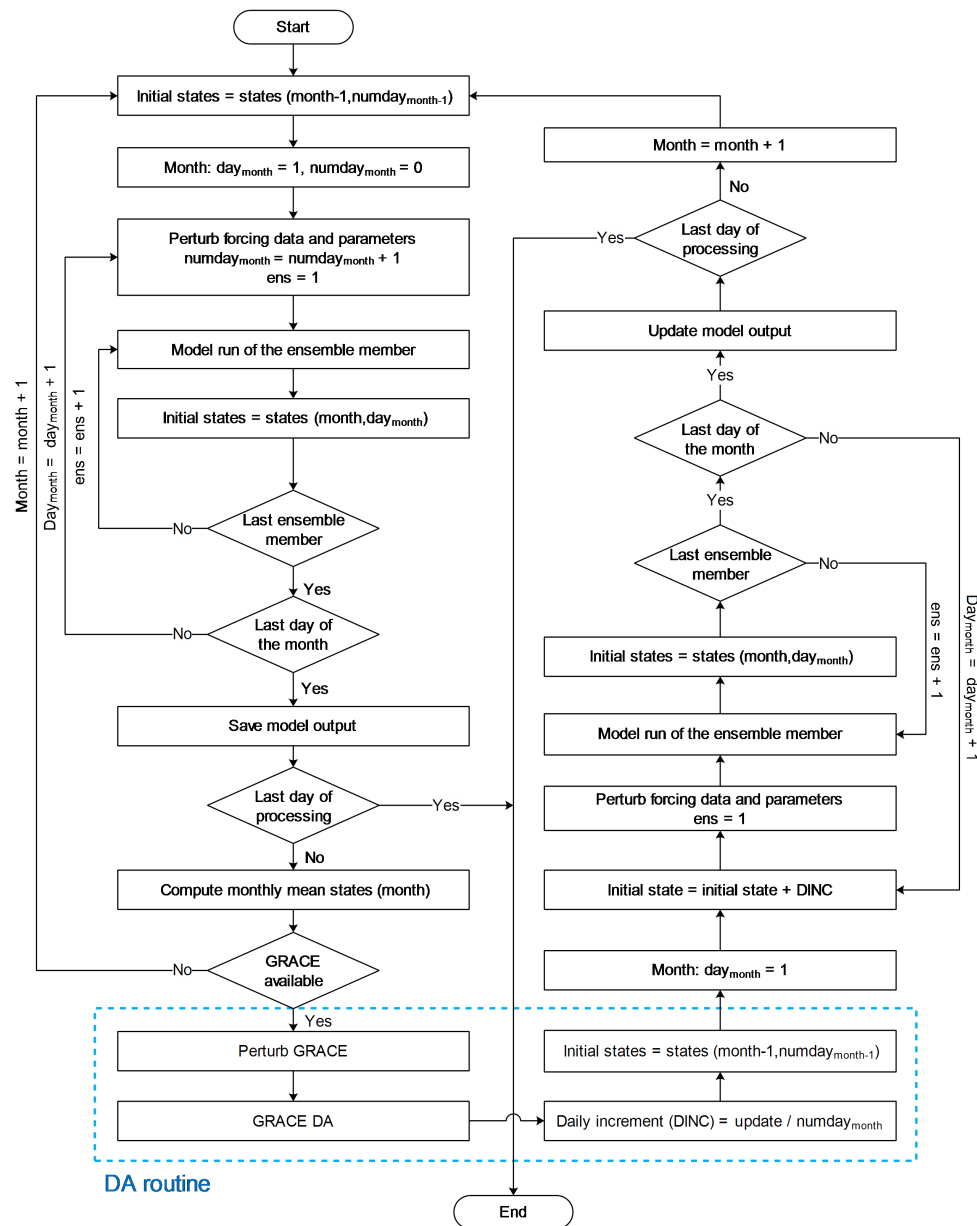
Figure 6. DA diagram representing the disaggregation of monthly averaged TWS from GRACE into the daily PCR-GLOBWB state estimates.

grow between updates. The precipitation and potential evapotranspiration were perturbed with additive lognormal noise while the temperature was perturbed with additive Gaussian noise. The forcing data uncertainties were assumed to be spatially correlated, which was accounted for using an exponential decay function. Based on a variogram analysis, the correlation distances of precipitation, temperature, and potential evapotranspiration were found to be approximately 150, 450, and 450 km, respectively.

As far as model parameters are concerned, a total of 15 TWS-related parameters (see Table 1; Sutanudjaja et al., 2011, 2014) were perturbed using additive Gaussian noise

without spatial correlations. The standard deviation of the perturbations of the parameters was set to 20 % of the range of the nominal values.

5.2.2 GRACE observation errors

Spatial correlations of GRACE observation errors were also taken into account in the DA scheme. The uncertainties in the GRACE-derived TWSV over the Hexi Corridor were computed using the monthly calibrated error variance–covariance matrix of the SHCs (Σ) provided by the CSR. Recalling the replacement of the low-degree SHCs (see Sect. 4.1), the er-

Table 1. PCR-GLOBWB model parameters related to the TWS estimate. Parameters are functions of spatial coordinates, except the degree-day factor (DDF), which is a constant.

| Parameter	Description	Unit
$K_{\text{sat,up}}$	Saturated hydraulic conductivity of the upper soil storage	m day^{-1}
$K_{\text{sat,low}}$	Saturated hydraulic conductivity of the lower soil storage	m day^{-1}
SC_{up}	Storage capacity of the upper soil	m
SC_{low}	Storage capacity of the lower soil	m
$f_g^{\min}, f_f^{\min}, f_p^{\min}, f_{np}^{\min}$	Minimum soil depth fraction of grassland (g), forest (f), paddy irrigation (p), non-paddy irrigation (np)	–
$f_g^{\max}, f_f^{\max}, f_p^{\max}, f_{np}^{\max}$	Maximum soil depth fraction of grassland (g), forest (f), paddy irrigation (p), non-paddy irrigation (np)	–
J	Groundwater recession coefficient	$1\,\text{day}^{-1}$
DDF	Degree-day factor in the snow pack	$^{\circ}\text{C m day}^{-1}$
$\text{KC}^{\min}$	Minimum crop coefficient	–
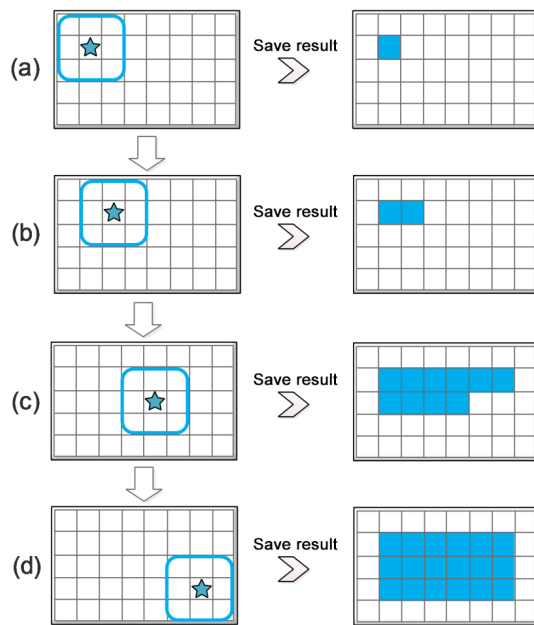
Figure 7. Demonstration of EnKF 3D scheme, accounting for the spatially correlated errors. For a centre grid cell, the state and observation matrices contain all TWS-related components of the neighbouring grid cells and the centre grid cell (left). The graphic demonstrates the case of one pixel (0.5°) correlation distance. The boundary stretches farther for larger correlation distance. The covariance matrices P_e and R are computed based on the data from these grid cells. Then, the EnKF is applied and the states of the centre grid cell are updated (right). The procedure is repeated through all grid cells.

ror (co-)variances of SHCs degree 2 were not provided by Cheng and Tapley (2004), and therefore the values obtained from the CSR were used. As for SHCs of degree 1, the error (co-)variances were not available from (Swenson et al., 2008) either and were set to 0. Note that $\boldsymbol{\Sigma}$ only reflects the error of the original GRACE data, i.e. before the GRACE processing described in Sect. 4.1 was applied. To obtain the

error variance–covariance matrix associated with the post-processed GRACE data, an ensemble of SHC noise realizations $\mathbf{Q}^c$ was first generated based on $\boldsymbol{\Sigma}$ as follows:

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

where $\mathbf{Q}^w = \left(q_1^w, q_2^w, q_3^w, \ldots, q_N^w\right)$ contains a set of white-noise realizations and has the dimension of $s \times N$, where $s = 1891$ is the number of SHCs, and $N = 100$ is the number of realizations. The matrix $\mathbf{Q}^c = \left(q_1^c, q_2^c, q_3^c, \ldots, q_N^c\right)$ has the same dimension as $\mathbf{Q}^w$ and contains an ensemble of correlated noise realizations in SHCs. Then, each noise realization (i.e. column of $\mathbf{Q}^c$) was post-processed in the same way as the GRACE data (Sect. 4.1), which resulted in $\hat{Q}^c = \left(\hat{q}_1^c, \hat{q}_2^c, \hat{q}_3^c, \ldots, \hat{q}_N^c\right)$. The post-processing included applying the destriping and Gaussian smoothing filters, as well as the signal restoration using the same number of iterations as was used in the GRACE data post-processing. The error variance–covariance matrix $\hat{\boldsymbol{\Sigma}}$ associated with the SHCs after post-processing was then computed as

$$\hat{\boldsymbol{\Sigma}} = \left[\hat{Q}^c \left(\hat{Q}^c \right)^T \right] / (N - 1). \tag{13}$$

Recalling Eq. (1), the TWSV over the Hexi Corridor can be computed as

$$\Delta \boldsymbol{\sigma} = \mathbf{YS}\boldsymbol{x}, \tag{14}$$

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

$$\mathbf{R} = \mathbf{YS}\hat{\boldsymbol{\Sigma}}(\mathbf{YS})^T. \tag{15}$$
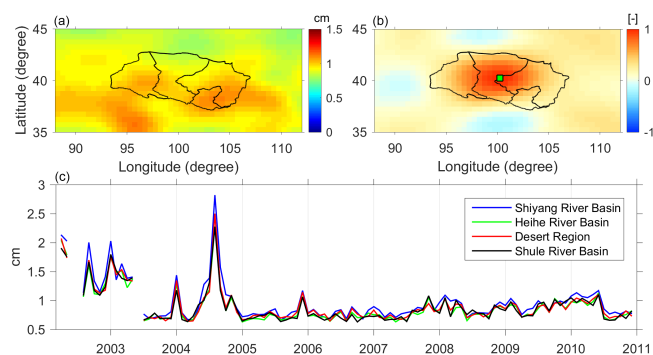
Figure 8. Some statistics of errors in GRACE-derived TWS variation over the Hexi Corridor. The standard deviation **(a)** and the correlation coefficient with respect to the green point **(b)** for a sample month, October 2002, are shown at the top. The time series of averaged standard deviation computed over four different basins are shown in the bottom plot **(c)**.

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

6 Results and discussion

The structure of this section is as follows. First, the impact of assimilation using EnKF 3D on the TWSV is considered in Sect. 6.1. Then, the impact of the EnKF 3D on the estimates of the individual stores is investigated in Sect. 6.2. The results of the EnKF 1D and EnKF 3D schemes are compared in Sect. 6.3 in terms of TWSV and the individual stores. Furthermore, the obtained results are validated against independent data in Sect. 6.4. Finally, in Sect. 6.5–6.6, the assimilation results are used together with ancillary remote sensing data to study water resources in the Hexi Corridor.

6.1 Impact of GRACE DA

To demonstrate the impact of DA, Fig. 9 shows the daily TWSV estimates over the Shiyang River basin between 1 April 2002 and 31 December 2003. Several features associated with the EnKF can be observed. Firstly, when a GRACE observation is available, the EnKF moves the estimated TWSV towards it. As a result, the estimated TWSV lies between the EnOL estimate and the GRACE observation most of the time. It is seen that GRACE-derived TWSV has a greater annual amplitude compared to the model-estimated TWSV. This can likely be attributed to the poor quality of the
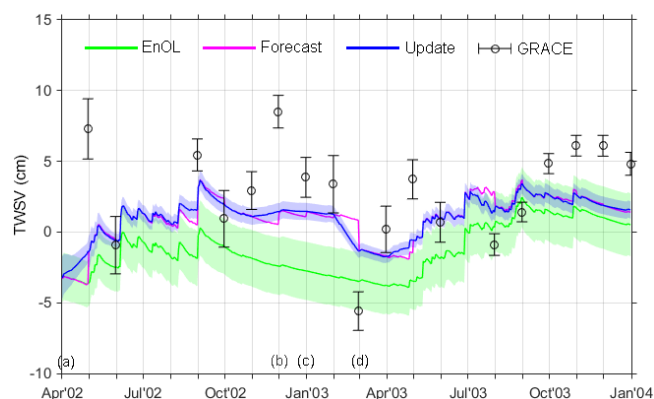

Figure 9. Daily TWS variations estimated between 1 April 2002 and 31 December 2003, averaged over Shiyang River basin. The mean value of the ensemble is given as the solid line, and the standard deviation is shown as the shaded envelope. The TWS estimates from model only (EnOL), GRACE DA forecast (EnKF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

model parameter calibration and the accuracy of the meteorological input data over the data-sparse regions. In the absence of observations, model parameters are difficult to determine and only the best available knowledge (or guess) is generally used, leading to inaccurate model state estimates. Updating the water storage estimates using GRACE DA showed a clear improvement in this case. Secondly, the standard deviation across the EnKF ensemble of TWSV values is smaller than that of the EnOL and smaller than the GRACE observation error. Thirdly, in the first month (April 2002) the TWSV estimates of the EnOL and EnKF were similar at the forecast step (as the initial states were the same, see point a in Fig. 9), but became different when the daily increment was applied to the EnKF. Finally, discontinuities in the time series before the update were observed at the end of the month, e.g. in November and December 2002 (points b and c), and February 2003 (point d). Applying the daily increment (see Sect. 5.3) served as a smoother, and these stepwise changes were reduced.

Similar features were also seen in the EnKF 1D TWSV estimates (not shown).

6.2 Impact of GRACE DA on individual stores

The monthly averaged values of the TWSV and individual stores in each of the four basins are presented in Fig. 10. Overall, TWSV estimates over the Hexi Corridor mostly reflect SMSV and GWSV components, while snow water storage variation (SNSV) and surface water storage variation (SFWV) are minor contributors, constituting less than 5 % in most basins. Clear seasonal variations in TWSV were seen in all basins for GRACE, EnOL, and GRACE DA (both EnKF
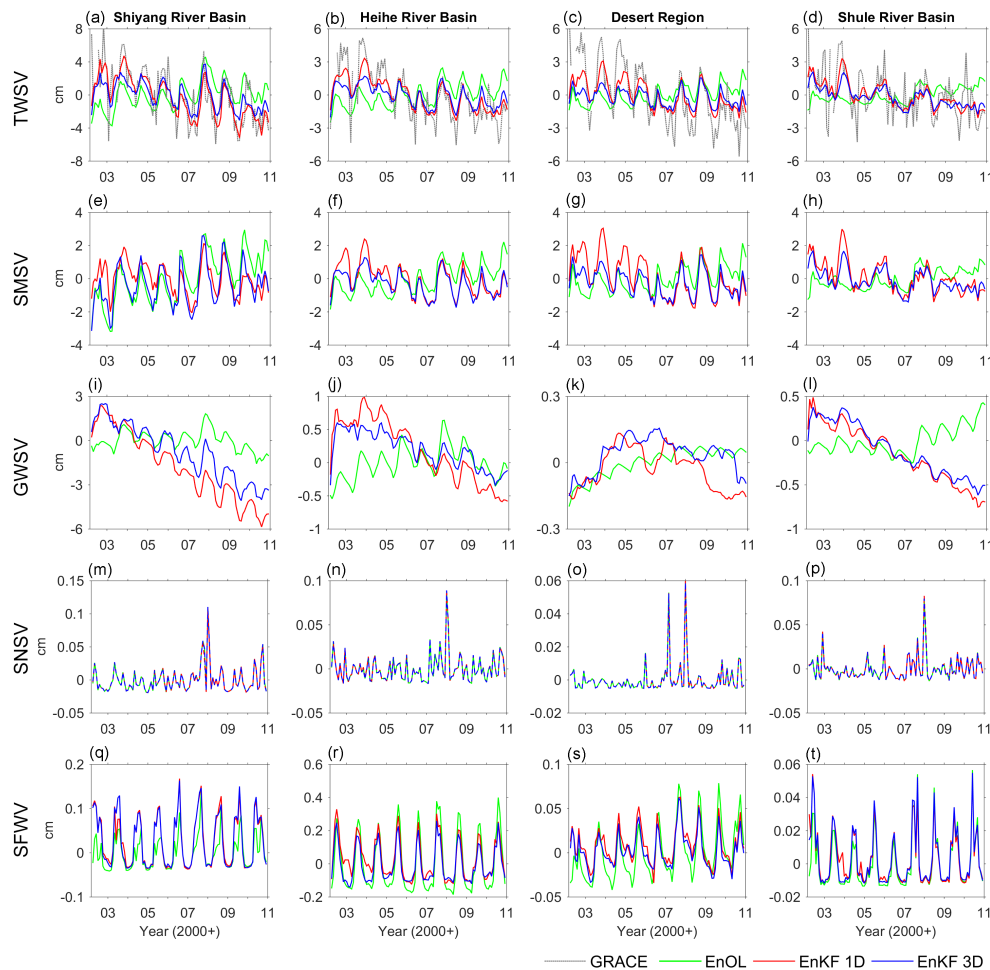
Figure 10. Monthly TWSV, SMSV, GWSV, snow water storage variation (SNSV), and surface water storage variation (SFWV) estimated between April 2002 and December 2010 from the EnOL, EnKF 1D, EnKF 3D, and GRACE observations over four basins.

1D and EnKF 3D) time series (Fig. 10a, b, c, d). As observed in Fig. 10, the GRACE DA estimated TWSVs are generally between the GRACE observations and the EnOL estimates. As a result of assimilating GRACE data, both the EnKF 1D and EnKF 3D added water to all basins between 2002 and 2005 and reduced it from the basins between 2006 and 2010. This is also seen in the time series of SMSV (Fig. 12e, f, g, h) and GWSV (Fig. 12i, j, k, l). Additionally, the annual amplitudes and phases of GRACE DA estimated TWSV were also found mostly in between the values computed from the GRACE observations and the EnOL results (see Table 2). In particular, the GRACE-DA estimated that TWSV's phase was always closer to the GRACE observation. The phase shifts of approximately 1 month were seen in both GRACE DA estimated TWS and GRACE observations compared to the EnOL results. Similar phase differences of approximately 1 month were also observed in SMSV and GWSV components.

Differences in the long-term trends were also detected between the TWSV estimates from the model alone (EnOL) and the GRACE DA. The GRACE DA results showed decreasing TWSV trends similarly to the GRACE data, while the EnOL showed increasing trends (Fig. 10a, b, c, d; see also Table 7). This change in the TWSV trend was clearly a result of assimilating GRACE observations. The negative trends were also observed after DA in the GWSV component in most basins (Fig. 10i, j, l). This indicates the potential of GRACE DA in adjusting GWSV. In this way, one can reveal continued groundwater consumption to support local agricultural activities (Li et al., 2013). Unlike over other basins, the negative trend of GWSV estimates was not clearly present over the desert region (Fig. 10k). This could be due to the small amplitude of the groundwater variation of this region (see also below), and most of the update took place in the SM component. As a result, a relatively large negative trend was seen in SMSV rather than GWSV after GRACE DA (see

Table 2. TWSV, SMSV, and GWSV estimated annual amplitude (A, cm) and phase (P, month) in four different basins computed between April 2002 and December 2010. Areally averaged values for the entire Hexi Corridor are also given.

			Shiyang	Heihe	Desert	Shule	Areal average
TWSV	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
SMSV	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
GWSV	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
	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

also Table 7). Further discussions on the trends are given in Sect. 6.4.

The impact of GRACE DA on different stores was influenced by both the model parameters and the forcing data. The four basins have similar soil water storage capacities (see Table 3), which indicates that the basins can store similar amounts of soil water and generate similar amounts of groundwater recharge under the same rainfall conditions. However, the four basins received different amounts of rainfall, which resulted in different SMSV and GWSV estimates. For example, the Shiyang River basin received the greatest amount of rainfall ($\sim$ twice of Heihe River basin), which led to the greatest amount of the SMSV estimate ($\sim$ 1 cm annual amplitude). Such a large amount was also sufficient to percolate into the groundwater layer, resulting in GWSV of $\sim$ 0.7 cm (see Fig. 10i and Table 2). In contrast, the desert region received approximately 3 times less rainfall, which led to a somewhat smaller amount of SMSV ($\sim$ 0.7 cm annual amplitude) and a much smaller amount of GWSV ($\sim$ 0.2 cm; see Fig. 10g, k). As the uncertainty of the water storage variation is associated with the signal amplitude, the greater (smaller) water storage variation leads to greater (smaller) uncertainty, resulting in a greater (smaller) update from GRACE DA. As such, a greater update (in particular, in GWSV) is seen over the Shiyang River basin, as compared to other basins.

Snow estimates (SWE plus SFW) were very small (less than 0.2 cm) over the Hexi Corridor and therefore were only slightly updated by GRACE DA. Note that the large variabil-

ity in the amount of snow seen as the sharp peaks (e.g. in January 2008) was caused by the precipitation and temperature variability. In January 2008, the precipitation records were 159 % higher than the January average value while the temperature was 2–3 °C lower. Such a condition resulted in a large amount of snow. Finally, GRACE DA influences the surface water, but the amplitude is still lower than that of the GRACE uncertainties. Validation of the surface water estimates in terms of river streamflow is given in Sect. 6.4.2.

6.3 Impact of taking spatial correlations of errors into account

The impact of accounting for the error correlations was clearly seen in the TWSV estimates (Fig. 10a, b, c, d). When the error correlations were ignored (EnKF 1D), the TWSV estimate received a larger update from GRACE, particularly between 2002 and 2005. Hence, the estimate was drawn significantly closer to the observation. The presence of error correlations effectively reduces the amount of information in the GRACE data since spatial averaging of such data mitigates noise to a much lesser extent than averaging of data with uncorrelated errors. Therefore, the impact of GRACE data in the EnKF 3D case is reduced. As such, the EnKF 3D estimated TWSV was always between the EnOL and EnKF 1D results. Validating against the in situ groundwater and streamflow data will quantitatively reveal the performance of each approach (Sect. 6.4).

Table 3. Averaged values and standard deviations of precipitation and model parameters for four different basins.

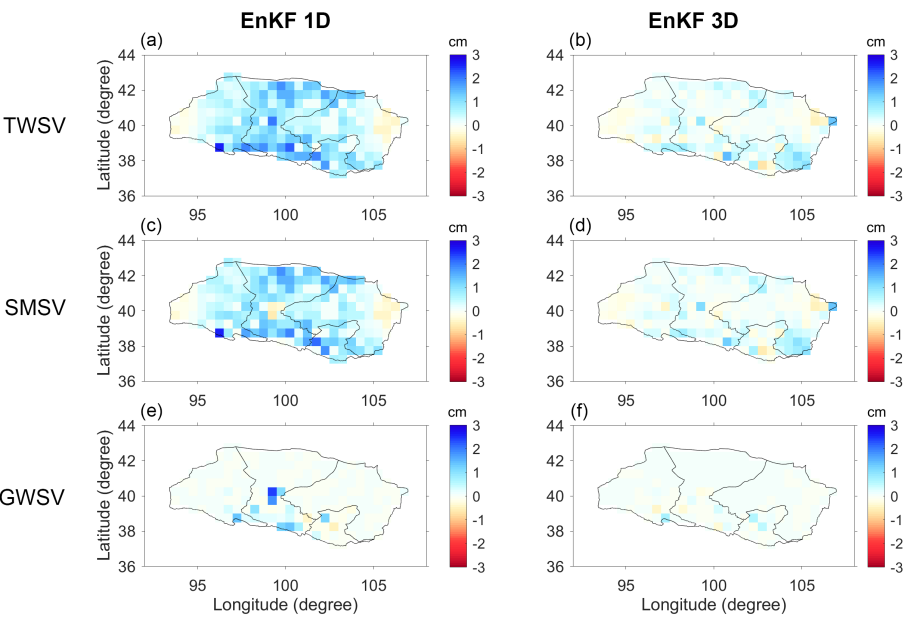|	Shiyang	Heihe	Desert	Shule
Precipitation (mm month^{-1})	21 ± 12	13 ± 12	11 ± 2	8 ± 6
SC$_{,up}$ (m)	0.08 ± 0.02	0.09 ± 0.02	0.09 ± 0.01	0.08 ± 0.01
SC$_{,low}$ (m)	0.33 ± 0.08	0.37 ± 0.07	0.35 ± 0.04	0.33 ± 0.08

Figure 11. TWSV, SMSV, and GWSV updates of October 2002 without the correlation error applied (EnKF 1D) and with the correlation error applied (EnKF 3D).

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

It is also worth discussing the impact of GRACE DA on the spatial pattern of the water storage estimates. To demonstrate this, the update term ($\Delta \mathbf{A}$ in Eq. 6) of October 2002 from EnKF 1D and 3D cases is shown in Fig. 11. Only TWSV, SMSV, and GWSV are presented, since other components (snow, surface water, and interception) are small. As discussed above, EnKF 3D shows a smaller update in all components. Due to a greater amplitude of GRACE-derived TWSV over northern and southern parts of the region (see Fig. 4), the update is mostly seen there. Almost all of the update is limited to the soil moisture layer. Higher precipitation is generally observed over the southern part, which leads to higher groundwater recharge (and GWSV) over that region. As such, a GWSV update is clearly seen over the southern part of the region.

6.4 Validation against independent data

6.4.1 Validation of groundwater estimates against well data

The GWSs estimated from GRACE DA were validated against the well measurements at five locations shown in Fig. 1c. Yang et al. (2001) showed that the specific yield values obtained from the field measurements over the Shiyang River basin were between 0.01 and 0.3. Although the measurements were not collected at the well stations used in this study, the values obtained can be used as a guidance of the
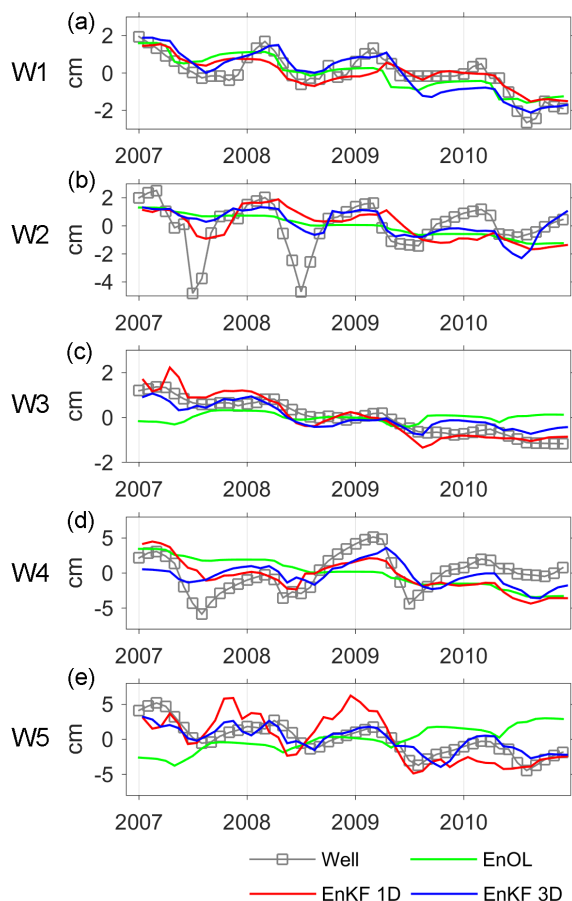
Figure 12. Monthly GWS variation estimates from the in situ well measurements, as well as EnOL, EnKF 1D, and EnKF 3D results, between January 2007 and December 2010 at five groundwater well locations. The chosen period is based on the availability of the well data.

specific yield of the Shiyang River basin. In this study, the head measurements were converted to storage units with the approach described in Sect. 4.3.1. The bias term in Eq. (2) was found to be very close to 0, as the variation (mean removed) was used in the regression analysis. The estimated scale factors were 0.23, 0.04, 0.24, 0.25, and 0.32 at W1–W5, respectively, which is in line with the values obtained from the field measurement.

The GWSV estimates at each well location are shown in Fig. 12. Compared to the EnOL results, GRACE DA results were visually closer to the well measurements at all five locations. The EnKF 1D and EnKF 3D showed a noticeable difference at each location. The updated GWSV estimates were evaluated in terms of the correlation coefficient, RMSD, and long-term trend (Tables 4, 5). Overall, the EnOL resulted in relatively poor correlation coefficients at most stations (except station W1), with the average value of only 0.06. Clear improvements were seen after GRACE DA was applied. The average correlation coefficient increased to approximately

0.6–0.7. Although the EnKF 1D introduced a greater update than the EnKF 3D, it only showed higher correlation coefficients at stations W1 and W3. Applying the EnKF 3D led to correlation coefficients greater than 0.45 in all stations, and on average it improved the correlation coefficient by approximately 0.1 over EnKF 1D. In terms of RMSD, applying GRACE DA reduced the difference by approximately 15–25 % compared to the EnOL. Compared to EnKF 1D, the EnKF 3D significantly improved the RMSD in most stations. The EnKF 1D only performed better than EnKF 3D at station W1, where it reduced the RMSD by approximately 16 % compared to the 8 % reduction by the EnKF 3D. The noticeably low GWSV observed by the well data at station W2 in the summers of 2007 and 2008 (Fig. 12b) was probably caused by significant groundwater abstraction. These local features could not be reproduced by the model and GRACE observations due to a limited spatial resolution. As a result, neither of the EnKF algorithms could improve the GWSV estimates at the W2 location during those periods.

The long-term trend estimated between 2007 and 2010 was also used to evaluate the impact of GRACE DA and the effect of taking the error correlations into account (Table 5). The EnOL trend estimates were considered poor as they showed the largest RMSD with respect to the in situ data. In fact, they were the least consistent with the in situ estimates at each individual station. Similar to the results in terms of correlation coefficient and RMSD (see Table 4), the EnKF 3D led to the largest improvement in the trend estimates (RMSD was 0.54 compared to 0.93 after EnKF 1D). However, while the EnKF 3D showed closer long-term trends to the in situ measurements at stations W2, W4, and W5, the EnKF 1D produced better estimates at stations W1 and W3.

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

6.4.2 Validation of streamflow estimates against river gauge data

The streamflow estimates were validated against the river gauge measurements at locations G1 and G2 (Fig. 1c). Results are shown in Fig. 13 and Table 6. Only modest improvements in the streamflow estimates were observed in terms of the correlation coefficient, NS coefficient, and RMSD. This behaviour is similar to what was observed previously for the Rhine River basin, when a different hydrology model and input data were used (Tangdamrongsub et al., 2015). Figure 13 shows that taking error correlations into account had little impact; i.e. similar streamflow estimates were seen for EnKF 1D and 3D results. At location G1 (Fig. 13a), GRACE DA added more water to the stream channel between 2002 and 2006 and reduced it between 2008 and 2010. This behaviour is consistent with the TWSV estimates discussed in Sect. 6.2.

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

		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
RMSDs (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

Table 5. Long-term trends and standard deviations of the in situ data and the DA estimates. The RMSDs between the in situ data and the DA trend estimates are also provided.
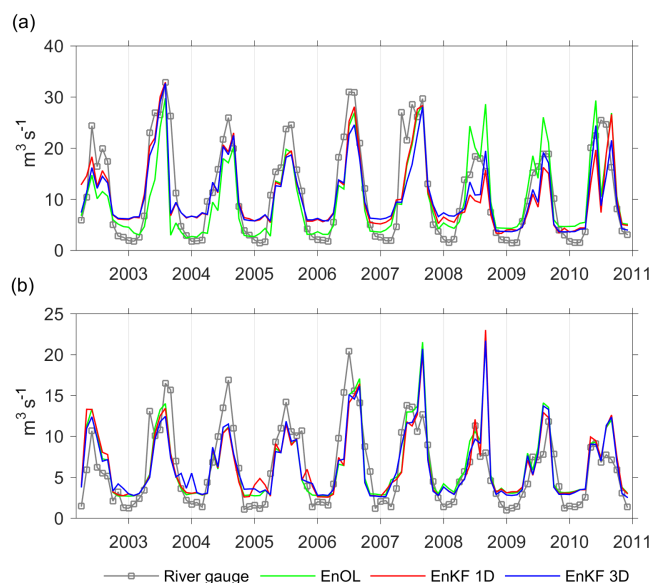
	W1	W2	W3	W4	W5	RMSD
In situ	−0.49 ± 0.03	0.01 ± 0.06	−0.60 ± 0.004	0.56 ± 0.12	−1.40 ± 0.03	0
EnOL	−0.57 ± 0.01	−0.64 ± 0.002	−0.01 ± 0.01	−1.69 ± 0.01	1.29 ± 0.02	1.62
EnKF 1D	−0.52 ± 0.02	−0.58 ± 0.04	−0.74 ± 0.02	−1.33 ± 0.08	−1.99 ± 0.13	0.93
EnKF 3D	−0.83 ± 0.02	−0.51 ± 0.03	−0.38 ± 0.01	−0.44 ± 0.08	−1.18 ± 0.06	0.54

Table 6. Statistical values of the streamflow computed from the river stream gauge measurement and GRACE DA estimates between April 2002 and December 2010. The average values are calculated by averaging the estimated statistical values from both gauge 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
RMSDs (cm)	EnOL	5.49	3.09	4.29
	EnKF 1D	5.18	3.08	4.14
	EnKF 3D	5.23	3.04	4.14

Figure 13. Monthly streamflow estimates from the in situ river gauge measurements, as well as EnOL, EnKF 1D, and EnKF 3D results, between April 2002 and December 2010 at two river gauge locations, G1 **(a)** and G2 **(b)**.

GRACE DA increased the correlation coefficient from 0.82 to 0.84, increased the NS coefficient from 0.65 to 0.69, and reduced the RMSD by approximately 5 % (Table 6). A lesser improvement was observed at G2.

Comparing to the gauge measurements, both the EnOL and GRACE DA overestimated the streamflow in September 2007 and September 2008 at G2. The sudden surge in the estimated streamflow resulted from heavy rainfall recorded by precipitation data while the soil was, according to the model, already saturated (Fig. 14). For example, in September 2007, the second highest amount of SM storage in the record (~ 19.5 cm) was obtained when the third largest

amount of rainfall (~ 90 mm month^{-1}) was observed. Similarly, in September 2008, large SM storage (~ 20 cm) and the heaviest rainfall (~ 100 mm day^{-1}) forced PCR-GLOBWB to generate a large amount of streamflow. In both cases, the modelled streamflow significantly exceeded the actual one observed at G2. Inaccurate precipitation data and model calibration likely led to these discrepancies. GRACE DA was un-

Table 7. TWSV, SMSV, GWSV, and precipitation estimated long-term trends in four different basins computed between April 2002 and December 2010. Areally averaged values for the entire Hexi Corridor are also given.

		Shiyang	Heihe	Desert	Shule	Areal average
TWSV (cm yr^{-1})	GRACE	-0.73 ± 0.04	-0.64 ± 0.03	-0.72 ± 0.03	-0.34 ± 0.04	-0.59 ± 0.03
	EnOL	0.30 ± 0.15	0.24 ± 0.09	0.20 ± 0.04	0.18 ± 0.06	0.22 ± 0.07
	EnKF 1D	-0.72 ± 0.08	-0.41 ± 0.04	-0.33 ± 0.05	-0.34 ± 0.04	-0.39 ± 0.07
	EnKF 3D	-0.36 ± 0.02	-0.21 ± 0.02	-0.11 ± 0.03	-0.25 ± 0.03	-0.20 ± 0.03
SMSV (cm yr^{-1})	EnOL	0.38 ± 0.05	0.21 ± 0.02	0.17 ± 0.03	0.14 ± 0.02	0.19 ± 0.02
	EnKF 1D	-0.11 ± 0.03	-0.20 ± 0.01	-0.29 ± 0.04	-0.22 ± 0.04	-0.23 ± 0.03
	EnKF 3D	0.10 ± 0.03	-0.12 ± 0.01	-0.12 ± 0.02	-0.14 ± 0.01	-0.11 ± 0.004
GWSV (cm yr^{-1})	EnOL	-0.08 ± 0.12	0.03 ± 0.07	0.02 ± 0.007	0.04 ± 0.02	0.02 ± 0.04
	EnKF 1D	-0.61 ± 0.01	-0.16 ± 0.004	-0.01 ± 0.005	-0.12 ± 0.02	-0.16 ± 0.02
	EnKF 3D	-0.39 ± 0.01	-0.09 ± 0.003	0.01 ± 0.004	-0.11 ± 0.001	-0.11 ± 0.002
Precipitation ((cm month^{-1}) yr^{-1})		0.04 ± 0.01	0.04 ± 0.01	0.05 ± 0.01	0.02 ± 0.01	0.04 ± 0.01
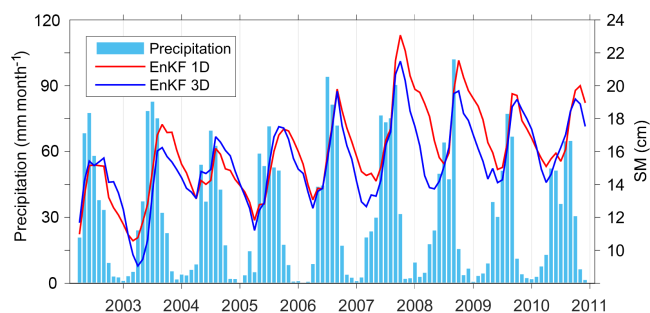
Figure 14. Monthly total precipitation (mm month^{-1}) and SM storage estimates (cm) from EnKF 1D and EnKF 3D results at the river gauge G2 location.

able to reduce these spurious peaks due to the limited spatial (~ 250 km) and temporal (1 month) resolution of GRACE data.

6.5 Declining water storages in the Hexi Corridor

The water resources situation over the Hexi Corridor was assessed using long-term trends estimated from the 9-year EnKF 3D results. This DA variant is primarily discussed here, as it provided better agreement with well observations than the EnKF 1D (see Sect. 6.4.1). For completeness, however, the values estimated from GRACE, EnOL, EnKF 1D, and precipitation are also provided. The trends in the TWSV, SMSV, and GWSV for the four basins, as well as the areally averaged values for the entire Hexi Corridor, are given in Table 7. The average EnKF 3D trends are all negative: approximately -0.2, -0.1, and -0.1 cm yr^{-1} for TWSV, SMSV, and GWSV, respectively. This reduction in the water storages is observed despite the increased amount of rainfall, which shows a positive trend of about 0.4 (mm month^{-1}) yr^{-1}. The

water storage reductions can likely be attributed to the extraction of groundwater to meet irrigation demands. In Sect. 6.6, it will be shown that groundwater extractions are essential for that purpose in the Hexi Corridor.

Focusing on individual river basins provides additional insight into the water storage issue, as the influence of the large desert area is removed. The water storage losses in the individual basins outside the desert are even more pronounced, particularly in the Shiyang River basin. This basin had the greatest TWS loss (approximately 0.4 cm yr^{-1}), which was entirely caused by the reduction of GWS. This can be explained by groundwater abstraction to meet the irrigation demand in the region. The Heihe and Shule River basins also experienced a TWS loss of ~ 0.2 cm yr^{-1}, which came from a reduction of both soil moisture and groundwater storages. Again, the negative GWS trend was likely caused by significant pumping of groundwater to maintain crop production. This is consistent with the extreme water stress over the Heihe River basin between 2001 and 2010, which was documented in Table 11.7 of the study by Chen et al. (2014). In the desert region, in contrast to other basins, the minor decreasing TWS trend of -0.1 cm yr^{-1} was dominated by loss of SM storage. This was likely caused by inaccurate model parameter calibration over the desert region (i.e. an SC value that is too large). Separation of the TWS into groundwater and soil moisture store was likely incorrect. As such, the annual signal in GWS is much less than in SM there. Therefore, the GRACE update was mostly attributed to the SM component, so that a groundwater-pumping signature (Jiao et al., 2015) was seen in the SM instead of the GWS layer.

6.6 Connection to agriculture activity

Figure 15 shows the monthly averaged groundwater head measurements at wells W1 to W5 in the Shiyang River basin (Fig. 1c). Monthly averaged precipitation and NDVI values
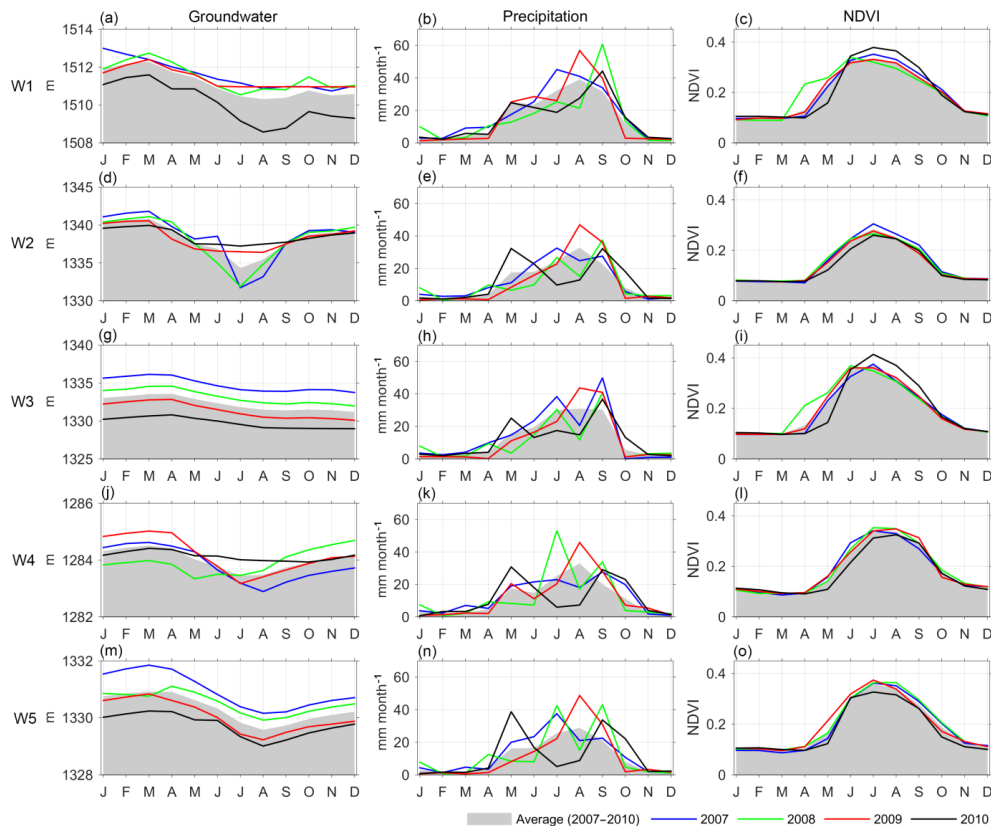
Figure 15. The monthly averaged groundwater head measurement (left), total precipitation (middle), and NDVI (right) for five groundwater well locations. Precipitation and NDVI data are reported as the average values within the circular areas of the 10 km radius. The long-term average values between January 2007 and December 2010 are shown in the grey shading, and the values in 2007, 2008, 2009, and 2010 are shown as blue, green, red, and black lines, respectively. The period is chosen based on the availability of the well data.

are shown as well. Since extracted water can be used to support agriculture not only at the well location but also in the nearby area, precipitation and NDVI are reported as the average values within a circular area of the 10 km radius. These data will be used to ascertain if groundwater extractions to support agriculture might be the source of the negative GWS trends observed in Fig. 12 and Table 6. From Fig. 14, it is noticed that the growing period is approximately between May and October, where the amount of rainfall is higher than 15 mm month^{-1} and the NDVI is typically greater than 0.2. By observing well measurements, precipitation, and NDVI together, some groundwater extraction signatures can be explained by the extension of the growing period over the dry season. For example, at station W1, the groundwater in 2010 was lower than the average, showing a gradual decrease in summer (Fig. 15a). One may attribute this to the shortage of rainfall in July and August 2010, which was lower than the average (Fig. 15b). However, the NDVI value was higher than the average during summer 2010 (Fig. 15c), which implies that water from other sources than precipitation was probably used to maintain the growing period. This additional water was likely extracted from the ground, and such

an activity led to a decreased groundwater table during summer 2010. A similar explanation can be applied to station W2, where low groundwater head, low rainfall, and high NDVI were observed in summer 2007 and summer 2008 (Fig. 15d, e, f). At station W3, the behaviour is similar to station W1: the extension of the growing period was observed in summer 2010, where the GWS and precipitation were lower than the average, while NDVI was significantly higher (Fig. 15g, h, i). Groundwater pumping signatures were not present at stations W4 and W5.

7   Conclusions

This study was focused on the estimation of water resources dynamics in the Hexi Corridor by assimilating GRACE-derived TWSV into the PCR-GLOBWB hydrological model. Validating against well data showed that DA led to noticeable improvement in the state estimates in terms of correlation, RMSD, and long-term trend. Furthermore, GRACE DA estimates revealed the reduction of water storages between 2002 and 2010. The Shiyang River basin – the southeastern part of the Hexi Corridor area – suffered the most from the water

loss, which was likely caused by the overuse of the groundwater for irrigation. Due to inaccurate groundwater abstraction information, PCR-GLOBWB alone could not properly capture the downward trend of water storages. This highlights the value of the GRACE DA in this situation. It should be emphasized that GRACE does not fix a technical problem of the hydrological model, but rather it provides information which is not available otherwise. Note that, in principle, the model may predict any long-term behaviour of water storage, but that information should be brought in "by hand" (e.g. via the groundwater abstraction parameter). As soon as that information is not available, reliable long-term predictions on the basis of hydrological modelling alone are conceptually impossible. GRACE DA acts as a provider of a missing puzzle piece here. Of course, the performance of GRACE DA needs to be further investigated in other geographical locations and with different hydrological models to confirm its benefits.

A substantial decrease in the water storage in the Hexi Corridor between 2002 and 2010, particularly over the Shiyang River basin, took place in spite of the increased precipitation. The amount of water from rainfall was likely insufficient to support irrigation water requirements. Irrigation water demands increased significantly to maintain the crop production and, as a result, the region was under extreme water stress. Water consumption from all available sources was essential for bridging the deficit, including a sizeable amount of groundwater extraction. This study illustrates how ground observations and remote sensing data may reveal the connection between groundwater pumping and agricultural activity.

The conversion approach between the groundwater head measurement and groundwater storage is proven feasible over the Shiyang River Basin. The scale factor estimates produced with this approach are consistent with the specific yield estimated from the field observations. However, it is noted here that the results of the conducted validation might be over-optimistic, since the well data processed with the adopted conversion procedure are not fully independent of the assimilated GRACE data. The specific yield from the field observation must be used when available.

Furthermore, we demonstrate how the error covariance matrix **R** of GRACE-derived TWSV can be obtained from the error covariance matrix of GRACE SHCs (which is currently provided together with the SHCs themselves). This study shows that it is necessary to use the **R** matrix in order to properly take into account the error correlations in the DA scheme. To come to that conclusion, we considered two variants of the error variance–covariance matrix in the data assimilation: excluding and including error correlations. Validating against well data showed that ignoring error correlations in DA tended to over-fit results to the observations, and in many cases led to less-accurate state estimates. This finding is in agreement with the recommendation in Schumacher et al. (2016). We explain this finding by the fact that GRACE errors at the neighbouring $0.5° \times 0.5°$ grid cells are highly correlated. As such, the simultaneous consideration of GRACE data at multiple neighbouring cells does not reduce data noise, as it would be the case if noise were white. In other words, the white-noise assumption may severely overestimate the information content of GRACE data. We recognize that the derivation of GRACE-derived TWSV error variance–covariance matrices is very computationally demanding. Still, we believe that this is a reasonable price to pay as deriving the error variance–covariance matrix from the full (and only full) error covariance matrix noticeably improves the results of GRACE data assimilation.

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

Finally, the usage of improved gravity solutions to be available after the launch of the GRACE Follow-On mission (Flechtner et al., 2014) will probably further increase the accuracy of the GRACE DA estimates.

Competing interests. The authors declare that they have no conflict of interest.

References

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

Carlson, T. N. and Ripley, D. A.: On the relation between NDVI,
fractional vegetation cover, and leaf area index, Remote Sens.
Environ., 62, 241–252, doi:10.1016/S0034-4257(97)00104-1,
1997.

Chen, J., Li, J., Zhang, Z., and Ni, S.: Long-term ground-
water variations in Northwest India from satellite grav-
ity measurements, Global Planet. Change, 116, 130–138,
doi:10.1016/j.gloplacha.2014.02.007, 2014.

Cheng, M. and Tapley, B.: Variations in the Earth's oblateness
during the past 28 years, J. Geophys. Res., 109, B09402,
doi:10.1029/2004JB003028, 2004.

Cui, Y. and Shao, J.: The Role of Ground Water in Arid/Semiarid
Ecosystems, Northwest China, Groundwater, 43, 471–477,
doi:10.1111/j.1745-6584.2005.0063.x, 2005.

Dahle, C., Flechtner, F., Gruber, C., König, D., König, R., Micha-
lak, G., and Neumayer, K.-H.: GFZ RL05: An Improved Time-
Series of Monthly GRACE Gravity Field Solutions, in: Observa-
tion of the System Earth from Space – CHAMP, GRACE, GOCE
and future missions, edited by: Flechtner, F., Sneeuw, N., and
Schuh, W.-D., GEOTECHNOLOGIEN Science Report, 20, Ad-
vanced Technologies in Earth Sciences, Berlin, Springer, 29–39,
doi:10.1007/978-3-642-32135-1_4, 2014.

Dahlgren, S. and Possling, B.: Soil Water Modelling In
Arid/Semiarid Regions of Northern China Using Land Informa-
tion System (LIS): A Minor Field Study in Shiyang River Basin,
Dept. of Water Resources Engineering, Lund University, ISRN
LUTVDG/TVVR-07/5019, ISSN: 1101-9824, 2007.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli,
P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G.,
Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bid-
lot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer,
A. J., Haiberger, L., Healy, S. B., Hersbach, H., Hólm, E. V.,
Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally,
A. P., Monge-Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey,
C., de Rosnay, P., Tavolato, C., Thépaut, J. N., and Vitart, F.: The
ERA-Interim reanalysis: configuration and performance of the
data assimilation system, Q. J. Roy. Meteor. Soc., 137, 553–597,
doi:10.1002/qj.828, 2011.

De Lannoy, G. J. M., Reichle, R. H., Houser, P. R., Ar-
senault, K. R., Verhoest, N. E. C., and Pauwels, R. N.:
Satellite-Scale Snow Water Equivalent Assimilation into a High-
Resolution Land Surface Model, J. Hydrometeor., 11, 352–369,
doi:10.1175/2009JHM1192.1, 2009.

Döll, P., Schmied, H. M., Schuh, C., Portmann, F. T., and
Eicker, A.: Global-scale assessment of groundwater deple-

tion and related groundwater abstractions: Combining hydro-
logical modeling with information from well observations
and GRACE satellites, Water Resour. Res., 50, 5698–5720,
doi:10.1002/2014WR015595, 2014.

Dong, J., Steele-Dunne, S. C., Judge, J., and van de Giesen, N.:
A particle batch smoother for soil moisture estimation using
soil temperature observations, Adv. Water Resour., 83, 111–122,
doi:10.1016/j.advwatres.2015.05.017, 2015.

Du, T., Kang, S., Zhang, Z., and Zhang, J.: China's food secu-
rity is threatened by the unsustainable use of water resources in
North and Northwest China, Food and Energy Security, 3, 7–18,
doi:10.1002/fes3.40, 2014.

Eicker, A., Schumacher, M., Kusche, J., Döll, P., and Müller
Schmied, H.: Calibration data assimilation approach for inte-
grating GRACE data into the WaterGAP Global Hydrology
Model (WGHM) using an Ensemble Kalman Filter: First Re-
sults, Surv. Geophys., 35, 1285–1309, doi:10.1007/s10712-014-
9309-8, 2014.

Evensen, G.: The ensemble Kalman filter: Theoretical formula-
tion and practical implementation, Ocean Dyn., 53, 343–367,
doi:10.1007/S10236-003-0036-9, 2003.

Flechtner, F., Morton, P., Watkins, M., and Webb, F.: Status of the
GRACE follow-on mission, in IAG symposium gravity, geoid,
and height systems, 141, Venice, Italy, Springer, 117–121, 2014.

Forman, B. A., Reichle, R. H., and Rodell, M.: Assimilation of ter-
restrial water storage from GRACE in a snow-domained basin,
Water Resour. Res., 48, W01507, doi:10.1029/2011WR011239,
2012.

Geng, G. T. and Wardlaw, R.: Application of Multi-Criterion De-
cision Making Analysis to Integrated Water Resources Manage-
ment, Water Resour. Manage., 27, 3191–3207, 2013.

Girotto, M., De Lannoy, G. J. M., Reichle, R. H., and Rodell, M.:
Assimilation of gridded terrestrial water storage observations
from GRACE into a land surface model, Water Resour. Res., 52,
4164–4183, doi:10.1002/2015WR018417, 2016.

Gong, D. Y., Shi, P. J., and Wang, J. A.: Daily precipitation changes
in the semi-arid region over northern China, J. Arid. Environ.,
59, 771–784, doi:10.1016/j.jaridenv.2004.02.006, 2004.

Hamill, T. M., Whitaker, J. S., and Snyder, C.: Distance-Dependent
Filtering of Background Error Covariance Estimates in an En-
semble Kalman Filter, Mon. Weather Rev., 129, 2776–2790,
2001.

Houborg, R., Rodell, M., Li, B., Reichle, R., and Zaitchik,
B. F.: Drought indicators based on model-assimilated Grav-
ity Recovery and Climate Experiment (GRACE) terrestrial wa-
ter storage observations, Water Resour. Res., 48, W07525,
doi:10.1029/2011WR011291, 2012.

Hu, Z. D.: Study on Study on Evolution Pattern of Water Resources,
Ecology and Oasis Migration in the Hexi Corridor, Doctoral Dis-
sertation, Tsinghua University, Beijing, China, 2015.

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., and Fer-
reira, L. G.: Overview of the radiometric and biophysical perfor-
mance of the MODIS vegetation indices, Remote Sens. Environ.,
83, 195–213, doi:10.1016/S0034-4257(02)00096-2, 2002.

Huffman, G. J.: Estimates of Root-Mean-Square Random Error for
Finite Samples of Estimated Precipitation, J. Appl. Meteor., 36,
1191–1201, 1997.

Huffman, G. J., Adler, R. F., Bolvin, D. T., Gu, G., Nelkin, E.
J., Bowman, K. P., Hong, Y., Stocker, E. F., and Wolf, D. B.:

The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales, J. Hydrometeor., 8, 38–55, doi:10.1175/JHM560.1, 2007.

Huo, Z., Feng, S., Kang, S., Dai, X., Li, W., and Chen, S.: The Response of Water-Land Environment to Human Activities in Arid Minqin Oasis, Northwest China, Arid Land Res. Manag., 21, 21–36, 2007.

Jekeli, C.: Alternative methods to smooth the Earth's gravity field, Rep., 327, Dept. of Geod. Sci. and Surv., Ohio State Univ., Columbus, 1981.

Jiao, J. J., Zhang, X., and Wang, X.: Satellite-based estimates of groundwater depletion in the Badain Jaran Desert, China, Nature Sci. Rep., 5, 8960, doi:10.1038/srep08960, 2015.

Kang, S., Su, X., Tong, L., Shi, P., Yang, X., Abe, Y., Du, T., Shen, Q., and Zhang, J.: The impacts of human activities on the water-land environment of the Shiyang River basin, an arid region in northwest China, Hydrolog. Sci. J., 49, 413–427, 2014.

Klees, R., Liu, X., Wittwe, T., Gunter, B. C., Revtova, E. A., Tenzer, R., Ditmar, P., Winsemius, H. C., and Savenije, H. H. G.: A Comparison of Global and Regional GRACE Models for Land Hydrology, Surv. Geophys., 29, 335–359, doi:10.1007/s10712-008-9049-8, 2008.

Kummerow, C., Barnes, W., Kozu, T., Shiue, J., and Simpson, J.: The Tropical Rainfall Measuring Mission (TRMM) sensor package, J. Atmos. Ocean. Tech., 15, 809–817, 1998.

Li, B., Rodell, M., Zaitchik, B. F., Reichle, R. H., Koster, R. D., and van Dam, T. M.: Assimilation of GRACE terrestrial water storage into a land surface model: Evaluation and potential value for drought monitoring in western and central Europe, J. Hydrol., 446–447, 103–115, doi:10.1016/j.jhydrol.2012.04.035, 2012.

Li, F., Zhu, G., and Guo, C.: Shiyang River ecosystem problems and countermeasures, Agr. Sci., 4, 72–78, doi:10.4236/as.2013.42012, 2013.

Long, D., Longuevergne, L., and Scanlon, B. R.: Uncertainty in evapotranspiration from land surface modeling, remote sensing, and GRACE satellites, Water Resour. Res., 50, 1131–1151, doi:10.1002/2013WR014581, 2014.

Ma, J. Z., Wang, X. S., and Edmunds, W. M.: The characteristics of ground-water resources and their changes under the impacts of human activity in the arid Northwest China – a case study of the Shiyang River Basin, J. Arid Environ., 61, 277–295, 2005.

Mitchell, T. D. and Jones, P. D.: An improved method of constructing a database of monthly climate observations and associated high-resolution grids, Int. J. Climatol., 25, 693–712, 2005.

Niu, G.-Y., Seo, K.-W., Yang, Z.-L., Wilson, C., Su, H., Chen, J., and Rodell, M.: Retrieving snow mass from GRACE terrestrial water storage change with a land surface model, Geophys. Res. Lett., 34, L15704, doi:10.1029/2007GL030413, 2007.

Peel, M. C., Finlayson, B. L., and McMahon, T. A.: Updated world map of the Köppen-Geiger climate classification, Hydrol. Earth Syst. Sci., 11, 1633–1644, doi:10.5194/hess-11-1633-2007, 2007.

Reichle, R. H. and Koster, R. D.: Bias reduction in short records of satellite soil moisture, Geophys. Res. Lett., 31, L19501, doi:10.1029/2004GL020938, 2004.

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C. J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D., and Toll,

D.: The global land data assimilation system, B. Am. Meteorol. Soc., 85, 381–394, 2004.

Schumacher, M., Kusche, J., and Döll, P.: A Systematic Impact Assessment of GRACE Error Correlation on Data Assimilation in Hydrological Models, J. Geod., 90, 537–559, doi:10.1007/s00190-016-0892-y, 2016.

Sheffield, J., Goteti, G., and Wood, E. F.: Development of a 50-yr high-resolution global dataset of meteorological forcings for land surface modeling, J. Climate, 19, 3088–3111, 2005.

Shiklomanov, I. A.: Assessment of water resources and water availability in the world, Comprehensive assessment of the freshwater resources of the world, World Meteorological Organization and the Stockholm Environment Institute, Stockholm, Sweden, 1997.

Steinfeld, H., Gerber, P., Wassenaar, T., Castel, V., Rosales, M., and de Haan, C.: Livestocks long shadow: Environmental issues and options, FAO, Rome, Italy, ISBN-13: 978-92-5-105571-7, 2006.

Su, H., Yang, Z. L., Dickinson, R. E., Wilson, C. R., and Niu, G. Y.: Multisensor snow data assimilation at the continental scale: The value of Gravity Recovery and Climate Experiment terrestrial water storage information, J. Geophys. Res., 115, D10104, doi:10.1029/2009JD013035, 2010.

Sutanudjaja, E. H., van Beek, L. P. H., de Jong, S. M., van Geer, F. C., and Bierkens, M. F. P.: Large-scale groundwater modeling using global datasets: a test case for the Rhine-Meuse basin, Hydrol. Earth Syst. Sci., 15, 2913–2935, doi:10.5194/hess-15-2913-2011, 2011.

Sutanudjaja, E. H., van Beek, L. P. H., de Jong, S. M., van Geer, F. C., and Bierkens, M. F. P.: Calibrating a large-extent high-resolution coupled groundwater-land surface model using soil moisture and discharge data, Water Resour. Res., 50, 687–705, doi:10.1002/2013WR013807, 2014.

Sutanudjaja, E. H., van Beek, L. P. H., Drost, N., de Graaf, I. E. M., de Jong, K., Peßenteiner, S., Straatsma, M. W., Wada, Y., Wanders, N., Wisser, D., and Bierkens, M. F. P.: PCR-GLOBWB 2.0: a 5 arc-minute global hydrological and water resources model, to be submitted to Geosci. Model Dev. Discuss., in preparation, 2017.

Swenson, S. and Wahr, J.: Post-processing removal of correlated errors in GRACE data, Geophys. Res. Lett., 33, L08402, doi:10.1029/2005GL025285, 2006.

Swenson, S., Chambers, D., and Wahr, J.: Estimating geocenter variations from a combination of GRACE and ocean model output, J. Geophys. Res., 113, B08410, doi:10.1029/2007JB005338, 2008.

Tangdamrongsub, N., Steele-Dunne, S. C., Gunter, B. C., Ditmar, P. G., and Weerts, A. H.: Data assimilation of GRACE terrestrial water storage estimates into a regional hydrological model of the Rhine River basin, Hydrol. Earth Syst. Sci., 19, 2079–2100, doi:10.5194/hess-19-2079-2015, 2015.

Tangdamrongsub, N., Ditmar, P. G., Steele-Dunne, S. C., Gunter, B. C., and Sutanudjaja, E. H.: Exploring irregular flood events over Tonlé Sap basin in Cambodia using GRACE and MODIS satellite observations combined with altimetry observation and hydrological models, Remote Sens. Environ., 181, 162–173, doi:10.1016/j.rse.2016.03.030, 2016.

Tapley, B. D., Bettadpur, S., Ries, J. C., Thompson, P. F., and Watkins, M.: GRACE Measurements of Mass Variability in the Earth System, Science, 305, 503–505, 2004.

van Beek, L. P. H.: Forcing PCR-GLOBWB with CRU data, Technical Report, Department of Physical Geography, Utrecht University, Utrecht, the Netherlands, available at: http://vanbeek.geo.uu.nl/suppinfo/vanbeek2008.pdf (last access: 8 October 2014), 2008.

van Beek, L. P. H. and Bierkens, M. F. P.: The Global Hydrological Model PCR-GLOBWB: Conceptualization, Parameterization and Verification, Technical Report, Department of Physical Geography, Utrecht University, Utrecht, the Netherlands, available at: http://vanbeek.geo.uu.nl/suppinfo/vanbeekbierkens2009.pdf (last access: 7 October 2014), 2009.

van Beek, L. P. H., Wada, Y., and Bierkens, M. F. P.: Global monthly water stress: 1. Water balance and water availability, Water Resour. Res., 47, W07517, doi:10.1029/2010WR009791, 2011.

Vermote, E. F., Kotchenova, S. Y., and Ray, J. P.: MODIS surface reflectance user's guide, verion 1.3, available at: http://modis-sr.ltdri.org/guide/MOD09_UserGuide_v1_3.pdf (last access: 31 July 2015), 2011.

Wada, Y., Wisser, D., and Bierkens, M. F. P.: Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources, Earth Syst. Dynam., 5, 15–40, doi:10.5194/esd-5-15-2014, 2014.

Wahr, J., Molenaar, M., and Bryan, F.: Time variability of the Earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE, J. Geophys. Res., 103, 30205–30229, 1998.

Wahr, J., Swenson, S., and Velicogna, I.: Accuracy of GRACE mass estimates, Geophys. Res. Lett., 33, L06401, doi:10.1029/2005GL025305, 2006.

Wanders, N., Bierkens, M. F. P., Jong, S. M., Roo, A., and Karssenberg, D.: The benefits of using remotely sensed soil moisture in parameter identification of large-scale hydrological models, Water Resour. Res., 50, 6874–6891, doi:10.1002/2013WR014639, 2014.

Wang, G., Y. Ding, Y., Shen, Y., and Lai, Y.: Environmental degradation in the Hexi Corridor region of China over the last 50 years and comprehensive mitigation and rehabilitation strategies, Environ. Geol., 44, 68–77, 2003.

Weerts, A. H. and El Serafy G. Y. H.: Particle filtering and ensemble Kalman filtering for state updating with hydrological conceptual rainfall-runoff models, Water Resour. Res., 42, W09403, doi:10.1029/2005WR004093, 2006.

Yang, Y. S., Kalin, R. M., Zhang, Y., Lin, X., and Zou, L.: Multi-objective optimization for sustainable groundwater resource management in a semiarid catchment, Hydrolog. Sci. J., 46, 55–72, doi:10.1080/02626660109492800, 2001.

Zaitchik, B. F., Rodell, M., and Reichle, E. H.: Assimilation of GRACE terrestrial water storage data into a land surface model: Results for the Mississippi basin, Amer. Meteor. Soc., J. Hydrometeor., 9, 535–548, doi:10.1175/2007JHM951.1, 2008.

Zheng, H., Wang, Z. J., Hu, S. Y., and Malano, H.: Seasonal Water Allocation: Dealing with Hydrologic Variability in the Context of a Water Rights System, J. Water Res. Pl. ASCE, 139, 76–85, 2013.

Zhu, J., Winter, C. L., and Wang, Z.: Nonlinear effects of locally heterogeneous hydraulic conductivity fields on regional stream-aquifer exchanges, Hydrol. Earth Syst. Sci., 19, 4531–4545, doi:10.5194/hess-19-4531-2015, 2015.