Abstract

Worsening water storage depletion (WSD) contributes to environmental degradation, land subsidence and earthquake and could disrupt food production/security and social stability. There is need for efficient water use strategies in North China, a pivotal agrarian, industrial and political base in China with a widespread WSD. This study integrates satellite, model and field data products to investigate WSD and land subsidence in North China. In the first step, GRACE (Gravity Recovery and Climate Experiment) mass rates are used to show WSD in the region. Next, GRACE total water storage (TWS) is corrected for soil water storage (SWS) to derive groundwater storage (GWS) using GLDAS (Global Land Data Assimilation System) data products. The derived GWS is compared with GWS obtained from field-measured groundwater level to show land subsidence in the study area. Then GPS (Global Positioning System) data of relative land surface change (LSC) are used to confirm the subsidence due to WSD. A total of ∼96 near-consecutive months (January 2002 through December 2009) of datasets are used in the study. Based on GRACE mass rates, TWS depletion is $23.76 \pm 1.74 \text{ mm yr}^{-1}$ or $13.73 \pm 1.01 \text{ km}^3 \text{ yr}^{-1}$ in the 578,000 km$^2$ study area. This is ∼31 % of the slated 45 km$^3$ yr$^{-1}$ water delivery in 2050 via the South–North Water Diversion Project. Analysis of relative LSC shows subsidence of $7.29 \pm 0.35 \text{ mm yr}^{-1}$ in Beijing and $2.74 \pm 0.16 \text{ mm yr}^{-1}$ in North China. About $11.53 \%$ ($2.74 \pm 0.18 \text{ mm}$ or $1.58 \pm 0.12 \text{ km}^3$) of the TWS and $8.37 \%$ ($1.52 \pm 0.70 \text{ mm}$ or $0.88 \pm 0.03 \text{ km}^3$) of the GWS are attributed to storage reductions accompanying subsidence in the region. Although interpretations of the findings require caution due to the short temporal and large spatial coverage, the concurrence of WSD and land subsidence could have adverse implications for the study area. It is critical that the relevant stakeholders embark on resource-efficient measures to ensure water availability, food security, ecological sustainability and social stability in this pivotal region.
1 Introduction

Groundwater is a vital source of water supply for humans, agriculture, industries and high-value ecosystems (Gleeson et al., 2012; Cao et al., 2013). Current global groundwater withdrawal is estimated at 750–800 km³ yr⁻¹ (Wada et al., 2010), providing some 50 % of drinking water needs, 40 % of industrial water demand and 20 % of irrigation water supply (Konikow, 2011). Along the benefits (e.g., increased food production and socio-economic growth) are also the consequences (e.g., storage depletion, water quality deterioration, ecological degradation, land subsidence, seawater intrusion, sea level rise and water conflicts) of groundwater exploitation (Gleeson et al., 2012; Feng et al., 2013). Excessive groundwater exploitation over the long period could cause land subsidence (Galloway et al., 1998; Xue et al., 2005; Gambolati et al., 2006) or even earthquake (Bawden et al., 2001; Gonzáles et al., 2012).

Groundwater withdrawal reduces pore pressure which increases effective stress; a condition that causes aquifer compaction, land subsidence or earthquake (Chen et al., 2007; Zhang et al., 2008; Gonzáles et al., 2012). Land subsidence is the elastic/inelastic compaction of crustal materials due to changes in stress dynamics (Galloway et al., 1999; Waltham, 2002). Groundwater-unloading subsidence is the compaction of clay-silt/sand-gravel deposits (Phienwej et al., 2006; Aeschbach-Hertig and Gleeson, 2012). Aquifer recharge restores lost pore-water pressure that induces elastic rebound of aquifer materials (Waltham, 2002).

Despite its significance or even because of it, current global rates of groundwater development are unsustainable (Yang et al., 2010; Aeschbach-Hertig and Gleeson, 2012; Moiwo et al., 2013). Groundwater-controlled crustal deformations like groundwater depression cones and land subsidence have been reported in Arizona/Texas (Galloway et al., 1999), Mexico City (Adrian et al., 1999), Jakarta (Abidin et al., 2001), Ravenna (Teatini et al., 2005), Taiwan (Hu et al., 2006) and Lorca/Spain (Gonzáles et al., 2012). Other crustal deformations and the related effects are detailed by Bawden et al. (2001) for the Los Angeles Basin in USA and Rodolfo and Siringan (2006) for East Asia.

Poland et al. (1975) noted a record groundwater-controlled subsidence of > 8.5 m in San Joaquin Valley of California. There are several other reports of subsidence due to long-term groundwater withdrawals in China and elsewhere across the globe (Ireland et al., 1984; Loefgren, 1991; Zhang et al., 2008; Li et al., 2011).

Monitoring land surface change (LSC) for crustal deformation is a critical first-step for effective intervention measures (Galloway et al., 1999; Galloway and Burbey, 2011). Several techniques exist for LSC monitoring, including spirit leveling, piezometric/reservoir pressure measurement, geodimeter measurement and extensometer measurement. Also radioactive marker techniques, modeling techniques and space-borne techniques like GPS (Global Positioning System) and InSAR (Interferometry Synthetic Aperture Radar) are now available. The procedural details of several LSC monitoring techniques are discussed by Loefgren (1991), Abidin et al. (2001), Gambolati et al. (2006), Chen et al. (2007) and Liu et al. (2008).

Here, GPS data products of relative LSC are used in combination with hydrological data products from GRACE (Gravity Recovery and Climate Experiment), GLDAS (Global Land Data Assimilation System) and field-measured groundwater level and soil water storage (SWS) to characterize water storage depletion (WSD) and land subsidence in North China.

Launched in March 2002 (Swenson et al., 2006), GRACE is a twin satellite system that orbits the Earth in tandem at a separation distance and altitude of ~ 200 and ~ 450 km, respectively (Syed et al., 2008; Longuevergne et al., 2010). Changes in the satellite separation distance due to variations in gravity pull are measured by an onboard K-band microwave to the nearest nanometer. At any point in time, the satellite orbit position is tracked by GPS receivers and related with the inter-satellite distance for gravity fields (Wahr et al., 2004; Rodell et al., 2009). Then mass redistributions within the Earth (e.g., atmospheric surface and ocean bottom pressures and terrestrial water, snow/ice storage) are inferred from the time-variable gravity fields (Chambers, 2009; Longuevergne et al., 2010). After removing non-hydrologic effects, GRACE gravity field residuals closely track terrestrial water storage changes derived from field measure-
ments or model estimates at various spatial scales (Ramillien et al., 2008; Strassberg et al., 2009; Famiglietti et al., 2011; Moiwo et al., 2011).

GLDAS mission models global land surface energy states (Rodell et al., 2004), including soil water, runoff and evapotranspiration (ET). GLDAS land surface models include Mosaic, Noah, Community Land Model, Variable Infiltration Capacity and Catchment Model (Hogue et al., 2005). As the accuracy levels of GRACE and GLDAS data products are comparable (Rodell et al., 2009; Longuevergne et al., 2010), their integrative use can improve water storage analysis (Moiwo et al., 2010; Immerzeel et al., 2010).

In this study, GRACE total water storage (TWS) is corrected for groundwater storage (GWS) using GLDAS-derived SWS and compared with GWS from field-measurements to show WSD and land subsidence in North China. Then the storage-unloading subsidence is confirmed by analysis of GPS data product of relative LSC in the region. This study extends current applications of GRACE/GLDAS data products well beyond the conventional field of water storage analyses into crustal deformation analyses. This will deepen existing knowledge on the implications of long-term storage depletion for human subsistence. It will also add greater urgency to the need to develop resource-efficient strategies for a truly sustainable world.

2 Materials and method

2.1 Study area

The study area, North China, lies between longitudes 107.1–119.2° E and latitudes 34.6–41.7° N in an area of ~ 578 000 km². In addition to Beijing and Tianjin, the area variously covers six provinces (Fig. 1, top plate) and has a total population exceeding 250 × 10⁶ people. The region accounts for over 45% of China’s grain production and 15% of its gross domestic product (Cao et al., 2013; Feng et al., 2013; Tang et al., 2013).

The average annual ET (985 mm yr⁻¹) in cultivated lands is over two times the average precipitation (500 mm yr⁻¹), sustained largely by irrigation (Yang et al., 2010; Moiwo et al., 2011). The thick Quaternary alluvial deposits (> 1000 m) are poorly drained with average hydrologic conductivity of > 100 m d⁻¹ (Foster et al., 2004; Hu et al., 2010; Cao et al., 2013). Thus the several decades of high water use (especially in agriculture and industry) has pushed water exploitation far beyond precipitation recharge in the region (Kendy et al., 2007; Cao et al., 2013). The difficulties in the water balance accounting and the increasing gap between water use and recharge have caused surface water depletion and groundwater level loss of 1–2 m yr⁻¹ (Yang et al., 2006; Hu et al., 2010).

This suggests that pumping rates are not balanced by increased recharge and/or decreased discharge due to groundwater irrigation (Yang et al., 2010; Cao et al., 2013).

The storage-depletion depression cones of 40–50 m in shallow and 60–80 m in deep aquifer systems are increasing concerns for land subsidence (Xue et al., 2005; Li et al., 2011). While groundwater recovery is cited in recent years (Cui et al., 2009), there are no confirmed reports of lessening water use or increasing precipitation/recharge in the region (Yang et al., 2010; Cao et al., 2013). This suggests uncertainties in the storage dynamics due to difficulties in accounting for water use in the region. There is thus the need to refine water use strategies towards greater efficiency and sustainability in the region.

2.2 Basic concept

Pumping wells generate stress disturbances which propagate through aquifer systems and cause pressure/head loss at magnitudes dependent on the hydrogeologic conditions (Gambolati et al., 2006). Stress change due to pumping could result in the compaction of hydrogeologic formations and land subsidence (Waltham, 2002). Subsidence, which is the response of compressible hydrogeologic formation to changes in fluid pressure within the formation, is quantified in terms of effective stress as (Poland and Davis, 1969; Poland, 1984; Galloway et al., 1998):
\[
\sigma_e = \sigma_T - p_w
\]

where \(\sigma_T\) is total stress (ML\(^{-1}\)T\(^{-2}\)), \(\sigma_e\) is effective stress (ML\(^{-1}\)T\(^{-2}\)), and \(p_w\) is pore-water pressure (ML\(^{-1}\)T\(^{-2}\)).

The three common compaction stresses are gravitational stress, hydrostatic stress and dynamic seepage stress (Lofgren et al., 1968). While hydrostatic stress is largely neutral, gravitational stress and dynamic seepage stress are additive in effect and together change the void ratio and mechanical properties of aquifer deposits. The combined effect, known as geostatic pressure or total stress \((\sigma_T; ML^{-1}T^{-2})\), of sediments and water at a reference plane below the water table is the unit weight \((\gamma_m; ML^{-2}T^{-2})\) times the thickness \((z_1; L)\) of moist sediments above the water table plus the unit weight \((\gamma_b; ML^{-2}T^{-2})\) times the thickness \((z_2; L)\) of buoyant saturated sediments below the water table (Poland and Davis, 1969; Poland, 1984):

\[
\sigma_T = z_1 \gamma_m + z_2 \gamma_b
\]

where \(\gamma_m = [g(1 - n) + r_s] \gamma_w\) and \(\gamma_b = [g(1 - n) + r_w] \gamma_w\); in which \(g\) is average specific gravity of the grain deposits (–), \(r_s\) is average specific retention of the moist grains (–), \(n\) is average porosity of the aquifer deposits (–), and \(\gamma_w\) is specific weight of water (ML\(^{-2}\)T\(^{-2}\)). The functional connections of the above equations with elastic/inelastic storage coefficients (e.g., specific storage/yield, aquifer thickness, etc.) and other aquifer properties (e.g., pore water pressure, hydraulic conductivity, etc.) are detailed by Poland and Davis (1969) and Galloway et al. (1998).

Since storage largely occurs as groundwater and soil water (especially in semiarid regions), most of GRACE analyses have focused on these storage components (Syed et al., 2008; Strassberg et al., 2009; Moiwo et al., 2011). Thus the time-variant change in soil water storage \((\partial S_{sws}/\partial t; LT^{-1})\) and in groundwater storage \((\partial G_{sws}/\partial t; LT^{-1})\) is related to the time-variant change in GRACE-derived total water storage \((\partial T_{sws}/\partial t; LT^{-1})\) as (Swenson et al., 2006):

\[
\frac{\partial T_{sws}}{\partial t} = \frac{\partial S_{sws}}{\partial t} + \frac{\partial G_{sws}}{\partial t}
\]

Note that the terms in Eq. (3) are running averages for the given region of interest.

Generally, \(\partial T_{sws}/\partial t\) can be computed from GRACE month-to-month gravity fields inverted for TWS anomaly (TWSA) (Rodell et al., 2009; Longuevergne et al., 2010). Storage anomaly is the residual storage content at a given time with respect to that at a reference epoch. Then storage change is the difference in storage anomalies between successive time steps (Moiwo et al., 2011). Similarly, GLDAS global SWS or field-measured SWS can be inverted for \(\partial S_{sws}/\partial t\) (Immerzeel et al., 2010; Moiwo et al., 2011). Then groundwater levels can be inverted for \(\partial G_{sws}/\partial t\) using storage coefficients (Feng et al., 2013; Tang et al., 2013); specific yield \((S_y)\) for unconfined aquifer and specific storage \((S_s)\) for confined aquifer (Syed et al., 2008).

Generally, TWS well corrected for SWS can track field-monitored GWS especially in semiarid hydrologic conditions (Strassberg et al., 2009). Sustained groundwater overdraft in North China since the 1978 land reforms (Yang and Tian, 2009) can potentially induce aquifer compaction and land subsidence. The change in the position of mass along the axis of gravity occurs when compaction/subsidence moves mass towards the center of mass of the Earth; causing change in gravity. However, this effect is small with respect to the mass changes addressed in this study and therefore not considered.

Thus the trends in SWS, GWS and TWS are likely to suggest subsidence or uplift, which either way needs verification by LSC analysis. In this study, GRACE TWS is corrected for SWS to derive GWS. The GRACE-corrected GWS anomaly (CGWSA) is then compared with field-measured GWS anomaly (GWSA). While negative trends in GWSA and CGWSA could suggest compaction and subsidence, positive trends suggest expansion and uplift of the crustal formations. To verify the LSC due to WSD in the study area, GPS data products of relative LSC are analyzed for limited region of the study area.
2.3 Data acquisition and processing

The GRACE, GLDAS and GPS data products are used in combination with field-measured SWS and GWS. Except for GRACE (which spans from April 2002 to December 2009), the datasets cover January 2002 through December 2009. The GRACE release-05 data products from CSR (Center for Space Research), JPL (Jet Propulsion Laboratory) and GFZ (German Research Center for Geosciences) are used available at http://geoid.colorado.edu/grace/dataportal.html.

The GRACE monthly solutions are filtered for leakages (using an averaging kernel for the study area), destriped for north–south trending errors and corrected for glacial isostatic adjustments and other non-hydrological effects (Swenson et al., 2006; Paulson et al., 2007; Immerzeel et al., 2010). The data are smoothed at 200 km Gaussian half-width, fitted with the Wahr et al. (2004) error bar and truncated to within one degree of the study area. The fields are afterwards spatially averaged (based on the averaging kernel) to create time series of TWS anomaly (TWSA), from which $\partial TWS/\partial t$ is computed. The spatial average of the GRACE-estimated monthly TWSA from the 3 centers for 2002–2009 is plotted in the bottom right plate of Fig. 1.

The average of the GLDAS-estimated optimal fields of SWS from the 5 model is used in this study (Rodell et al., 2009). The GLDAS data product is verified with field-measured SWS data in the study area. Like GRACE, the GLDAS data field is truncated to within one degree of the study area and spatially averaged (based on the averaging kernel) to create time series of SWS anomaly (SWSA) and then $\partial SWS/\partial t$. The GLDAS data products used are also available at http://gdata1.sci.gsfc.nasa.gov/daac-bin/G3/gui.cgi?instance_id=GLDAS10_M.

The groundwater data are from 205 fairly distributed monitoring wells in the study area (Fig. 1, top plate). About one-third of the wells monitor groundwater level in the unconfined aquifer while the rest are for the confined aquifer. The data, also for January 2002 through December 2009, are monthly averages of observations on the 8th, 18th and 28th of every month. The data are converted into GWS using storage coefficients. To do this, the monthly groundwater level data are interpolated (Inverse Distance Weighted interpolation in ArcMap) separately for the unconfined and confined aquifers. For 2002–2009, groundwater level range across the study area was 66 to −26 m for the unconfined aquifer and 14 to −87 m for the confined aquifer.

Storage coefficients of the unconfined ($S_u$) and confined ($S_c$) aquifers are 0.012–0.264 and 0.5639–1.0668 × 10$^{-3}$ mm$^{-1}$, respectively (Feng et al., 2013; Tang et al., 2013). The mean $S_u$ is 0.07829 × 10$^{-3}$ mm$^{-1}$, which is ~33% (0.235) of $S_c$ – see http://en.cgs.gov.cn/ or http://www.cgs.gov.cn/ for details. GWSA is derived by multiplying the separately-interpolated GWS with storage coefficient for the respective aquifer systems. The GWSA for the separate aquifers are then aggregated and interpolated for $\partial GWS/\partial t$ in the same way as described for GRACE and GLDAS data products.

Note that the concept of storage coefficient is a simplified assumption of instantaneous drainage of aquifer systems with dropping pressure heads as it ignores delayed drainage effects. While storage coefficients could vary spatially, the change as a function of depth to groundwater can result in non-linear relationship between in-situ groundwater level and storage. However, these effects are not critical in this study as the analyses are long-term averages.

The GPS data of relative LSC (frame IGS08) are part of the Global Navigation Satellite System (GNSS) or International GNSS Service (IGS) available at http://www.unavco.org/crosscutting/cc-data.html. The data product is processed in GAMIT/GLOBK for daily loosely-constrained solutions of long-term surface deformation (González et al., 2012; Palano et al., 2012) in the region. GAMIT/GLOBK, developed by Massachusetts Institute of Technology (MIT) and sponsored by National Science Foundation (NSF), is a comprehensive package of programs for analyzing GPS measurements, primarily to study crustal deformations.

Generally, GRACE accuracy increases at higher spatial/temporal scales (Swenson et al., 2006), so the analysis is scaled up from monthly to yearly cycles. Here, monthly cycles are the month-to-month time series for 2002–2009 spatially averaged over the study area. Average monthly cycles are the time series averaged temporally along the
12 months of the years in 2002–2009 and spatially over the study area. The seasonal
and average seasonal cycle definitions follow the definitions for the monthly and aver-
age monthly cycles. In the study area, the period from December to February is winter,
March to May is spring, June to August is summer and September to November is
autumn.

2.4 Uncertainty/bias

Because GLDAS generally underestimates SWS, the standard deviation of the five-
contributing products is used as a measure of uncertainty (Rodell et al., 2009). The
contributing model-estimated standard deviation is 2.16 mm month\(^{-1}\), < 7 % of the av-
erage value and therefore acceptable for this study. The GLDAS SWS is in good agree-
ment with field-observed SWS (adjusted for irrigation) with \(R^2 = 0.899\) (Fig. 1, bottom
right plate). The SWS data from 80 monitoring sites in the study area are spatially aver-
ged on monthly basis (for January 2002 through December 2009) and plotted against
the GLDAS-estimated SWS in Fig. 1.

The GRACE dataset is processed based on the one-to-zero (from center-to-
peripheral) averaging kernel method (Swenson et al., 2006) to minimize leakage. The
kernel is spatially smoothed and truncation as a function of the Stokes coeffi-
cient, noise filtering and Gaussian convolution. Amplitude damping due to the kernel averaging is
corrected following the method used by Strassberg et al. (2009). Least squares anal-
ysis suggests that average random errors in the GRACE monthly data are < 5.32 mm,
which is < 5 % of the average error. Also the standard deviation of the three-contributing
GRACE data products is used as a measure of uncertainty. The estimated standard deviation is 1.64 mm, < 6 % of the average value. Because errors generally cancel
out with higher spatial/temporal averaging and storage change is derived from storage
anomaly but also for clarity, error bars are only plotted for monthly storage anomalies.
Also linear trends and regression equations are shown on the plots for a quick grasp of
the storage dynamics.

3 Results and analyses

3.1 Monthly cycles

Figure 2 plots the time series of the monthly GWS, SWS and TWS anomalies and
changes in the study area (see details in Table 1). The range of storage change (“Dif” in
Table 1) is the difference between the maximum and minimum amplitudes. It is smallest
for SWS and highest for GWS, suggesting that the storage is mainly driven by change
in GWS (Feng et al., 2013; Tang et al., 2013).

The time series (especially storage anomalies in Fig. 2a, c and e) are similar in
phase and the amplitude range smallest for SWS. The trend-lines and regression func-
tions for all the storage terms suggest declining storage in the region (Liu et al., 2006;
Zheng et al., 2010). Because up-scale averaging reduces errors by limiting outlier ef-
ects (Swenson et al., 2006), the datasets are represented at seasonal and annual
cycles to more clearly show storage dynamics in the study area.

3.2 Seasonal cycles

The time series of the seasonal storage anomalies and changes are plotted in Fig. 3
(see details in Table 1). Due to suppressed outlier effects by the up-scale averaging, the
amplitudes are smaller and phases more similar at the seasonal than at the monthly
cycles. Based on the trend lines, GWS (Fig. 3a and b), SWS (Fig. 3c and d) and TWS
(Fig. 3e and f) generally decline in the study area. Like in the monthly cycles, the
amplitude change is smallest for SWS. The storage trends show not only seasonality,
but are also significant at \(p < 0.05\) and \(\alpha = 0.05\). The seasonality is attributed to the
hydro-climatic and agronomic conditions in the study area (Moiwo et al., 2011).

3.3 Yearly cycles

The amplitudes of the yearly storage dynamics in Fig. 4 are further reduced (due to
the up-scale averaging), with the smallest amplitude range still for SWS. The phases
of the storage anomalies (Fig. 4a, c and e) and changes (Fig. 4b, d and f) are a lot more similar at this scale. Like in the other cycles, all the trends in TWS, SWS and GWS terms are negative (Feng et al., 2013; Tang et al., 2013). The low SWS variability is attributed to irrigation during non-raining cropping periods. In fact, SWS change is generally minimal in irrigated regions (Rodell et al., 2009).

The annual storage trends are significant at \( p < 0.01 \) and \( \alpha = 0.05 \). The storage anomalies (Fig. 4a, c and e) and changes (Fig. 4b, d and f) show that the storage is highest in 2004. Although the trends suggest an apparent recovery in 2006–2008, the overall storage remains negative. This contradicts recent reports of groundwater recovery in the region (Cui et al., 2009). The continuous WSD could cause deformation of the geo-matrix and land subsidence in the region (Zheng et al., 2010).

### 3.4 Average monthly cycles

Figure 5 depicts the dynamics of the average monthly storage terms in the study area. GWS increases from January through February and then decreasing through March. It is lowest in June and again rebounds through December (Fig. 5a). The corresponding GWS change decreases from January through April, recovers through September and again decreases through December (Fig. 5b). The rates of precipitation (storage input), irrigation (storage input/output) and ET (storage output) greatly influence storage dynamics in the study area (Hu et al., 2010; Cao et al., 2013a; Tang et al., 2013).

The SWS and TWS terms have different amplitudes but similar phases. The phases generally decline from January through May/June, increase to peak values in August and again decline through December (Fig. 5c and e). The corresponding trends in SWS and TWS change track a trough-shaped curve for the period from May through September. The trends in storage change are not so pronounced before May and after September (Fig. 5d and f). As GRACE is relatively less sensitive to short-term storage changes, TWS (Fig. 5b and f) is a bit out of phase with measured GWS (Fig. 5a and e) in the study area.

### 3.5 Average seasonal cycles

Figure 6 presents the dynamics of the storage terms averaged seasonally for 2002–2009. The dynamics of GWS and SWS are similar but altogether different from that in TWS; further pointing out the low sensitivity of GRACE to short-term storage changes (Swenson et al., 2006). GWS anomaly declines from winter through summer, before rebounding in autumn (Fig. 6a). That of SWS declines from winter to spring and rebounds through autumn (Fig. 6c). Also GWS as well as SWS change (Fig. 6b and d) decline from winter through spring and rebound through autumn.

The trends in GRACE average seasonal TWS anomaly (Fig. 6e) and change (Fig. 6f) track a forward-leaning S-shaped-like curve. It is lowest in spring and highest in summer; depicting the overall storage characteristics in the region. The storage dynamics is the combined effects of the water use, precipitation and ET in the region (Yang et al., 2010; Cao et al., 2013).

### 3.6 Groundwater storage anomaly

GWS anomaly derived from field observations (GWSA) is plotted along with that derived from GRACE/GLDAS datasets (CGWSA) in Fig. 7. Here, storage anomaly is the storage variation relative to the mean (monthly, seasonal or annual) storage for 2002–2009 and then storage change is the difference in two successive time periods. Thus the trend in GWS anomaly in effect reflects the rate of groundwater depletion in the study area. The degree of the trends (given by the trend-lines and regression functions) in GWSA (Fig. 7, left plots) and in CGWSA (Fig. 7, right plots) are negative at the monthly, seasonal and yearly scales, suggesting that there is storage loss in the study area (Yang et al., 2006; Hu et al., 2010).

Water budget analysis based on the GRACE data shows WSD at 23.76 ± 1.74 mm yr\(^{-1}\), the equivalent of 13.73 ± 1.01 km\(^3\) yr\(^{-1}\) for the 578 000 km\(^2\) study area. Also GLDAS data product analysis shows that 20.75 % (4.93 ± 0.36 mm yr\(^{-1}\); 2.85 ± 0.21 km\(^3\) yr\(^{-1}\)) of that loss is in SWS. Groundwater and other storage sources account
for 79.25% (18.83 ± 1.38 mm yr\(^{-1}\) or 10.88 ± 0.80 km\(^3\) yr\(^{-1}\)) of the storage loss. Feng et al. (2013) noted that about 67% of storage loss in North China is from aquifer storage. The estimated contribution of GWS loss to TWS loss in the region is 76.84% (18.26 ± 1.02 mm yr\(^{-1}\) or 10.55 ± 0.41 km\(^3\) yr\(^{-1}\)). The cumulative storage loss for the 8 year period (2002–2009) is 109.87 ± 8.05 km\(^3\) for TWS, 84.40 ± 3.28 km\(^3\) for GWS and 22.80 ± 1.68 km\(^3\) for SWS. Thus as noted by Strassberg et al. (2009) for the highland semiarid region in the US, surface water use has little affect (<3%) on overall storage dynamics in the study area.

Sustained WSD over long periods of time could cause land subsidence and further limit the storage and water supply (Gambolati et al., 2006; González et al., 2012). To determine that there could be subsidence due to sustained WSD in the study area, GPS data products is used to analyze LSC in the region.

### 3.7 Land subsidence

Because of the dense population, high agro-industrial expansion and semiarid climatic conditions, water use in North China exceeds natural recharge rate (Hu et al., 2010; Wang et al., 2011; Feng et al., 2013). Since there are hardly any occurrences of earthquakes or large-scale earth-faults in the region, land surface deformation could only be caused by abstractions of groundwater, hydrocarbons or coal (Liu et al., 2008). Thus GPS data product of relative LSC is used to analyze for land subsidence due to loss of water storage in the region.

The GPS data analysis suggests the occurrence of subsidence at an estimated rate of 7.29 ± 0.35 mm yr\(^{-1}\) in the vertical component of the IGS08 station in Beijing (Fig. 7g). Note that Beijing is one of the zones with severe groundwater depression cones in China. The vertical dip of hydraulic gradient from moderate-to-worse WSD zones is 4–30% in Beijing (Wang et al., 2011). Groundwater accounts for >61% of water supply in Beijing, 26% in Tianjin, 80% in Hebei province and 58% in Shanxi province (Cao et al., 2013).

As the degree of subsidence is loosely related to the magnitude of groundwater depression cone (González et al., 2012), subsidence in Beijing (where depression cones are severe) could be worse than the average for the rest of North China (where there are many areas with little/no depression cones). Analysis of well data shows that storage loss is 62.41% higher in Beijing than the average for the 578,000 km\(^2\) North China study area. Given the spatial averaging, a similar hydrogeologic conditions is assumed across the study area. Then a relational analysis suggests the occurrence of subsidence at an average rate of 2.74 ± 0.16 mm yr\(^{-1}\) in the entire study area. Because the assumption here ignores the basis for complex subsidence processes, the estimated subsidence only conservative and must be treated with caution. Irrespective, it could guide future subsidence analyses at specific strategic locations of the study area (e.g., Tianjin, Shijiazhuang and Baoding) with WSD comparable to that in Beijing.

Shi et al. (2006) noted that since land reforms in the 1970s, subsidence has been occurring at the rate of 2.97 mm yr\(^{-1}\) in North China Plain. Skeletal matrix can be re-arranged under increased effective stress exceeding pre-consolidation stress in fine grain deposits, resulting in reduced porosity that is permanent and non-recoverable. Skeletal matrix generally compresses under decreased effective stress (not exceeding pre-consolidation stress), also resulting in reduced porosity. However, this deformation is recoverable under stresses in the elastic range. Crushing of granular components of the matrix is possible in sand and especially diatomaceous deposits, but this is rare and tends to occur under very large increases in effective stress (Poland, 1994). Thus high storage depletion (Wang et al., 2011; Cao et al., 2013) in North China could cause any of the above subsidence conditions.

### 4 Discussions

Groundwater accounts for some 80% of irrigation (Yang et al., 2010; Hu et al., 2010) and 20% of industrial and domestic (Feng et al., 2013) water use in North China. The year-on-year groundwater level decline (of 1–2 m) suggests that withdrawal exceeds
recharge (Cao et al., 2013). This unsustainable reliance on groundwater is also reflected in the widespread WSD in the region (Yang et al., 2006; Moiwo et al., 2010; Wang et al., 2011).

Based on GPS data analysis, subsidence due to WSD is $7.29 \pm 0.35 \text{ mm yr}^{-1}$ in Beijing and $2.74 \pm 0.16 \text{ mm yr}^{-1}$ in the rest of North China. Assuming that the subsidence is all drainable water, it is the equivalent of $0.12 \pm 0.01 \text{ km}^3$ for the $16,080 \text{ km}^2$ Beijing area and $1.58 \pm 0.09 \text{ km}^3$ for the $578,000 \text{ km}^2$ North China region. There is also a considerable water unloading via mineral/coal mining operations in the region (Feng et al., 2013). As WSD will continue into the foreseeable future due to the lack of reliable alternatives, pumping-induced subsidence could worsen environmental conditions in the region in the long run (Wang et al., 2011).

Droughts, degenerated water/wetland ecosystems and earthquakes are variously associated with long-term WSD. Pumping-drained groundwater level loss of 100 m in Las Vegas Valley (USA) in the 1950s resulted in $\sim 2 \text{ m}$ subsidence, severely damaging infrastructures in the region (Bell, 1981). Intense pumping in the 1940s in Antelope Valley caused $\sim 2 \text{ m}$ subsidence in Lancaster and $>1 \text{ m}$ subsidence in Rogers Lake region, USA (Galloway et al., 1998; Hoffmann et al., 2003). Groundwater drawdown of $>250 \text{ m}$ during 1960–2010 triggered the 2011 earthquake of 5.1 moment magnitude in Lorca, Spain (González et al., 2012).

Numerous reports have pointed to WSD and/or subsidence in a number of regions North China (Li, 2003; Liu et al., 2006; Moiwo et al., 2010; Wang et al., 2011; Feng et al., 2013). Consistent with other studies (Foster et al., 2004; Kendy et al., 2007; Tang et al., 2013), this study shows WSD in North China. The storage loss ($1.58 \pm 0.09 \text{ km}^3$) due to subsidence is $\sim 12 \%$ of TWS loss ($13.73 \pm 1.01 \text{ km}^3$) in the ($578,000 \text{ km}^2$) study area. This underscores the need for water-wise strategies which will enhance recharge and limit discharge in the region (Cui et al., 2009; Cao et al., 2013).

There is SWS, GWS and TWS loss in the study area. After correcting for storage loss due to subsidence, annual WSD in the region is $21.02 \pm 1.58 \text{ mm} (12.15 \pm 0.91 \text{ km}^3)$ for TWS and $15.52 \pm 0.76 \text{ mm} (8.97 \pm 0.44 \text{ km}^3)$ for GWS. Of course root-zone SWS is largely insensitive to subsidence. The small SWS change (Moiwo et al., 2011) is due to extensive irrigation in the region (Han et al., 2008; Yang et al., 2010). The storage-unloading subsidence mainly occurs in the deep aquifer systems (Xue et al., 2005; Zhang et al., 2008; Li et al., 2011; Cao et al., 2013).

The storage trends are significant at the seasonal and yearly scales and have clear seasonality. Currently, there are no significant storage recovery measures such as to reduce groundwater use or increase recharge in the region (Cui et al., 2009). In fact, GRACE signal can detect storage recoveries driven by such measures (Feng et al., 2013). GWS is lowest and SWS highest during summer precipitation months (Fig. 5), suggesting that precipitation is most effective for root-zone SWS in the region (Han et al., 2008). High SWS in the other seasons is due to irrigation (Yang et al., 2010). Over-reliance on groundwater causes a steady decline in GWS (Yang et al., 2006). GRACE captures the overall conditions as storage loss (Fig. 5e). While the average monthly trends suggest that the lowest storage is in summer, the average seasonal trends clearly reflect the prevailing hydro-climatic conditions.

The monthly, seasonal and yearly storage anomalies are negative for GWS derived from well data (GWSA) and that from GRACE/GLDAS data (CGWSA) for 2002–2009 (Fig. 7). As argued in Eq. (2) and the subsequent discussions, GWS from GRACE/GLDAS should track that from well data. The storage trends at the various scales are negative, suggesting storage loss. Even historical records of water use in the region suggest WSD (Foster et al., 2004; Zhang et al., 2008). Excessive WSD can cause land subsidence (Shi et al., 2006; Wang et al., 2011; González et al., 2012).

The condition for subsidence is verified using GPS data of relative LSC. The analysis suggests the occurrence of subsidence at $7.29 \pm 0.35 \text{ mm yr}^{-1}$ in Beijing and $2.74 \pm 0.16 \text{ mm yr}^{-1}$ in the North China study area (Fig. 7g). Subsidence driven by WSD could have disastrous implications for ecological sustainability, water availability, food security and social stability in this pivotal region. Thus there is the need to reorient water-use policies towards greater efficiency and sustainability in this semiarid region.
5 Conclusions

Our current understanding of WSD and the approaches towards sustainable water use is limited. This study contributes to the broader efforts of deepening this knowledge for greater efficiency and sustainability in water use. It integrates satellite, model and field data products to show WSD and subsidence in North China. Error analysis shows that the results are not the effects of data biases or artifacts of the analysis.

The rate of TWS loss in the study area is $23.76 \pm 1.74 \text{ mm yr}^{-1}$ ($13.73 \pm 1.01 \text{ km}^3 \text{ yr}^{-1}$), which is $30.52\%$ of the planned south–north water transfer ($45 \text{ km}^3 \text{ yr}^{-1}$) in 2050. Cumulative WSD over 2002–2009 is $190.08 \pm 13.92 \text{ mm}$ ($109.87 \pm 8.05 \text{ km}^3$), which is $244.12\%$ of the slated water transfer. Subsidence due to WSD is $7.29 \pm 0.35 \text{ mm yr}^{-1}$ in Beijing and $2.74 \pm 0.16 \text{ mm yr}^{-1}$ in North China. About $11.53\%$ ($2.74 \pm 0.18 \text{ mm}; 1.58 \pm 0.12 \text{ km}^3$) of TWS and $8.37\%$ ($1.52 \pm 0.70 \text{ mm}; 0.88 \pm 0.03 \text{ km}^3$) of GWS are attributed to storage reductions accompanying subsidence in the region. Excessive groundwater use causes storage depletion and subsidence in the long-term. Subsidence further limits GWS by reducing crustal pore spaces. Because GRACE merely detects mass change, it is used in combination with LSC data to show subsidence in this study. The concurrence of groundwater depletion and subsidence could grow into water crisis in the region.

North China is a pivotal agricultural, industrial and political base where WSD and subsidence could destabilize millions in the region. It is critical for relevant stakeholders to embark on greater efficiency in water use. Use of alternative water resources could mitigate the WSD and subsidence. Aquifer recharge, saltwater use and inter-basin water transfer could enhance GWS, food security and social stability in the region.

Acknowledgements. The study is supported by National Natural Science Foundation of China (No. 91325302) and the Chinese Academy of Sciences Visiting Fellowship for Developing Countries (2013FFZA0019).

References


Immerzeel, W. W., van Beek, L. P. H., and Bierkens, M. F. P.: Climate change will affect the Asian water towers, Science, 328, 1382–1384, 2010.


Table 1. Details of the GLDAS soil water, well-monitored groundwater and GRACE total water storage trends for the North China study area.

<table>
<thead>
<tr>
<th>Storage variable</th>
<th>Storage parameter</th>
<th>Min (mm)</th>
<th>Max (mm)</th>
<th>Dif (mm)</th>
<th>Mean (mm)</th>
<th>SD (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GLDAS Anomaly</td>
<td>Month</td>
<td>−58.56</td>
<td>83.41</td>
<td>141.97</td>
<td>0.56</td>
<td>31.93</td>
</tr>
<tr>
<td></td>
<td>Season</td>
<td>−30.00</td>
<td>66.38</td>
<td>96.38</td>
<td>2.31</td>
<td>27.29</td>
</tr>
<tr>
<td></td>
<td>Avgmon</td>
<td>29.80</td>
<td>18.93</td>
<td>53.10</td>
<td>0.56</td>
<td>16.36</td>
</tr>
<tr>
<td></td>
<td>Change</td>
<td>−10.95</td>
<td>16.80</td>
<td>27.75</td>
<td>2.31</td>
<td>11.57</td>
</tr>
</tbody>
</table>

|                  | Month             | −30.00   | 66.38    | 96.38    | 2.31      | 27.29   |
|                  | Season            | −33.24   | 40.39    | 73.63    | 0.00      | 22.72   |
|                  | Avgmon            | −29.80   | 18.93    | 53.10    | 0.56      | 16.36   |
|                  | Change            | −10.95   | 16.80    | 27.75    | 2.31      | 11.57   |

| Groundwater Anomaly | Month             | −112.18  | 113.96   | 226.14   | 18.67     | 47.78   |
|                     | Season            | −96.45   | 78.20    | 174.66   | −14.35    | 43.20   |
|                     | Avgmon            | −58.12   | 52.95    | 111.07   | −7.83     | 34.79   |
|                     | Change            | −49.03   | 73.12    | 122.15   | −4.17     | 47.44   |
|                     | Avgmon            | −36.03   | 22.15    | 58.18    | 0.00      | 16.90   |
|                     | Change            | −30.25   | 31.76    | 62.01    | −4.87     | 29.52   |

| GRACE Anomaly      | Month             | −77.59   | 98.91    | 176.50   | −1.41     | 34.58   |
|                    | Season            | −57.11   | 58.64    | 115.75   | −1.30     | 23.78   |
|                    | Avgmon            | −34.47   | 35.56    | 70.04    | 0.00      | 21.12   |
|                    | Change            | −25.55   | 36.63    | 62.18    | −4.18     | 24.53   |
|                    | Avgmon            | −33.06   | 45.79    | 88.65    | 1.41      | 18.73   |
|                    | Change            | −30.13   | 10.95    | 40.88    | 2.28      | 7.60    |

Note that the column highlighted grey is in meter (m); then Min is minimum value; Max is maximum value; Dif is the difference between Max and Min; SD is standard deviation; Avgmon is average monthly value; and Avgseas is average seasonal value.
Figure 1. Map of North China depicting groundwater monitoring well and GPS relative land surface data sites, provincial boundaries and the ambient topography (top plate), GRACE averaged monthly total water storage anomaly (bottom left plate), and a plot of field-measured vs. GLDAS soil water storage (bottom right plate) for the North China study area.

Figure 2. Time series (2002–2009) of monthly groundwater storage anomaly (GWSA, a), groundwater storage change (δGws/δt, b), soil water storage anomaly (SWSA, c), soil water storage change (δSws/δt, d), GRACE total water storage anomaly (TWSA, e), and GRACE total water storage change (δTws/δt, f) for the North China study area. Error bars denote standard deviation of the storage variables.
Figure 3. Time series (2002–2009) of seasonal groundwater storage anomaly (GWSA, a), groundwater storage change (\(\partial Gws/\partial t\), b), soil water storage anomaly (SWSA, c), soil water storage change (\(\partial Sws/\partial t\), d), GRACE total water storage anomaly (TWSA, e), and GRACE total water storage change (\(\partial Tws/\partial t\), f) for the North China study area.

Figure 4. Time series (2002–2009) of yearly groundwater storage anomaly (GWSA, a), groundwater storage change (\(\partial Gws/\partial t\), b), soil water storage anomaly (SWSA, c), soil water storage change (\(\partial Sws/\partial t\), d), GRACE total water storage anomaly (TWSA, e), and GRACE total water storage change (\(\partial Tws/\partial t\), f) for the North China study area.
Figure 5. Time series (2002–2009) of average monthly groundwater storage anomaly (GWSA, a), groundwater storage change (∂Gws/∂t, b), soil water storage anomaly (SWSA, c), soil water storage change (∂Sws/∂t, d), GRACE total water storage anomaly (GRACE TWSA, e), and GRACE total water storage change (∂Tws/∂t, f) for the North China study area.

Figure 6. Time series (2002–2009) of average seasonal groundwater storage anomaly (GWSA, a), groundwater storage change (∂Gws/∂t, b), soil water storage anomaly (SWSA, c), soil water storage change (∂Sws/∂t, d), GRACE total water storage anomaly (GRACE TWSA, e), and GRACE total water storage change (∂Tws/∂t, f) for the North China study area.
Figure 7. Parallel plots of monthly (a and b), seasonal (c and d) and yearly (e and f) anomalies of groundwater storage separately derived from well-monitored groundwater data (GWSA) and GRACE/GLDAS data (CGWSA) products for the North China study area. The bottom plate (g) depicts the vertical component of relative land surface derived from the GPS data station IGS08 (in Beijing) for January 2002 through December 2009.