Improved methods for satellite-based groundwater storage estimates: A decade of monitoring the high plains aquifer from space and ground observations

Jose Agustin Breña-Naranjo1,2, Anthony D. Kendall1, and David W. Hyndman1

1Department of Geological Sciences, Michigan State University, East Lansing, Michigan, USA, 2Instituto de Ingeniería, Universidad Nacional Autónoma de México, Coyoacán, México

Abstract

The impacts of climate extremes and water use on groundwater storage across large aquifers can be quantified using Gravity Recovery and Climate Experiment (GRACE) satellite monitoring. We present new methods to improve estimates of changes in groundwater storage by incorporating irrigation soil moisture corrections to common data assimilation products. These methods are demonstrated using data from the High Plains Aquifer (HPA) for 2003 to 2013. Accounting for the impacts of observed and inferred irrigation on soil moisture significantly improves estimates of groundwater storage changes as verified by interpolated measurements from ~10,000 HPA wells. The resulting estimates show persistent declines in groundwater storage across the HPA, more severe in the southern and central HPA than in the north. Groundwater levels declined by an average of approximately 276 ± 23 mm from 2003 to 2013, resulting in a storage loss of 125 ± 4.3 km3, based on the most accurate of the three methods developed here.

1. Introduction

Globally, aquifer systems face enormous pressure from unsustainable groundwater withdrawals [Wada et al., 2010]. Although groundwater depletion is most commonly attributed to agricultural activities, its magnitude and extent generally depend on pumping rates as well as variable climate and baseflow discharge to surface water bodies [Konikow and Kendy, 2005; Scanlon et al., 2012]. Over the last decade, satellite-based estimates of global groundwater depletion have been possible due to the Gravity Recovery and Climate Experiment (GRACE) satellite [Rodell and Famiglietti, 1999]. GRACE provides nearly continuous estimates of total water storage changes from groundwater, soil moisture, and surface water systems over large river basins and aquifers around the world [Schmidt et al., 2006; Wada et al., 2010].

The High Plains Aquifer (HPA) in the central United States has been one of the most analyzed groundwater systems in the world due to the regional importance of irrigated agriculture. The HPA has been the focus of several groundwater assessment studies in which GRACE data are used for monitoring and/or validation of models [e.g., Strassberg et al., 2009; Döll et al., 2012]. Although groundwater pumping is the main driver of depletion [Stanton et al., 2011], the role of irrigation has not been explicitly included in groundwater storage estimates based on GRACE data.

The standard approach in GRACE-driven hydrogeological studies is to subtract simulated unsaturated zone storage changes from total water storage anomalies to estimate saturated zone changes. However, substantial differences in global and regional groundwater depletion estimates [Wada et al., 2010; Konikow, 2011] have highlighted the importance of processes that are typically neglected in GRACE studies. Moreover, irrigation practices enhance seasonal evapotranspiration and increase atmospheric moisture and cooling [Cochran and Brunsell, 2012]. Neglecting such processes during the irrigation season can lead to severe underestimation of groundwater storage depletion, with implications for quantifying the impacts of global change [e.g., Döll et al., 2012; Konikow, 2013; Pokhrel et al., 2012].

Previous attempts to estimate the impact of aquifer pumping on groundwater depletion in the HPA using GRACE data had large uncertainties [e.g., Döll et al., 2012]. Here we account for added soil moisture during the irrigation season, which is traditionally neglected in GRACE studies, by estimating coefficients to match measured storage changes for three regions of the aquifer integrated from ~10,000 groundwater level monitoring wells across the HPA. We develop several new methods that provide more accurate estimates of groundwater storage changes from GRACE data and compare these to measured groundwater level changes. Groundwater withdrawal in the HPA
and its subsequent enhancement of soil moisture storage in irrigated areas will likely induce significant changes in water storage below the root zone due to annual variations in irrigation return flow. Our study covers the last decade of observations and model estimates, which includes a significant drought that began in the southern High Plains in 2010 and expanded to the northern High Plains by 2012. This study (1) provides improved estimates of water storage variations across the HPA over the last decade using the latest GRACE and data assimilation products, (2) tests the sensitivity of the results to enhanced soil moisture from seasonal irrigation, and (3) demonstrates an approach to correct estimates of groundwater withdrawals in regions with limited pumping data.

2. Data and Methods

Processed and quality-controlled estimates of total water storage anomalies $dTWS/dt$ from monthly GRACE CSR-RL05 [Landerer and Swenson, 2012] land mass grids were obtained from March 2003 to February 2013, with $1 \times 1$ degree resolution expressed in mm month$^{-1}$. Estimates of water storage changes in the top 200 cm of the unsaturated zone were calculated using the Noah model from the North American Land Data Assimilation System (NLDAS) phase 2 with 0.125 degree resolution [Xia et al., 2012], since in situ soil moisture data are too sparse to estimate soil moisture anomalies across the domain. Daily soil water content estimates for the study period were aggregated into a monthly time series for each pixel and converted to anomalies ($dSM/dt$) relative to a 2004–2009 baseline period.

GRACE and NLDAS cells that are at least 75% within the HPA were considered to be representative of storage changes in the aquifer and were thus aggregated within three aquifer regions: Southern (SHP), Central (CHP), and Northern (NHP) High Plains (see Figure 1). The remaining cells were neglected in the study. After this filtering, there were 7 cells in the SHP, 9 cells in the CHP, and 23 cells in the NHP, over a total area of approximately 400,000 km$^2$, which remains within the typical GRACE analysis footprint. As noted by Strassberg et al. [2009], the sum of storage changes in surface water, snow, ice, and biomass are minor relative to the total HPA regional water storage and are thus usually neglected. Under such assumptions, GRACE-based groundwater storage anomalies ($dGWS/dt$) are typically approximated as:

$$\frac{dGWS}{dt} = \frac{dTWS}{dt} - \frac{dSMS}{dt}$$

where total water storage is TWS, and soil moisture storage is SMS.

We developed three methods to modify GRACE estimates to account for increased soil moisture due to irrigation, which are not incorporated in NLDAS-2 [Xia et al., 2012]. The first method called Soil
Moisture + Irrigation \((SM + IRR)\), assumes that additional storage not accounted for by NLDAS is proportional to the total applied irrigation water. For this method, the product of a coefficient \((C_1)\), interpreted as the fraction of soil water provided by irrigation that remains in the unsaturated zone and is not accounted for by NLDAS-2, and applied irrigation \((I)\) is subtracted from equation (1).

\[
\frac{dGWS}{dt} = \frac{dTWS}{dt} - \frac{dSMS}{dt} - I \times C_1
\]  

(2)

A second method, called Soil Moisture Deficit \((SMD)\), assumes that the irrigation effect on storage is proportional to the difference between the NLDAS estimates of potential evapotranspiration rate \((PET)\) and the evapotranspiration \((ET)\) rate.

\[
\frac{dGWS}{dt} = \frac{dTWS}{dt} - \frac{dSMS}{dt} - C_2(PET_{GS} - ET_{GS})
\]  

(3)

where \((PET_{GS} - ET_{GS})\) is the difference between the two NLDAS simulated rates in irrigated areas during the growing season, from March through September. This method assumes that irrigation demand is proportional to the ET deficit, and therefore irrigated areas evaporate close to the PET rate, where \(C_2\) has a similar meaning to \(C_1\), except it is a proxy for irrigation water use defined: the difference between PET and ET. Thus, the coefficients \(C_1\) and \(C_2\) can be interpreted as proportional to the sum of irrigation-enhanced ET and recharge (see supporting information).

The third method, called Simple Irrigation \((IRR)\), provides the simplest estimate of groundwater storage variations by neglecting near-surface processes and assuming that all soil moisture anomalies in heavily irrigated regions, such as over the HPA, can be estimated using irrigation pumping data (observation or estimates). Here, changes in soil moisture from NLDAS are replaced by the total irrigation water applied to cropland multiplied by a coefficient \((C_3)\), which is a direct measure of the proportion of irrigation water applied that remains in the unsaturated zone. This assumes that soil moisture variation is negligible on non-irrigated lands.

\[
\frac{dGWS}{dt} = \frac{dTWS}{dt} - \left(I \times C_3\right)
\]  

(4)

For the \(SM + IRR\) and \(IRR\) methods, \(I\) is the annual reported (estimated for years when observations are not available) groundwater withdrawal volume for irrigation purposes in the SHP and CHP, while in Nebraska it is the sum of estimated groundwater and surface water withdrawals (see supporting information, Table A2). These annual observations and estimates were only available until 2010; thus, groundwater withdrawals in 2011 and 2012 were assumed as the 2003–2010 average.

The groundwater anomalies from each of the four methods were compared to aggregated storage based on kriged annual water level change measurements (between consecutive spring seasons) using 8182 to 10,851 observation wells monitored by state and federal agencies each year. Simulated \(dGWS/dt\) anomalies from equations (2)–(4) were expressed as monthly changes in groundwater storage in each region. The 2003–2013 period was divided in three subperiods that reflect different climate regimes: near-normal rainfall from 2003 to 2007, a wet period in the NHP from 2008 to 2010, and an exceptional drought since 2011. The average interpolated \(dGWS/dt\) values for the three regions were multiplied by an estimated HPA-average specific yield of 0.15 [Strassberg et al., 2009]. Monthly estimated groundwater storage anomalies were sequentially added to an initial groundwater storage estimate, calculated as the specific yield times the interpolated saturated thickness in 2003.

For each method, coefficients were estimated to fit observed groundwater anomalies for three climate periods from 2003 to 2013, including a major drought (Table A1), as well as for the entire period.

Uncertainties for: GRACE data were derived from Landerer and Swenson [2012], NLDAS soil moisture and potential evapotranspiration products as the standard deviation of the month-to-month estimates [Famiglietti et al., 2011], and for pumping as the standard deviation of the annual pumping time series for each region. The uncertainty in pumping for the NHP was replaced with the CHP estimate due to the lack of a complete time series for that region, which artificially reduced the variance. Total method uncertainties were added according to equations (1)–(4), with the uncertainties in pumping and PET scaled by the estimated coefficients as appropriate for each method. Uncertainties were first calculated in mm and then scaled to km^3 based on the area of each region. Estimates for HPA total uncertainties were calculated as the area-weighted
averages of the regional uncertainties. See Table A6 for a complete list of uncertainty terms. Trend uncertainties were calculated according to Taylor [1997].

3. Results

Trends in monthly total water storage and soil moisture storage anomalies (relative to 2004 through 2009 baseline) across the HPA from 2003 to 2013 are shown in Figure 1 for the SHP, CHP, and NHP. Average precipitation increases by nearly a factor of three from west to east, with only a slight decrease from North to South (Figure 1a). Changes in monthly total water storage show recent periods with anomalies as high as +177 ± 22 mm in the NHP and as low as −141 ± 22 mm in the SHP (Figure 1b). The unsaturated storage anomalies from NLDAS-2 (Figure 1c) clearly show the effect of the recent drought on simulated soil moisture, starting in the SHP then propagating to the NHP.

The monthly total water storage anomalies from 2003 to 2008 were similar for the three HPA regions, except for a late 2004 to early 2005 wet period in the SHP (Figure 1d). A marked divergence then began in 2008, with significant increases in total water storage in the NHP and declines in the SHP and the CHP (starting in 2009). A temporary rise of \( dTWS/dt \) is likely due to the increase in seasonal precipitation during the growing season (May–September) from 179 mm in 2006 to 430 mm in 2010 (see supporting information, Figure A1). The drought that started in 2010 resulted in significant storage declines for the entire HPA (Figure 1d), with large regional variability and a later arrival to the NHP. The difference in total water storage anomalies between the NHP and SHP rose from ~50 mm in 2008 to ~150 mm in 2011 during the early phase of the drought. Estimates of soil moisture storage anomalies from NLDAS show similar trends in regional anomalies from 2003 to 2008, followed by generally more positive soil moisture anomalies in the North relative to the CHP and SHP, with some convergence in 2012 (Figure 1e).

Variations in groundwater storage for the three HPA regions are shown in Figure 2 using the standard method of subtracting NLDAS soil moisture (SM) anomalies from the GRACE estimates (gray lines) along with our three new correction methods: two (SM + IRR and SMD) that add a term to the standard method to account for irrigation effects (red and blue lines, respectively), and another (IRR) that uses only irrigation data (green line). All three correction methods provided much better estimates of observed depletion rates than the traditional NLDAS-based approach (SM); mean absolute errors and \( R^2 \) values for the four methods with respect to observed values are given in supporting information Table A3. The IRR approach proved to be substantially better than the other methods, especially for the SHP (Figure 2). The results are based on coefficients estimated for three climatic periods between 2003 and 2012 (see supporting information, Table A1).

Surprisingly, the simplest of the three methods, which uses a single coefficient multiplied by irrigation as the anomaly (green line, IRR method), performed better than the more complex methods (see supporting information, Table A3). \( R^2 \) values for the IRR method were 0.79, 0.93, and 0.72 for the NHP, CHP, and SHP, respectively, compared to the much lower 0.18, 0.00, and 0.02 values for the same regions with the traditional SM approach. Estimated coefficients for this simple method across the entire period were 0.28, 0.47, and 0.27 for the NHP, CHP, and SHP, respectively (see supporting information, Table A1), implying a larger impact of irrigation on soil moisture in the CHP than either the NHP or SHP. Optimal \( C_1 \) values in the \( SM + IRR \) method for the NHP, CHP, and SHP were 0.19, 0.44, and 0.24, while \( C_2 \) values for the SMD method were estimated to be 0.07, 0.08, and 0.03 for the three regions (see Figure A4). As discussed above, \( C_1 \) and \( C_2 \) have similar physical interpretation, but different magnitudes, and should exhibit similar patterns among the regions if the SMD approach provides a reasonable estimate of applied irrigation. The estimated storage changes were similar for both the \( SM + IRR \) and SMD methods, as shown in Figure 2 and Table A3, indicating that the SMD method can provide reasonable estimates of applied irrigation.

Based on average trends from the best correction method (IRR), we estimate that the net groundwater change across the entire HPA is \(-27.6 ± 3.3\) mm/yr or \(-12.5 ± 0.61\) km\(^3\)/yr, from March 2003 to February 2013. During the same period, the average trends for the \( SM + IRR \) and SMD methods were \(-17.2 ± 6.0\) mm/yr and \(-7.9 ± 1.1\) km\(^3\)/yr, respectively. The SM method showed much milder trends on the order of \(-1.9 ± 5.6\) mm/yr or \(-0.9 ± 1.1\) km\(^3\)/yr. Trends for the SHP, CHP, and NHP were calculated separately for three periods to highlight changes due to climate variability (see supporting information Tables A5 and A7). From 2003 to 2008, when GRACE anomalies were similar for the three regions, the IRR method indicates that the HPA storage changed by \(-7.5 ± 1.2\) km\(^3\)/yr. Wetter conditions in the NHP from 2008 to 2010 resulted in storage increases in the NHP,
partially offsetting declines in the other two regions (see supporting information, Figure 1). Observed groundwater storage changes for the 2003 to 2010 period were approximately $43 \, \text{km}^3$, similar to the IRR estimate of $37.8 \pm 4.3 \, \text{km}^3$. In contrast, the significant drought from 2010 to 2013 resulted in substantial storage declines of $93.6 \pm 6.8 \, \text{km}^3$ across the entire HPA. The largest storage declines during this drought period were in the CHP and NHP, accounting for $\sim 84\%$ of the total aquifer depletion during this period.

4. Discussion

Traditional NLDAS-corrected GRACE estimates significantly underestimate the amount of groundwater depletion in areas of intensive irrigation, such as the HPA. The three correction methods we present can account for irrigation-induced soil moisture changes in addition to the unsaturated zone storage changes that respond directly to seasonal climate variations. Corrected estimates with constant coefficients reasonably describe temporal variations in observed groundwater levels across all three regions of the HPA. Overall, levels of depletion prior to the recent drought are similar to those described by earlier studies [e.g., Konikow, 2011; Scanlon et al., 2012]. From March 2003 to February 2013, the IRR method estimates that the NHP and CHP regions contributed groundwater storage declines of 53.0 and 52.1 $\, \text{km}^3$, respectively, whereas the SHP experienced a smaller decline of 19.7 $\, \text{km}^3$. Annual estimates by region, estimation method and period are shown in Figure 2 and in supporting information Table A1.

Unsaturated zone water storage anomalies were similar to, or even larger than, total water storage anomalies (Figure 1). $d\text{SMS}/dt$ exceeded $d\text{TWS}/dt$ several times, especially during the drought. This helps explain why remotely sensed estimates of groundwater changes may underestimate observed values (Figure 2, gray lines). Corrections that account for irrigation-driven processes provide a solution to the recognized issue that data assimilation products such as GLDAS and NLDAS do not capture the response of aquifers to irrigation and other anthropogenic activities, especially important during extended droughts [Famiglietti et al., 2011; Thomas et al., 2014; Xia et al., 2012]. These assimilation products can describe natural variations due to climate variability but they cannot yet estimate surface water consumption [Anderson et al., 2012] or enhanced soil moisture during the irrigation season [Voss et al., 2013].

[Figure 2. Left column (A–C): Monthly groundwater storage derived from the standard (SM) and simple irrigation (IRR) methods (gray and green lines) and observed annual storage from groundwater levels (black circles). Right column (D–F): Monthly groundwater storage from the soil moisture + irrigation (SM + IRR) and soil moisture deficit (SMD) methods (red and blue lines) relative to observed storage values (black circles). Observed groundwater storage in 2003 was estimated as saturated thickness multiplied by the storage coefficient and the area of each region. Coefficients were optimized for from 2003 to 2012, as data for 2013 were not yet available. Vertical axes differ by region. Since limited data were available for the Southern High Plains (SHP) and Central High Plains (CHP) in 2010, the observed $d\text{GWS}/dt$ was set to zero.]
The similarity of the estimated change in groundwater storage in the SM + IRR and SMD methods across the HPA (see Table A1) suggests that actual irrigation or PET – ET rates can be derived by simultaneously solving equations (2) and (3). By knowing the variable with the lowest error, estimates of the variable with the highest error could be improved. In cases where in situ observed irrigation volume is less certain than data assimilation products, new estimates of IRR could be calculated as (PET – ET) × C2 / C1. In cases where data assimilation error is higher than the error associated with irrigation, the ET deficit (PET – ET) could be calculated as I × C1 / C2. The best match between estimated groundwater storage changes and those that were observed was provided using the simplest approach, with a correction based on a constant fraction of groundwater pumping (Figure 2, green lines). Accounting for irrigation in addition to NLDAS anomalies (Figure 2, blue and red lines) also provided good estimates. Since the SMD method does not require pumping data, it is likely applicable for a much broader region where such data are not available. However, the simple IRR method proved to be the best for estimating groundwater storage changes, and it complements other GRACE-based studies that estimate surface water consumption of irrigated croplands from observed dGWS/dt [Anderson et al., 2012].

Our results provide good comparisons with observed water level changes although the coefficient is variable across different climatic periods (see supporting information Table A1). Seasonal variability in evaporative demand and/or soil moisture conditions cause monthly fluctuations; however, a more parsimonious constant coefficient model provided here showed good representations of water level changes even during the recent drought condition in 2012. Observed groundwater pumping in the HPA range from 21 to 25 km³/year (see supporting information, Figure A2 and Table A2). Our results indicate that the observed and estimated groundwater depletion volumes are much lower than the volumes pumped from the aquifer, supporting previous evidence that a significant amount of the irrigation water returns to the aquifer via recharge in the NHP [Scanlon et al., 2012]. There is significant uncertainty in the annual pumping estimates used here as these are derived from direct [Frenzel, 1985] and indirect observations [Stanton et al., 2011]. Uncertainty in groundwater storage decline (Table A6) varies between 23 mm for the IRR method and ~40 mm for the SM + IRR and SMD methods.

The conceptual framework of groundwater depletion presented here allows recharge to be estimated where irrigation rates and GRACE/NLDAS estimated dGWS/dt values are available. For instance, by assuming recharge as the difference between the estimates of dGWS/dt and the observed pumped volume for the three HPA regions, average annual estimates of recharge (both natural and irrigation return flow recharge) were found to be 34.5 ± 5.2, 30.6 ± 2.5, and 62.0 ± 4.2 mm/yr for the SHP, CHP and NHP. This likely underestimates recharge in the NHP because groundwater discharge to surface water is a significant portion of the regional water balance. Independent ground-based estimates of recharge are in a similar range despite the uncertainty that characterizes estimates of groundwater recharge in the HPA (see Table A4, supporting information).

The average estimated groundwater level decline from 2003 to 2010 across the HPA of 27.6 ± 3.3 mm/yr based on the IRR method is similar to the estimated rate of 20 ± 4 mm/yr for California’s Central Valley for the same period, including an extreme drought in 2007–2009 [Famiglietti et al., 2011]. Other prominent aquifer systems worldwide have shown significantly larger declines, such as the Indus Plains aquifer system in northern India where storage declined by 40 ± 10 mm/yr during normal conditions from 2002 to 2008 [Rodell et al., 2009] and the Tigris-Euphrates region which experienced declines of 34 ± 10 mm/yr [Voss et al., 2013]. Our estimates, using corrections for soil moisture deficit and enhanced evapotranspiration in irrigated areas, suggest that the amount of groundwater depleted over the last decade represents 23% to 38% of the total depletion of the HPA since large-scale aquifer development in the 1950s. These values would have been significantly larger had groundwater levels in the NHP not recovered from 2008 to 2011, which was partly due to an increase in growing season precipitation during those years. The IRR method indicates that groundwater storage in the HPA declined 29.1 ± 2.0 km³/yr from 2010 to 2013. The IRR method shows a significant impact of the recent drought on water levels, while none of the other methods showed such a significant impact. Figure 2 illustrates the greater skill of the IRR method at capturing the drought impact. The estimated decline rates during the recent drought represent almost one third of the estimated total annual groundwater depletion volume in the United States for the year 2000 [Wada et al., 2012].

This study demonstrates that groundwater monitoring in the High Plains aquifer using the GRACE satellite can be improved by accounting for the impact of irrigation on soil moisture. The three irrigation correction
methods developed here provided reasonable estimates of mean annual recharge for the three main regions of the HPA. Future versions of NLDAS and other land surface assimilation products would benefit by incorporating irrigation volumes in the analysis where such data are available since it is clearly an important factor to consider when estimating groundwater storage changes. While most regions lack good irrigation/pumping data, it appears to be possible to estimate pumping rates based on crop water demands [Famiglietti et al., 2011] as well as regional knowledge of irrigation practices or data assimilation of evapotranspiration [Sun et al., 2012]. In regions with high resolution observations of pumping and/or irrigation, the temporal and spatial heterogeneity of processes that influence the groundwater balance such as recharge, evapotranspiration, and return flow could be inferred.

To transfer the methods presented here to other sites requires sufficient groundwater level measurements to interpolate water levels through time and optionally irrigation data to estimate the coefficients at some point during GRACE data availability. We did not relate the calibrated coefficients to other landscape and climatic factors that vary by region, which would be necessary to estimate the coefficients in regions without groundwater level data. To date, GRACE studies have only been validated to extensive water level data in the High Plains aquifer [Strassberg et al., 2009] and the Mississippi river basin [Döll et al., 2012]. Recent global initiatives including the International Groundwater Resources Assessment Centre are beginning to resolve this problem by sharing observed groundwater data from the USA, France, South Africa, and Brazil, among others.

Acknowledgments

This work was supported by the National Science Foundation (NSF) Grant EAR-1039180. Any opinions, findings, and conclusions or recommendations expressed are those of the authors and do not necessarily reflect the views of the NSF. GRACE data were processed by S. Swenson, supported by the NASA MEaSUREs Program. We thank Jim Butler, Bridget Scanlon, and M. Bayani Cardenas for their insightful comments, Jim Butler, Bridget Scanlon, and M. Bayani Cardenas for their insightful comments, and the Kansas Geological Survey for processing groundwater level data.

The Editor thanks Bridget Scanlon for her assistance in evaluating this manuscript.

References


