Effects of climate and irrigation on GRACE-based estimates of water storage changes in major US aquifers

Bridget R Scanlon1, Ashraf Rateb1, Donald R Pool1, Ward Sanford1, Himanshu Save4, Alexander Sun1, Di Long5 and Brian Fuchs6

1 Bureau of Economic Geology, Jackson School of Geosciences, University of Texas at Austin, Austin, TX, United States of America
2 U.S. Geological Survey, Arizona Water Science Center, Tucson, AZ, United States of America
3 U.S. Geological Survey, Eastern Region, Reston, VA, United States of America
4 Center for Space Research, University of Texas at Austin, Austin, TX, United States of America
5 Department of Hydraulic Engineering, Tsinghua University, Beijing, People’s Republic of China
6 National Drought Mitigation Center, University of Nebraska, Lincoln, NE, United States of America

E-mail: bridget.scanlon@beg.utexas.edu

Keywords: GRACE satellite data, US major aquifers, sustainable water management, drought, irrigation, climate extremes, groundwater depletion

Abstract

Understanding climate and human impacts on water storage is critical for sustainable water-resources management. Here we assessed climate and human drivers of total water storage (TWS) variability from Gravity Recovery and Climate Experiment (GRACE) satellites compared with drought severity and irrigation water use in 14 major aquifers in the United States. Results show that long-term variability in TWS tracked by GRACE satellites is dominated by interannual variability in most of the 14 major US aquifers. Low TWS trends in the humid eastern U.S. are linked to low drought intensity. Although irrigation pumpage in the humid Mississippi Embayment aquifer exceeded that in the semi-arid California Central Valley, a surprising lack of TWS depletion in the Mississippi Embayment aquifer is attributed to extensive streamflow capture. Marked storage depletion in the semi-arid southwestern Central Valley and south-central High Plains totaled ~90 km³, about three times greater than the capacity of Lake Mead, the largest U.S. reservoir. Depletion in the Central Valley was driven by long-term droughts (≤5 yr) amplified by switching from mostly surface water to groundwater irrigation. Low or slightly rising TWS trends in the northwestern (Columbia and Snake Basins) US are attributed to dampening drought impacts by mostly surface water irrigation. GRACE satellite data highlight synergies between climate and irrigation, resulting in little impact on TWS in the humid east, amplified TWS depletion in the semi-arid southwest and southcentral US, and dampened TWS depletion in the northwest and north central US. Sustainable groundwater management benefits from conjunctive use of surface water and groundwater, inefficient surface water irrigation promoting groundwater recharge, efficient groundwater irrigation minimizing depletion, and increasing managed aquifer recharge. This study has important implications for sustainable water development in many regions globally.

1. Introduction

Water sustainability is a critical issue globally because of the importance of water security for humans and ecosystems [1, 2]. Water sustainability is strongly linked to human water use and climate extremes (droughts and floods) [3, 4]. Water-related disasters have impacted ~4.2 billion people since 1992, representing the most economically destructive (~1.3 trillion US dollars, 63% of all damages) of all natural disasters [5]. Unsustainable groundwater (GW) development with human GW use exceeding recharge rates occurs in regions with ~1.7 billion people [3]. One of the overarching goals of the United Nations is ‘Securing Sustainable Water for All’ with particular emphasis on sustainable GW development [3, 6]. Irrigation is the dominant water user globally, accounting for ~70% of water withdrawal and ~90%
of water consumption, with heavy reliance on GW in (semi)arid regions [7, 8].

The sustainability of water resources can be evaluated using monitored or modeled water fluxes and/or water storage. However, monitoring data and regional modeling are limited in most regions, including parts of the US. Water storage changes reflect the balance of fluxes:

\[
\text{Water Storage Change} = \text{Input flux} - \text{Output flux}. \tag{1}
\]

Water storage declines or unsustainable development can result from decreasing input fluxes, increasing output fluxes, or both, considering natural fluxes (e.g. climate forcing) and human fluxes from various processes (e.g. irrigation).

The Gravity Recovery and Climate Experiment (GRACE) satellites have revolutionized water storage monitoring at regional to global scales, providing data on vertically integrated terrestrial total water storage (TWS) changes. These composite TWS values include snow storage, surface water (SW) storage, soil moisture (SM) storage, and GW storage (GWS) [9]. Many studies emphasize GRACE-derived GWS variability, which requires separate estimation of storage changes in the other components. Previous studies have related changes in GRACE TWS to climate variability/change, GW use, or both in 34 regions globally that have large trends [9]. GRACE data have been used to delineate GW depletion in the North China Plain [10], NW India [11, 12], and US aquifers [13, 14]. Hydraulic connections between SW and GW in many regions underscore the importance of managing both of these conjunctively to maintain streamflow for aquatic ecosystems [15, 16].

A number of factors can contribute to water storage changes:

(a) climate (arid versus humid) and climate extremes (droughts and floods), and

(b) human intervention through water use (often dominated by irrigation), source of water use (SW, GW), and surface reservoir management.

Climate controls hydrologic systems with GW discharging to SW in humid regions, whereas SW often recharges GW in arid regions. Climate extremes are generally more prevalent in arid areas with many longer-term droughts than in humid regions. If there is no human intervention in a region, then any water storage changes should reflect climate variability. Climate can impact water storage changes directly through fluxes (precipitation (P), evapotranspiration (ET), runoff (\(R_{df}\)), surface infiltration (I), and GW recharge (GWR)) or indirectly through changes in human water use in response to climate extremes. Floods generally increase (droughts generally decrease) storage (SW, SM, and GWS) through increased (decreased) fluxes (P, \(R_{df}\), I, and GWR), although these generalizations may not apply everywhere.

Indirect linkages between climate and GW storage can occur through human water use, particularly when climate extremes lead to changes in irrigation water demand or sources of irrigation (SW or GW), resulting in amplification or dampening of direct climate-driven storage changes. Irrigation is the dominant water use in the US, accounting for 63% of freshwater use (excluding power generation) [17]. The source of irrigation water, GW or SW, is critical because GW irrigation amplifies drought impacts by reducing GWS, whereas SW irrigation generally dampens drought impacts by increasing GW recharge from irrigation return flows. However, this is not necessarily the case from a producer’s perspective, as access to GW can mitigate drought impacts for producers. Humans can also affect TWS variability by constructing and managing surface reservoirs and through managed aquifer recharge (MAR). Difficulties in attributing causes of water storage changes arise in regions where both climate extremes and human intervention are prevalent. A recent study emphasized the importance of climate on GWS trends in the US with GW use contributing less than 25% to GWS trends [18]. Synergies between climate extremes and human water use have been recognized in some recent studies emphasizing the importance of climate-driven human water use [19, 20].

The objective of this study was to address the following questions:

(a) What are the factors controlling TWS variability as estimated from GRACE data? Natural climate variability, human intervention approximated by irrigation water use, or both?

(b) How can we use insights from GRACE data to better inform sustainable management of water resources, particularly GW resources?

A flow chart describes the data sources and approaches used (figure 1). This is the first detailed analysis of linkages between GRACE TWS variability, climate, and irrigation water use for 14 major aquifers throughout the US (figures 1, 2 and S1). Novel aspects of this study include:

(a) detailed analysis of the severity of TWS trends in the major US aquifers using different metrics;

(b) comparison of TWS variability to climate variability based on precipitation and US Drought Monitor [USDM] data;

(c) in-depth evaluation of impacts of irrigation water use and source (SW and GW) on TWS variability during dry and wet climate cycles; and

(d) extension of water storage records over several decades using output from regional GW models.
In addition, this study leverages a recent study that assessed the reliability of GRACE-derived GWS variability through detailed comparisons with GW-level monitoring data and regional and global models in major US aquifers [21]. The results of this companion study show that TWS and GWS time series plot very close to each other for most aquifers, indicating that GWS is the dominant contributor to long-term variability in TWS in most systems with limited contribution from snow, SM, and reservoir storage, except Powell and Mead reservoirs in Arizona. There was good correspondence between GWS trends from GRACE and those from regional models for most aquifers with the exception of the Mississippi Embayment aquifer. This companion study forms the foundation for the current study, which focuses on the causes of long-term TWS variability, emphasizing climate variability and human water use, focusing on irrigation. While we focus on the current climate in this analysis, we recognize the importance of climate change with megadroughts projected for the South-west and High Plains regions of the US in the latter half of the 21st century [22, 23]. We examined various approaches to more sustainable water management, particularly GW management, based on insights from GRACE data with implications for critically stressed aquifers globally.

2. Materials and methods

We selected 14 major aquifers throughout the US that are generally intensively monitored and modeled by the US Geological Survey (figure 2). These aquifers are described in supporting information (SI), section 1 (available online at stacks.iop.org/ERL/16/094009/mmedia).

2.1. Water storage from GRACE satellite data

GRACE satellites monitor TWS variability at continental to global scales. TWS variability in this study is based on GRACE data Release 06 from the University of Texas Center for Space Research mascons (CSR-M) solutions and NASA Jet Propulsion Lab mascons (JPL-M) solutions. The data are based on the original GRACE mission extending from April 2002 through June 2017 (15.25 yr). More details are provided in SI, section 2.

Water storage changes reflect the balance of fluxes at the land surface in regional models, as follows:

\[ P + \text{ Irrig.} + Q_{\text{on}} - ET - Q_{\text{off}} - \text{GWP} = \Delta \text{TWS} \]  

(2)

where \( P \) is precipitation, Irrig. is irrigation return flow, \( Q_{\text{on}} \) and \( Q_{\text{off}} \) represent surface and subsurface (GW) flow into and out of the system, respectively, ET is evapotranspiration, GWP is total GW pumpage, including irrigation, and \( \Delta \text{TWS} \) is the change in TWS [25].

Raw time series of \( \Delta \text{TWS} \) from GRACE (TWSAraw) was disaggregated into long-term (linear trend + interannual variability), annual, and residual (mostly sub-annual) variability using Seasonal Trend decomposition using Loess (STL) (SI, section 2.1) [26]:

\[ \text{TWSA}_{\text{Raw}} = \text{TWSA}_{\text{Long-term}} + \text{TWSA}_{\text{Annual}} + \text{TWSA}_{\text{Residual}} \cdot \]  

(3)

Linear trends were fit to the long-term variability (TWSA_{long-term}) using a nonparametric regression tool (e.g. Sen slope) [27] and the remaining long-term signal reflects interannual variability (equation (4)):

\[ \text{TWSA}_{\text{Long-term}} = \text{TWSA}_{\text{Linear trend}} + \text{TWSA}_{\text{Interannual}} \cdot \]  

(4)

This study focuses on long-term (trend + interannual) variability in TWS based on the ensemble mean.
Figure 2. US map shows apparent trends in total water storage (TWS) based on the mean of CSR and JPL mascon solutions for 14 major aquifers represented at total volume change over the 15.25 yr GRACE period (April 2002–June 2017) (table 1). Hachures are used to identify aquifers with reliable trends (i.e. greatly exceeding interannual variability, San Joaquin/Tulare, Central and Southern High Plains, and Columbia Plateau. Aquifer outlines are based on the Groundwater Atlas of the United States [24]. Time series: upper panels: long-term variability (trend + interannual variability) in TWS (black line) based on mean of CSR and JPL mascons with apparent linear trend shown as a dashed line. Uncertainties in TWSA are shown in figure S2. Gray bars represent the annual cumulative precipitation anomaly based on 2002–2017 period (defined in SI, section 3.1). Lower panels: percentage of aquifer areas under drought based on the US Drought Monitor (USDM) classes: D0: abnormally dry; drought categories: D1: moderate; D2: severe; D3: extreme; and D4: exceptional. Additional drought indices are shown in figure S5. The GRACE data are provided in table S7, CPA data in table S10, and drought data in table S12.
of CSR-M and JPL-M solutions (figure 2, tables 1 and S1(c)).

While many GRACE studies emphasize linear trends, the 15yr GRACE record is relatively short for estimating trends [28]. We assessed the robustness of the GRACE apparent trends against current and historical natural variability at interannual and multi-decadal scales using two metrics: (a) the goodness of fit of the linear trend (coefficient of determination of the TWS trend, $R^2$) relative to current interannual variability and (b) the severity of the current TWS trend relative to historical natural multidecadal variability (1901–2014) (trend to interannual variability ratio, TIVR) [29, 30]. The TIVR was calculated from the GRACE TWS annual trend multiplied by the GRACE period (15.25 yr) and divided by the standard deviation (SD) of the reconstructed climate-driven TWS variability (1901–2014) using precipitation and temperature forcing data [29]. TIVRs ranging from ±2 to ±3 (i.e. trends greater than 2–3 SD of interannual variability) are considered extreme, whereas TIVRs outside ±3 are considered exceptional and very unlikely to reflect natural climate variability [30].

2.2. Climate data
Monthly precipitation data were derived from the PRISM (Parameter-elevation Regressions on Independent Slopes Model) climate data (www.prism.oregonstate.edu/). The gridded (4 km) PRISM precipitation data were aggregated to the aquifer scale using aquifer polygons. The precipitation time series was analyzed using STL, similar to the TWS variability. Anomalies were calculated by subtracting the long-term mean over the GRACE period and the anomalies were accumulated between 2002 and 2017 to determine the cumulative precipitation anomaly (CPA) (SI, section 3). Drought data were derived from the USDM using aquifers polygons with details in SI section 3. TWS interannual variability was compared with CPA and USDM data using Pearson correlation.

2.3. Regional groundwater models
Regional GW models were used to assess GWS variability over much longer time periods than the GRACE record to put the GRACE data within a longer-term context. The water balance applied to aquifers is as follows:

$$R - D - GWP = \Delta GWS$$

where $R$ is GW recharge, $D$ is GW discharge to streams, springs, and ET, and $\Delta GWS$ is change in GWS [32]. Regional GW models have been developed for seven of the 14 major aquifers, including the California Central Valley Hydrologic Model (CVHM) [33, 34], Columbia Plateau Regional Aquifer System (CPRAS) [35], Eastern Snake River Plain Aquifer [36], Northern High Plains [37], Southern High Plains [38], Mississippi Embayment Regional Aquifer System (MERAS 2.1) [39], and a portion of the Texas Gulf Coast, the Houston Area Groundwater Model (HAGM) [40] (SI, section 4). These comparisons allow us to further evaluate the persistence of GRACE-derived TWS trends.

2.4. Irrigation wateruse
Irrigation water use data were obtained from the US Geological Survey (USGS) National Water-Use Science Project (NWUSP) that compiles data on water use (withdrawals) for different sectors by county in the US, every 5 yr since 1985 (table S15). Additional details are provided in SI, section 5.

3. Results and discussion
The main findings from this study are as follows with more details provided in later sections. Long-term variability (trend + interannual variability) in GRACE-derived TWS is the focus of this study and is dominated by interannual variability in most aquifers (figures 2 and 3; table 1). Linear trends are generally within ±2–3 SDs of reconstructed interannual TWS variability (1900–2014) in 10 out of the 14 aquifers (TIVRs $\leq$ 2–3), indicating that the calculated apparent trends in most of these aquifers reflect natural interannual variability and are unlikely to persist into the future (table 1, figure 3). Linear trends do not provide a good fit for TWS variability in many aquifers (low $R^2$ values) (table 1).

There are distinct differences in TWS variability between the humid east and semi-arid west with the 98th meridian often considered the boundary between these regions (figure 3) [41]. The humid east is characterized by large interannual TWS variability with stable or slightly increasing apparent linear trends related to low drought intensities (Accumulated Drought Severity and Coverage Index, ADSCI, mostly 37–64; table 1, figures 2 and SS) and variable GW irrigation pumping that can capture extensive surface water networks, such as in the Mississippi Embayment aquifer [32, 42]. The semi-arid western US has varying TWS trends between the southwest and northwest US. Large decreasing trends were restricted to the semi-arid southwest and south-central US, markedly exceeding ±2–3 SDs of interannual variability (TIVRs: $\sim$3.5 to $\sim$5.6; figure 3) and moderate to high TWS trend $R^2$ values ($\sim$0.5–$\sim$0.7) (table 1, figure 2). Impacts of intense droughts in the southwest (up to 5 yr long) were amplified by switching from mostly SW irrigation (wet periods) to increased GW irrigation (drought) in the Central Valley resulting in long-term net declines in TWS, consistent with regional modeling studies (figure 4) [33]. In contrast, apparent trends in TWS in the remaining aquifers in the northwest and north-central US
Table 1. Mean annual precipitation (P) in mm yr\(^{-1}\) (1980–2018, PRISM) for 14 major US aquifers. Results for the entire Central Valley as a unit are shown to facilitate comparison with previous studies. ADSCI is Accumulated Drought Severity and Coverage Index (April 2002–June 2017, 15.25) (SI, section 3.2). Trends in total water storage (TWS) (mm yr\(^{-1}\) and km\(^3\) yr\(^{-1}\), 2002–2017) and trend \(R^2\) value. Uncertainty in TWS trends is based on the standard deviation (SD) of five GRACE solutions (mascons and spherical harmonics). TWS trends were statistically insignificant in the Upper Colorado and Mississippi Embayment aquifers (shown by *). SD interannual TWS is SD of interannual TWS for the GRACE period (2002–2017 based on mean of CSR and JPL mascons) and of the reconstructed data (mean of JPL and GSFC mascons, 1901–2014). TIVR is the Trend (mean CSR-M and JPL-M) to interannual variability ratio, based on SD of reconstructed interannual variability (JPL-M and GSFC, 1901–2014). Correlation coefficient (\(R\)) based on Best and Roberts [31] was estimated between interannual TWS variability and cumulative precipitation anomaly (CPA, defined in SI, section 3.1), and between interannual TWS variability and the US Drought Monitor (USDM: D0 through D4). Detailed comparison of TWSA and different USDM drought categories are included in table S13. Statistically insignificant correlations are bolded based on \(p\) value <5% and a 99% confidence interval. Irrig. is the total irrigation from surface water and groundwater derived from US Geological Survey county data in 2010 and 2015 [17]. Supporting data are provided in supporting information. More detailed information is provided in table S1 and in SI.

| ID | Aquifer                          | P (1980–2018) (mm yr\(^{-1}\)) | ADSCI | TWS trend (2002–2017) (mm yr\(^{-1}\)) | SD interannual TWS (mm) | TIVR | TWSA vs CPA \(R\) | TWSA vs USDM \(R\) | Irrig. 2010 (km\(^3\) yr\(^{-1}\)) | Irrig. 2015 (km\(^3\) yr\(^{-1}\)) |
|----|----------------------------------|---------------------------------|-------|--------------------------------------|--------------------------|------|-------------------|-------------------|--------------------------|--------------------------|
| 1  | Columbia Plateau                 | 431                             | 91    | 5.6 ± 0.8                            | 9.7 ± 1.3                | 0.48 | 25.04             | 3.49               | 0.16                     | 0.50                     | 5.49                     | 4.45                     |
| 2  | Snake River Plain                | 595                             | 143   | 1.5 ± 0.2                            | 4.4 ± 0.7                | 0.06 | 25.74             | 3.28               | 0.72                     | 0.19                     | 18.42                    | 18.76                    |
| 3  | Central Valley                   | 375                             | 174   | −1.29 ± 2.9                          | −30.5 ± 7                | 0.41 | 71.12             | 50.36              | −3.90                    | 0.72                     | 25.96                    | 20.49                    |
| 4  | Sacramento                        | 918                             | 155   | −7.8 ± 0.3                           | −8.6 ± 0.4               | 0.22 | 67.37             | 49.21              | −2.43                    | 0.74                     | 7.69                     | 7.66                     |
| 5  | San Joaquin + Tulare             | 549                             | 190   | −17.1 ± 1.0                          | −22.1 ± 1                | 0.55 | 67.12             | 46.93              | −5.55                    | 0.71                     | 18.27                    | 12.83                    |
| 6  | U Colorado                        | 348                             | 172   | −0.04 ± 0.3*                         | −0.2 ± 0.4               | 0.01 | 19.11             | 18.61              | −0.03                    | 0.83                     | 6.59                     | 9.4                      |
| 7  | Arizona Alluvial                  | 315                             | 190   | −5.2 ± 1.4                           | −18 ± 5                  | 0.50 | 18.94             | 22.52              | −3.51                    | 0.50                     | 6.41                     | 6.35                     |
| 8  | N High Plains                     | 544                             | 140   | −5.7 ± 0.3                           | −22 ± 1.3                | 0.30 | 41.67             | 59.86              | 1.46                     | 0.71                     | 10.45                    | 11.21                    |
| 9  | C + S High Plains                | 506                             | 150   | −1.29 ± 0.3                           | −4.0 ± 1.0               | 0.74 | 35.58             | 37.24              | −5.29                    | 0.88                     | 10.17                    | 9.22                     |
| 10 | Edwards Trinity                   | 520                             | 142   | −4.7 ± 0.2                           | −8.4 ± 0.8               | 0.33 | 29.98             | 24.74              | −2.92                    | 0.71                     | 0.62                     | 0.52                     |
| 11 | Texas Gulf Coast                  | 983                             | 124   | −40 ± 0.5                            | −14 ± 1.7                | 0.03 | 61.38             | 30.48              | −2.00                    | 0.61                     | 2.49                     | 1.62                     |
| 12 | Mississ. Embay.                   | 1370                            | 63    | −1.0 ± 0.0*                          | −3.1 ± 0.1               | 0.01 | 52.82             | 36.64              | −0.41                    | 0.67                     | 16.05                    | 19.41                    |
| 13 | Coastal Lowlands                  | 1590                            | 64    | 2.1 ± 0.7                            | 4.7 ± 1.5                | 0.02 | 46.8              | 21.41              | 1.52                     | 0.49                     | 1.04                     | 1.24                     |
| 14 | Floridian aquifer                 | 1297                            | 74    | 3.4 ± 1.5                            | 11.3 ± 4.8               | 0.09 | 42.58             | 25.77              | 2.04                     | 0.27                     | 4.15                     | 3.61                     |
| 15 | Pennsylvania                      | 1215                            | 37    | 2.1 ± 0.3                            | 5.1 ± 0.73               | 0.03 | 27.77             | 18.75              | 1.69                     | 0.51                     | 0.03                     | 0.04                     |
Figure 3. Map of the ratio of total water storage (TWS) trends (15 yr; 2002–2017) to interannual variability in TWS (trend to interannual variability). This map is for demonstration purposes and is based on the gridded JPL-M total trends (TWS trend (mm yr$^{-1}$ × 15.25 yr) (figure S3). The interannual variability is represented by the standard deviation (SD) of the reconstructed TWS interannual variability (114 yr; 1901–2018) based on JPL mascons (resampled at 0.5° grid scale) using precipitation and temperature data (figure S4) [29]. The 98th meridian represents the boundary between the humid east and the semi-arid west. Apparent TWS trends exceeding ±2–3 SD of interannual variability are considered persistent and are found in the southwestern and southcentral US. The outlines of the major aquifers are also shown. The TIVR data in tables 1 and S1 are based on trends in TWS from CSR-M and JPL-M and the reconstructed interannual variability is based on JPL-M and GSFC.

are mostly within ±2–3 SDs of interannual variability and apparent trend $R^2$ values are low (mostly 0.01–0.3), resulting in temporary trends that reflect natural interannual variability that are less likely to persist in the future (table 1). Impacts of lower drought intensities in the NW were dampened by more widespread SW irrigation and recent MAR resulting in limited TWS trends, mostly within ±2–3 SDs of interannual variability (table 1). In summary, GW use in humid regions is more sustainable than in arid regions in general, and GW sustainability can be enhanced in arid regions through conjunctive use of SW and GW and MAR.

3.1. Relationship between GRACE total water storage, climate, and irrigation

3.1.1. Humid eastern US

Aquifers in the humid eastern US (east of the 98th meridian) are generally characterized by moderate to high interannual variability (SD: ∼30–60 mm; table 1, figure S4). Apparent linear trends in TWS are close to 0 (figure 2), with $R^2$ values mostly <0.1 and TIVRs <2–3 (figure 3), indicating that these trends reflect natural interannual variability. Drought intensities are generally low (ADSCI: 37–74) but higher in Texas aquifers (ADSCI: 124–142) (figure S5).

Human intervention is variable in these humid regions (figure 4). Irrigation water use in the Mississippi Embayment was high, surprisingly similar to or 50% greater than that in the California Central Valley, mostly sourced from GW (∼84%–88%). This level of irrigation and GW source would be expected to greatly reduce TWS, as suggested by regional GW models (∼−120 km$^3$ over the 15 year GRACE period) [39]. Lack of irrigation impact on GRACE TWS trends may result from 90% of GW irrigation pumpage being derived from the shallow Mississippi River Valley alluvial aquifer [43] that is likely well connected to a dense stream network in this humid region (SI, section 6). While storage depletion may have occurred prior to GRACE monitoring (from ∼1980s on), the shallow aquifer may have reached a quasi-equilibrium status with irrigation pumpage linked to water capture (SW, ET) rather than storage depletion. Preliminary results from the new regional model suggest up to $10 \times$ less GW depletion, likely linked to increased recharge and stream capture than the earlier Mississippi Embayment model [44]. The new model has a much denser stream network (∼1000s of streams) relative to the original model (43 of the largest streams) (SI, section 6). Irrigation water use is likely derived from stream baseflow, induced stream recharge,
3.1.2. Western US

Semi-arid regions west of the 98th meridian include a number of aquifers with varying climate forcing and irrigation amounts and sources. Many aquifers receive large water inputs from outside of the aquifers (Central Valley, Arizona Basin and Range, Columbia Plateau, and Snake River Plain), whereas groundwater in the High Plains aquifer is derived primarily from precipitation over the plain. As a result, TWS variability differs markedly between the southern and northern aquifers in the semi-arid western US (figure 2).

3.1.3. Southwest and south-central US

TWS depletion in the major aquifers in the southwest and south-central U.S. (Central Valley, Arizona, and High Plains) totaled almost 90 km$^3$ over the 15 yr GRACE record, strongly linked to long-term droughts and changes from SW to GW irrigation in the Central Valley during drought and GW irrigation pumpage greatly exceeding recharge in the Central and Southern High Plains.

The Central Valley shows a net TWS trend of $\sim-30$ km$^3$ (2002–2017), similar to the capacity of Lake Mead (32 km$^3$), the largest US surface reservoir (table 1, figure 2). Intrannual variability is high in both the northern Sacramento and southern San Joaquin/Tulare basins (SD: 67 mm in both) and is highly correlated with the USDM (D0–D4: $R=\sim 0.9$) and with the cumulative precipitation anomaly (CPA) ($R=\sim 0.7$ (table 1). The TWS trend in the southern San Joaquin/Tulare Basin ($\sim-22$ km$^3$ 15 yr$^{-1}$) exceeds natural variability and is likely to persist ($R^2=0.55$; TIVR, $-5.6$), whereas the trend in the northern Sacramento Basin ($\sim-9$ km$^3$ 15 yr$^{-1}$) reflects interannual variability and is less likely to persist.

Figure 4. US irrigation lands derived by merging Moderate Resolution Imaging Spectroradiometer (MODIS) Irrigated Agriculture Dataset (MIrAD-US) [45]. Total Water Storage for 14 major aquifers over the 15 yr GRACE period (2002–2017). Histograms show the volume of water use in 2010 and 2015 summed over counties outlined in aquifers (table S15). The dominant water use is irrigation, including irrigation source (surface water and groundwater) and other water uses (livestock, domestic, industrial, mining, public supply and aquaculture for surface water and groundwater). Thermoelectric power was excluded. Water use data are provided in table 1 and details of irrigation water sources and use are provided in table S16. Central Valley includes Sacramento and San Joaquin/Tulare Basins. TX ETP: Texas Edwards Trinity Plateau; TX Gulf Coast: Texas Gulf Coast. Annual water use data are available for Kansas (Central High Plains), Southern High Plains, TX ETP, and TX Gulf Coast (figure S7, table S16).
Declines in the central and southern High Plains were driven primarily by GW pumpage because there is almost no SW available for irrigation. Low recharge rates in these aquifers preclude direct connections between climate and TWS variability [50]. Interannual variability in TWS change is attributed to indirect linkages between climate and TWS variability through variations in irrigation water demand and GW pumping linked to climate variability, as shown in earlier studies in the central High Plains [51]. Previous studies show that GW depletion exceeded recharge by ~10× in the Central High Plains [52]. TWS declines in Texas aquifers south of the High Plains (Edwards Trinity and Gulf Coast aquifers) were moderate, partly because irrigation was only ~5%–25% of that in the High Plains aquifer. Trends in these Texas aquifers are within ±2–3 SD of interannual variability (TIVR < 2–3) and low TWS trend $R^2$ values (0.03–0.3) indicate predominantly natural interannual variability (table 1). The high correlation between TWS variability and drought ($R = 0.62–0.79$) indicates that climate is the primary driver of TWS variability in these Texas aquifers.

### 3.1.4. Northwest and north-central US

In contrast to declining TWS trends in the south-west and south-central US, apparent TWS trends were stable or slightly rising in the northwest (Snake River Plain, 4.4 km$^3$; Columbia Plateau, 9.7 km$^3$, 2002–2017) (figure 2). The TIVR is 3.5 in the humid Columbia Plateau but is only 0.7 in the Snake River Plain, indicating the latter may reflect interannual variability mostly (table S1). Drought intensity was low in the Columbia (ADSCI, 91) and higher in the Snake River Plain (143), less severe than in the south-west US. Correlation between TWS and precipitation was low in both the Columbia and Snake basins ($R = ±0.2$) (table 1). Widespread flooding in 2011 may have contributed to increasing TWS [53].

SW irrigation accounted for 68%–72% of total irrigation in the Columbia Plateau and Snake River Plain (table S15(a)), likely dampening drought impacts on TWS changes because of GW recharge from flood irrigation and transmission losses along unlined canals, partially disconnecting storage changes from climate variability (figure 4, table S15). However, irrigation pumpage from deeper confined basals in the Columbia River Basalt Group is disconnected from the shallow system and would not benefit from SW irrigation. In contrast, the Snake River Plain aquifer is unconfined with strong interconnections between shallow and deeper systems.

Although drought intensity was high in the Upper Colorado (ADSCI, 172), similar to the Central Valley, irrigation was derived primarily from SW (~97%) and likely dampened drought impacts, resulting in very low apparent TWS trends, similar to interannual
variability (TIVR, \(\sim 0.03\)) with low TWS trend \(R^2\) values (0.01) (table 1, figures 2–4).

Further east, the Northern High Plains shows moderately high interannual variability (SD: 42 mm) correlated with the USDM data (\(R = 0.73\)) (table 1). The apparent TWS trend (\(\sim 22 \text{ km}^3\)) primarily reflects natural interannual variability because the trend is within \(\pm 2\) SD of interannual variability (TIVR: 1.5, figure 3) and the TWS trend \(R^2\) is low (0.3) (table 1). Climate effects on TWS variability may have been partially dampened by SW irrigation, accounting for \(\sim 20\%–30\%\) of total irrigation, with inefficient flood irrigation representing \(\sim 80\%\) of total irrigated area (figure 4, table S15), recharging aquifers adjacent to the Platte and other rivers, as shown by long-term GW level monitoring [54]. Variations in TWS trends between the Northern (22 \(\text{ km}^3\)) and Central and Southern (\(-40 \text{ km}^3\)) High Plains may be explained by shorter droughts in the north (2012–2013) relative to further south (2011–2015) and some SW irrigation and sandier soils in the Northern High Plains (e.g. Nebraska Sand Hills) resulting in higher recharge and more dynamic storage response to climate variability relative to absence of SW irrigation and more clay-rich soils further south, limiting GW recharge [55].

3.2. Long-term system evolution from regional groundwater models

GRACE-derived TWS changes are restricted to the recent 15 year period; however, trends in GWS were evaluated over much longer decadal timescales using regional GW models supported by GW level monitoring for seven of the 14 US aquifers (figures 5 and S8). Results from a previous analysis show that GRACE-derived GWS variability compares favorably with regional models for many aquifers, except the Mississippi Embayment, although the overlap period of GRACE and models is limited [21].

In the Central Valley, results from the regional GW model (1963–2014) are consistent with GRACE results from this study showing drought-driven TWS changes amplified by switching from predominately SW irrigation during wet periods to increased GW irrigation pumpage during drought. Modeled GWS declined by \(\sim 15–20 \text{ km}^3\) during each short-term drought (1976–1977; 1999–2003; 2007–2009) and by \(\sim 40 \text{ km}^3\) during a 5 year drought (1987–1992), with only partial recovery during wet periods in the early 1980s and late 1990s [33, 34]. The model also shows the impact of the irrigation source shifting from up to \(\sim 70\%\) SW irrigation during wet periods to up to \(\sim 70\%\) GW irrigation during dry periods, amplifying drought impacts on GW. The model emphasizes the importance of conjunctive use of SW and GW, with pipeline development (\(\leq 1000 \text{ km}\)) transferring SW from the more humid north to the semi-arid south, resulting in GWS recoveries in some regions [33]. The importance of conjunctive SW and GW use is highlighted by recent new land subsidence linked to irrigation expansion in areas relying entirely on GW without access to SW [34].

In the Northern High Plains, the regional model shows no net change in GWS from \(\sim 1980\) to mid-2000s, similar to the GW level monitoring data (figure S8(d), table S17) [37, 54]. The importance of conjunctive use of GW and SW is evident in modeled and monitored GW-level rises from SW irrigation near rivers (e.g. Platte River) and modeled reductions in baseflow to some streams by up to 50\% [56]. The regional model of the Southern High Plains shows \(\sim 350 \text{ km}^3\) of GW depletion since the 1950s related to intensive GW irrigation greatly exceeding GW recharge rates [38]. The much greater depletion relative to the Northern High Plains is attributed to \(\sim 10 \times\) lower recharge relative to irrigation pumpage, lack of SW for irrigation and related recharge (return flow and leakage from distribution systems), and lower permeability soils in the Southern High Plains [55].

In the northwestern US aquifers, losses to the aquifer from SW irrigation increased GWS by \(\leq \sim 20 \text{ km}^3\) from \(\sim 1940\) to \(\sim 1970\) in the Columbia Plateau with \(\sim 70\%–80\%\) from inefficient flood irrigation [55] and by \(\leq \sim 20 \text{ km}^3\) from 1912–1950 in the Eastern Snake River Plain aquifer [36]. Increasing GW-based irrigation in the Snake River Plain depleted GWS from an excess of \(\sim 20 \text{ km}^3\) in the early 1950s down to \(\sim 6 \text{ km}^3\) in the mid-2010s. Recent MAR has partially replenished GWS.

Re-evaluation of the original Mississippi Embayment regional model [44] suggests that the fraction of GW pumpage derived from storage depletion may have been substantially overestimated while the amount derived from capture of stream baseflow, induced stream recharge, and ET, may have been greatly underestimated, particularly in recent decades [39, 57]. These GW/SW modeling issues are not as prevalent in other systems where interactions between GW and SW can be monitored or bounded.

3.3. Study limitations

The large regional scale output provided by GRACE is considered an advantage when conducting aquifer to continental scale water storage analyses. However, the low spatial resolution of GRACE data (\(\sim 100,000 \text{ km}^2\)) is often viewed as a limitation by hydrologists when evaluating water storage in smaller scale aquifers and river basins. This regional scale GRACE data may mask local scale variations in water storage. The coarse resolution provided by GRACE data could be partially overcome in the future by supplementing GRACE TWS changes with groundwater-based gravity monitoring that has much higher spatial resolution (\(\sim 100 \text{ m}\)), as shown in previous studies [58, 59].

This study focused on current and historical climate extremes; however, climate change is also a critical issue with projected megadroughts in the
US Southwest and Plains regions that should be addressed in future studies [22, 23]. Comparison of GRACE data with climate extremes in this study focuses primarily on droughts and benefits from the detailed data available from the US Drought Monitor. Conversely, comparable data are not available for flooding in the US. The Dartmouth Flood Observatory relies primarily on subjective reporting rather than independent monitoring data. Therefore, it is much more difficult to compare GRACE data with floods than droughts in the studied aquifers.

Irrigation water use is one of the primary drivers considered in this study; however, the data are based primarily on estimates of SW and GW use for irrigation that are provided once every 5 yr for most aquifers. These are some of the primary limitations of this study; however, they do not impact the main findings of this analysis.

3.4. Implications for sustainable water management in the US

GRACE satellites may be extremely valuable in assessing the sustainability of future water management projects designed to resolve spatial and temporal disconnects between water supplies and demands caused by climate extremes, irrigation, and SW availability.

Low regional storage changes in the humid eastern US underscore the importance of high precipitation, low to moderate drought intensities, and extensive perennial stream networks that can be captured even by GW irrigation resulting in more sustainable GW management. However, impacts of GW pumpage on streams need to be considered to maintain environmental flows for healthy ecosystems.

SW irrigation (mostly flood irrigation) has been extremely valuable in recharging GW and increasing aquifer storage in the northwest US, as shown by GW level rises during irrigation development in the early to mid-1900s, regional models, and is consistent with stable or slightly increasing TWS trends in GRACE data (figure 5) [21]. In Idaho, up to 0.5 km³ yr⁻¹ of Snake River water has been transferred to the Eastern Snake River Plain aquifer within the past few years to promote recharge in unlined canals and adjacent spreading basins (MAR) [60]. In the northern High Plains, pilot studies transporting excess SW during wet periods in unlined irrigation canals promote recharge [61]. Although MAR has been practiced in some parts of the Central Valley since the 1960s, water volumes transferred from SW to GW were low (~14 km³ from 1960 to 2013) and impacts were generally localized [48]. More recent studies have been applying flood MAR in California, capturing excess SW using irrigation infrastructure to flood cropped and fallow fields in winter to promote aquifer recharge [62]. Flood irrigation and MAR were also effective in recharging GW in Arizona Alluvial Basins sourced from the Colorado River [48].

These approaches counter the mantra of 'more crop per drop’ to maximize irrigation efficiency because the latter fails to recognize that losses from inefficient SW irrigation (mostly from flood irrigation) actually recharge GW systems and are somewhat similar to MAR. Some recent studies in California and Texas have estimated how much high magnitude streamflow (>90th–95th percentiles) could be captured to recharge depleted aquifers that would otherwise discharge to the ocean [63, 64].

Historical GW depletion provides subsurface reservoirs to complement surface reservoirs. The GRACE data show TWS depletion of ~90 km³ in the southwest and southcentral US (2002–2017; Central Valley, Arizona, and Central and Southern High Plains). Previous studies show that TWS in these systems is dominated by GW [21] and this level of depletion would provide reservoir storage almost 3× the capacity of Lake Mead. Estimated depletion of US aquifers over approximately the last century (1900–2008) from modeling and monitoring data totals ~1000 km³ [3, 65]. Not all of this depleted aquifer storage would be available as some storage is permanently lost because of aquifer compaction (e.g. ~20% in the Central Valley) [33]. However, additional subsurface storage may be available from natural deep water tables in aquifers in semi-arid regions.

3.5. Implications for other systems globally

Net increases in TWS in the northwest US from GRACE data in this study are consistent with net increases in GWS modeled by the WaterGap Global Hydrologic Model (WGHM) in this region and also in other regions globally where SW irrigation recharges GW (e.g. NW India, SE Asia) [66]. Although numerous GRACE studies delineated GW overexploitation in NW India [12], analysis of earlier data indicate that GW levels in some regions of the Indo-Gangetic Basin rose by median values of ~20–30 m from
leaky SW irrigation canals (1900s–1950s, 1960s) with recent net depletion since the 1980s ≤10 m, highlighting the importance of considering recent GRACE data within a longer-term context [67]. Future water management in NW India could move towards more sustainable development by conjunctive use of SW and GW. More recent analysis on the River Ganges in Bangladesh indicates that GW pumping may be enhancing capture of SW by inducing recharge, similar to what may be occurring in the Mississippi Embayment aquifer [68, 69]. To increase environmental flows in the Murray Darling Basin, the Australian Government spent almost 6 billion dollars on water infrastructure (e.g. lining irrigation canals and piping irrigation water) [70]. However, failure to monitor and account for irrigation return flows resulted in little improvement of river flows in the basin [70]. These limited examples of the importance of SW irrigation in sustainable development is similar to the recent expansion of MAR in many regions.

3.6. Application of findings for sustainable water management

The results of this study indicate that GRACE satellites can be extremely valuable in monitoring regional water storage changes to evaluate sustainable management approaches. GW irrigation pumpage is the primary driver of GRACE TWS declines in the central and southern High Plains. Climate is a major driver of TWS variations in many of the other aquifers, resulting in high interannual water storage variability in response to wet and dry climate cycles. GW irrigation amplifies GRACE TWS changes in response to drought (e.g. Central Valley) but SW irrigation dampens TWS changes (e.g. NW US aquifers) (figures 2 and 5). Therefore, sustainable GW management would benefit from irrigation sourced by SW because it is more renewable than GW, recognizing that inefficient SW irrigation (mostly flood irrigation) can contribute to GW recharge (e.g. Columbia Plateau Aquifer). The use of SW irrigation needs to ensure that it does not negatively impact environmental flow regimes in streams that include preservation of extreme/peak flows. Inter-basin SW transfers to semi-arid regions with limited SW increase opportunities for GW recovery and more sustainable management (e.g. US Central Valley and Arizona Alluvial Basins). Conjunctively managing SW and GW is also beneficial, optimizing their use considering floods and droughts. Inefficient SW irrigation (e.g. flood) and efficient GW irrigation (e.g. drip) should be optimal from a water storage perspective; however, energy consumption for SW pumping and potential contamination during water transfer for inefficient SW irrigation should also be considered. In the past, inefficient SW irrigation (e.g. flood irrigation) recharged aquifers unintentionally. More recently, excess SW is recharged using MAR from high magnitude stream flows, as quantified in California and Texas, in large depleted GW reservoirs that are a legacy of previous overexploitation. Systems with little or no SW, such as Central and Southern High Plains., may decrease aquifer overexploitation and extend aquifer lifespan by need maximizing irrigation efficiency and reducing pumpage.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://dataverse.tdl.org/dataset.xhtml?persistentId=doi:10.18738/T8/1JMYNL.

Acknowledgments

This work was conducted as a part of the ‘Integrating GRACE Satellite and Ground-based Estimates of Groundwater Storage Changes’ Working Group supported by the John Wesley Powell Center for Analysis and Synthesis, funded by the US Geological Survey. We are very grateful for detailed review comments from Dr Leonard Konikow (USGS) which greatly improved the manuscript. We also acknowledge the input from three anonymous reviewers. Additional support was also provided by NSF (IAA 1702174) to the Powell Center for this work and from the Jackson Endowment at the University of Texas at Austin to the senior author. B R S, A R, D P, W S, and H S participated in the Powell Center meetings. A R conducted the data analyses. B R S wrote the paper. All authors reviewed the paper. The authors declare that they have no competing interests.

ORCID iDs

Bridget R Scanlon https://orcid.org/0000-0002-1234-4199
Ashraf Rateb https://orcid.org/0000-0002-8875-1508
Alexander Sun https://orcid.org/0000-0002-6365-8526

References

[1] United Nations 2013 Water security & the global water agenda: a UN-water analytical brief United Nations Water p 47 (available at: www.unwater.org/publications/water-security-global-water-agenda/) (8 May 2013)
[2] Bernauer T and Böhmelt T 2020 International conflict and cooperation over freshwater resources Nat. Sustain. 3 350–6
[3] United Nations 2014 A post-2015 global goal for water: synthesis of key findings and recommendations from UN-water United Nations-Water p 41 (available at: www.un.org/waterforlifedecade/pdf/27_01_2014_un_water_paper_on_a_post2015_global_goal_for_water.pdf) (27 January 2014)
[4] Wada Y, Van Beek L P H, Van Kempen C M, Reckman J W T M, Vasak S and Bierkens M F P 2010 Global depletion of groundwater resources Geophys. Res. Lett. 37 L20402
[47] Melton F et al 2017 Mapping drought impacts on land following in California with Satellite Data (available at: https://nex.nasa.gov/nex/projects/1372)

[48] Scanlon B R, Reedy R C, Faunt C C, Pool D and Uhlman K 2016 Enhancing drought resilience with conjunctive use and managed aquifer recharge in California and Arizona Environ. Res. Lett. 11 035013

[49] Central Arizona Project CAP 2021 CAP water deliveries (available at: www.cap-az.com/departments/water-operations/deliveries) (accessed 15 March 2021)

[50] McMahon P B, Dennehy K F, Bruce B W, Bölhke J K, Michel R L, Gurdak J J and Hurlbut D B 2006 Storage and transit time of chemicals in thick unsaturated zones under rangeland and irrigated cropland, High Plains, United States Water Resour. Res. 42 W03413

[51] Whittemore D O, Butler J J and Wilson B B 2016 Assessing the major drivers of water-level declines: new insights into the future of heavily stressed aquifers Hydrolog. Sci. J.-J. Des. Sci. Hydrol 61 134–45

[52] Scanlon B R, Reedy R C, Gates J B and Gowda P H 2010 Impact of agroecosystems on groundwater resources in the Central High Plains, USA Agric. Ecosyst. Environ. 139 700–13

[53] Vining K C, Chase K J and Loss G R 2013 General weather conditions and precipitation contributing to the 2011 flooding in the Mississippi River and Red River of the North Basins, December 2010 through July 2011 U.S. Geological Survey Professional Paper 1798–B p 22

[54] McGuire V L 2017 Water-level and recoverable water in storage changes, High Plains aquifer, predevelopment to 2013 and 2013–15 U.S. Geological Survey Scientific Investigations Report 2017–5040 p 14

[55] Scanlon B R, Faunt C C, Longuevergne L, Reedy R C, Alley W M, McGuire V L and McMahon P B 2012 Groundwater depletion and sustainability of irrigation in the US High Plains and Central Valley Proc. Natl Acad. Sci. 109 9320–5

[56] Peterson S M, Stanton J S and Bradley J R 2008 Simulation of ground-water flow and effects of ground-water irrigation on base flow in the Elkhorn and Loup River Basins, Nebraska U.S. Geological Survey Scientific Investigations Report 2008–5243 p 65

[57] Clark B R, Hart R M and Gurdak J J 2011 Groundwater availability of the Mississippi embayment U.S. Geol. Surv. Prof. Paper 1785 p 62

[58] Scanlon B R, Zhang Z, Reedy R C, Pool D R, Save H, Long D, Chen J, Wolock D M, Conway B D and Winester D 2015 Hydrologic implications of GRACE satellite data in the Colorado River Basin Water Resour. Res. 51 9891–903

[59] Pool D R and Anderson M T 2007 Ground-water storage change and land subsidence in Tucson Basin and Avra Valley, Southeastern Arizona, 1998–2002 U.S. Geol. Surv. Sci. Inv. Rept. 2007–5275 p 34

[60] O’Connell J 2019 State exceeds aquifer recharge goal for third consecutive year. Idaho State Journal, April 1, 2019

[61] GRIPP 2020 Managed aquifer recharge (MAR) in Nebraska: using flood river flows to recharge groundwater for irrigation and enhance aquatic habitats GRIPP: Groundwater Solutions Initiative for Policy and Practice (available at: http://gripp.iwmi.org/natural-infrastructure/water-storage/managed-aquifer-recharge-mar-in-nebraska/) (accessed 19 January 2020)

[62] Dahlke H E et al 2018 Advances in Chemical Pollution, Environmental Management and Protection vol 3, ed J Friesen and I Rodriguez-Sinobas (Amsterdam: Elsevier) pp 215–75

[63] Yang Q and Scanlon B R 2018 Availability of high magnitude streamflows for potential storage to mitigate floods in Texas AGU Fall meeting 2018 Abs H21Q-1946

[64] Kocis T N and Dahlke H E 2017 Availability of high-magnitude streamflow for groundwater banking in the Central Valley, California Environ. Res. Lett. 12 084009

[65] Konikow L F 2013 Groundwater depletion in the United States (1900–2008) U.S. Geol. Surv. Scientific Investigation Report 2013–5079 p 63

[66] Doll P, Hoffmann-Dobrev H, Portmann F T, Siebert S, Eicker A, Rodell M, Strassberg G and Scanlon B R 2012 Impact of water withdrawals from groundwater and surface water on continental water storage variations J. Geodyn. 59–60 143–56

[67] MacDonald A M et al 2016 Groundwater quality and depletion in the Indo-Gangetic Basin mapped from in situ observations Nat. Geosci. 9 762

[68] Shamsudduha M, Taylor R G, Ahmed K M and Zahid A 2011 The impact of intensive groundwater abstraction on recharge to a shallow regional aquifer system: evidence from Bangladesh Hydrogeol. J. 19 901–16

[69] Nowreen S, Taylor R G, Shamsudduha M, Salehin M, Zahid A and Ahmed K M 2020 Groundwater recharge processes in an Asian mega-delta: hydrometric evidence from Bangladesh Hydrogeol. J. 28 2917–32

[70] Grafton R Q 2019 Policy review of water reform in the Murray-Darling Basin, Australia: the ‘do’s’ and ‘do’nots’ Aust. J. Agric. Resour. Econ. 63 116–41