AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE
10.1002/2016WR019494

# Global evaluation of new GRACE mascon products for hydrologic applications

**Bridget R. Scanlon[1], Zizhan Zhang[1], Himanshu Save[2], David N. Wiese[3], Felix W. Landerer[3], Di Long[4], Laurent Longuevergne[5], and Jianli Chen[2]**

[1]Bureau of Economic Geology, Jackson School of Geosciences, University of Texas at Austin, Austin, Texas, USA, [2]Center for Space Research, University of Texas at Austin, Austin, Texas, USA, [3]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA, [4]Department of Hydraulic Engineering, Tsinghua University, Beijing, China, [5]Géosciences Rennes, Université de Rennes, Campus de Beaulieu, Rennes Cedex, France

**Abstract** Recent developments in mascon (mass concentration) solutions for GRACE (Gravity Recovery and Climate Experiment) satellite data have significantly increased the spatial localization and amplitude of recovered terrestrial Total Water Storage anomalies (TWSA); however, land hydrology applications have been limited. Here we compare TWSA from April 2002 through March 2015 from (1) newly released GRACE mascons from the Center for Space Research (CSR-M) with (2) NASA JPL mascons (JPL-M), and with (3) CSR Tellus gridded spherical harmonics rescaled (sf) (CSRT-GSH.sf) in 176 river basins, ~60% of the global land area. Time series in TWSA mascons (CSR-M and JPL-M) and spherical harmonics are highly correlated (rank correlation coefficients mostly >0.9). The signal from long-term trends (up to ±20 mm/yr) is much less than that from seasonal amplitudes (up to 250 mm). Net long-term trends, summed over all 176 basins, are similar for CSR and JPL mascons (66–69 km$^3$/yr) but are lower for spherical harmonics (~14 km$^3$/yr). Long-term TWSA declines are found mostly in irrigated basins (−41 to −69 km$^3$/yr). Seasonal amplitudes agree among GRACE solutions, increasing confidence in GRACE-based seasonal fluctuations. Rescaling spherical harmonics significantly increases agreement with mascons for seasonal fluctuations, but less for long-term trends. Mascons provide advantages relative to spherical harmonics, including (1) reduced leakage from land to ocean increasing signal amplitude, and (2) application of geophysical data constraints during processing with little empirical postprocessing requirements, making it easier for nongeodetic users. Results of this product intercomparison should allow hydrologists to better select suitable GRACE solutions for hydrologic applications.

**Key Points:**
- Performance of new GRACE mascons in 176 global basins is shown to be equivalent to or superior to traditional-scaled spherical harmonics
- Net long-term trends from mascons summed over all basins agree (66–69 km$^3$/yr) but are higher than spherical harmonics (14 km$^3$/yr)
- Net depletion in total water storage occurs mostly in irrigated basins, ranging from −41 to −69 km$^3$/yr among different GRACE solutions

**Supporting Information:**
- Supporting Information S1

**Correspondence to:**
B. R. Scanlon,
bridget.scanlon@beg.utexas.edu

**Citation:**
Scanlon, B. R., Z. Zhang, H. Save, D. N. Wiese, F. W. Landerer, D. Long, L. Longuevergne, and J. Chen (2016), Global evaluation of new GRACE mascon products for hydrologic applications, *Water Resour. Res.*, *52*, 9412–9429, doi:10.1002/2016WR019494.

Received 12 JUL 2016
Accepted 7 NOV 2016
Accepted article online 15 NOV 2016
Published online 17 DEC 2016



## 1. Introduction

There is increasing interest in applying GRACE data to hydrologic problems, including evaluation of water storage changes related to droughts [*LeBlanc et al.*, 2009; *Chen et al.*, 2010a; *Long et al.*, 2013; *Brena-Naranjo et al.*, 2014; *Thomas et al.*, 2014; *Humphrey et al.*, 2016], and floods [*Chen et al.*, 2010b; *Reager and Famiglietti*, 2013] and estimation of groundwater depletion [*Strassberg et al.*, 2007; *Rodell et al.*, 2009; *Longuevergne et al.*, 2010; *Famiglietti et al.*, 2011; *Scanlon et al.*, 2012; *Feng et al.*, 2013; *Richey et al.*, 2015]. Advances in processing and increasing applications of GRACE data are seen in the rise in publications, up to 150/yr within the past few years (see Figure S1 in the Supporting Information). Understanding the different GRACE solutions is essential for assessing the advantages and limitations of various products for different hydrologic applications.

The primary objective of the GRACE satellites is to monitor the time variable gravity field which reflects changes in mass resulting primarily from movement of water through the Earth's water cycle. GRACE satellites are used to monitor spatiotemporal variability in terrestrial total water storage anomalies (TWSA), which vertically integrate water storage changes from the land surface to the deepest aquifers. While many studies estimate component storages from GRACE data, particularly groundwater storage [*Chen et al.*, 2010a; *Richey et al.*, 2015], reliable TWSA data are a necessary prerequisite for assessing component water storages. Because of their integrative nature, TWSA measurements are difficult to validate with independent data. GRACE satellites have been likened to a giant weighing scale in the sky, tracking spatiotemporal variations in water mass globally. Variations in water mass or storage are generally expressed as an equivalent water thickness (EWT; mm water). The spatial resolution of the water storage changes from the GRACE satellites is

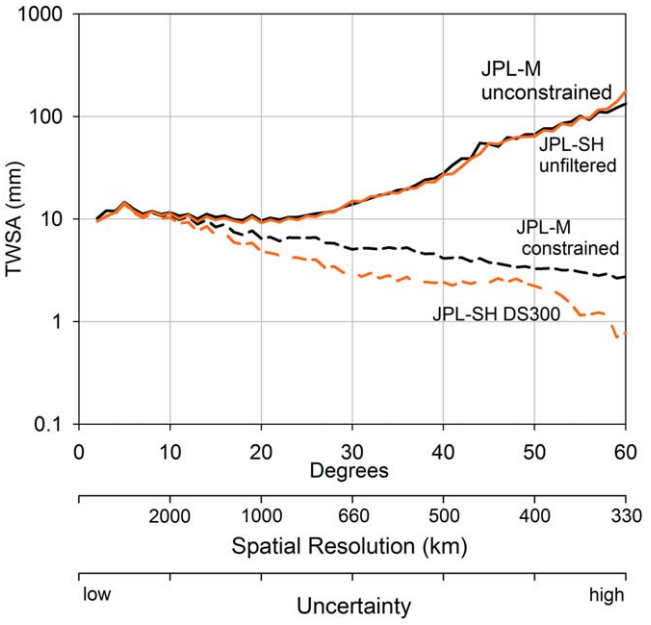


**Figure 1.** GRACE-observed total water storage anomalies (TWSA) expressed in equivalent water thickness (EWT, mm) relative to spherical harmonic degrees. The data show that the unconstrained JPL mascons (JPL-M) and unfiltered spherical harmonics (JPL-SH) are similar. The constrained mascons have higher EWT than the filtered SH solution (DS, destriped, 300GF = 300 km Gaussian filter). (Data modified from *Watkins et al.* [2015]). The spatial resolution corresponding to the SH degrees is shown, with increasing degree/order and increasing spatial resolution. The uncertainty related to the SH is also shown with low degrees corresponding to low uncertainties and high degrees corresponding to high uncertainties (noisy data).

controlled primarily by the altitude of the satellites (~450 km throughout most of record), that results in a resolution of ~100,000 km$^2$ (~300 km × 300 km). While some GRACE products are gridded at 0.5°–1° (~55–110 km at the Equator), these grids do not reflect the GRACE resolution and grid cells should be aggregated to the spatial resolution of GRACE data (~100,000 km$^2$) [*Landerer and Swenson*, 2012; *Watkins et al.*, 2015] in stand-alone applications.

GRACE satellites monitor changes in mass by measuring distance variations between a pair of satellites with 1 gigaton mass change equivalent to 1 km$^3$ of water storage [*Tapley et al.*, 2004]. Various approaches have been developed to process GRACE range-rate observations. The traditional processing approach that has been applied over the decade or more is parameterizing the Earth's gravity field using global spherical harmonics (SH) basis functions [*Wahr et al.*, 1998; *Bettadpur*, 2012]. An alternative processing approach involves parameterizing the gravity field with regional mass concentration functions ("mascons") which has become operational within the past couple of years [*Luthcke et al.*, 2015; *Watkins et al.*, 2015; *Save et al.*, 2016].

What are the similarities and differences between SH and mascon data? Spherical harmonic and unconstrained mascon data are similar, as shown in the degree variance plot (Figure 1) because they are both based on the same fundamental GRACE satellite data (level 1) and models to remove atmospheric, oceanic, and tidal signals. A basic difference between SH and mascons is that SH solutions are global whereas mascons can be applied at regional to global scales. This difference is important because we cannot distinguish land and ocean areas in global SH; therefore, the generally higher land signals leak into the lower ocean signals, reducing signal amplitudes. In contrast, land and ocean areas can be explicitly defined during mascon processing, reducing leakage errors relative to the SH solutions.

GRACE data are noisy with noise increasing at higher SH degrees, corresponding to higher spatial resolution or smaller basins (Figure 1). Therefore, GRACE data are processed to reduce noise. However, methods of reducing noise often involve losing some signal also. Therefore, the challenge is to optimize the trade-offs between noise reduction and signal loss. Noise in the GRACE data is evident in north-south stripes in the gravity field in the SH solutions (Figure S2) that result from correlated errors, attributed to the near-polar orbit with poor observability of the east-west component of the gravity gradient. Most of the early hydrologic studies using SH data include significant postprocessing of GRACE data at the basin scale to reduce noise by truncating the signal mostly to degree and order of 60, applying a destriping filter [*Swenson and Wahr*, 2006], and Gaussian smoothing filter (300–500 km radius). Various approaches were developed to restore the lost signal, the most widely applied method involving scaling factors developed by applying the same processing as in GRACE to simulated TWSA by a land surface model (LSM, truncation, and filtering) and determining the scaling factor by minimizing the differences between the unfiltered and filtered simulated TWSA [*Landerer and Swenson*, 2012; *Long et al.*, 2015].

In contrast to the SH approach, in which noise reduction and signal restoration are applied as post processing steps, the regional mascon solutions provide a framework to implement noise reduction during processing. Correlated errors can be suppressed during the gravity inversion process through regularization or stabilization of the solution, which is more optimal than removing such errors after the fact with postprocessing destriping/smoothing techniques in the SH approach. The postprocessing techniques have a tendency to attenuate signals along with suppressing the errors because these filter techniques have no information from the raw observations or the model parameters. Applying constraints or regularization in the mascon estimation and letting the data drive the estimate while suppressing the correlated errors is a more optimal way to handle these errors to minimize signal attenuation. In short, the estimates for all parameters improve in the mascon approach.

Noise is reduced during mascon processing by applying constraints. The CSR and JPL mascon solutions represent two fundamentally different approaches to applying constraints; CSR-M constraints are based on GRACE data only whereas JPL-M constraints are based on both GRACE data and geophysical models, e.g., hydrology and ocean models with loose constraints to avoid biasing the GRACE solutions toward the models [*Watkins et al*., 2015; *Save et al*., 2016]. The CSR and JPL mascon approaches also differ in terms of applying time constraints. JPL-M implements time correlation using a sequential Kalman filter with stationary process noise to account for trends in the data that are not included in the models. In contrast, CSR-M forward models secular trends derived from an intermediate regularized mascon solution, while implementing a time variable regularization matrix [*Save et al*., 2016]. While both regularization approaches are quite different, they should lead to similar estimates of gravity field variations within assessed uncertainties. The hydrologic implications of different approaches to time constraints are shown by *Save et al*. [2016] who found that fixed regularization relative to time variable regularization attenuates flood signals in some regions, such as the northern High Plains (2010 floods) and Lake Eyre (2011).

CSR and JPL mascons also differ in terms of the scale of the grid cells. CSR-M uses a geodesic equal area grid with hexagonal tiles ~120 km wide or ~1° at the Equator, resulting in ~41,000 equal area (12,400 $km^2$) hexagonal grid cells globally, similar to a soccer ball [*Save et al*., 2016]. JPL-M parameterizes the gravity field using $3^\circ \times 3^\circ$ grid cells (native resolution, ~ 330 km at the Equator) resulting in ~10 times less grid cells than CSR-M (4551 grid cells) [*Watkins et al*., 2015]. The native resolution of GRACE is ~300 km, primarily limited by the altitude of the satellites. Hence, JPL-M samples the gravity field at approximately the native resolution of the GRACE mission at the Equator, while CSR-M deliberately oversamples the gravity field at the Equator to increase the solution resolution at higher latitudes where there are more ground tracks.

One of the basic questions related to all processing approaches applied to GRACE data is how much signal is lost. *Watkins et al*. [2015] note that the scale factors for SH are a simple approximation of signal loss and can be highly variable spatially (Figure S3a). It is difficult to quantify signal loss based on the applied constraints to mascons data. *Save et al*. [2016] compute postfit residuals by fitting the GRACE range-rate data to the mascon solutions. These residuals are a part of the observed range-rate signal that is not adjusted in the mascon estimates and is discarded. If a map of these range-rate residuals has any geospatial correlation, it would suggest that the signal is attenuated during the mascon estimation process. By analyzing these postfit residuals, the CSR mascon computation process makes sure that all of the signal observed by the GRACE satellite is captured by the estimated mascon solutions by recomputing the mascon regularization and the solutions if required. Assessing postfit residuals is considered a necessary but not a sufficient condition to evaluate signal attenuation.

Previous studies evaluated the mascon solutions by comparing GRACE TWSA in the oceans to in situ ocean bottom pressure (OBP) data, noting coefficients of determination ($r^2$) between GRACE JPL and CSR mascons and in situ OBP data up to 0.35–0.85 higher relative to those with JPL and CSR-SH solutions [e.g., *Watkins et al*., 2015]. However, it is not possible to validate GRACE TWSA on land because TWSA is an integrative measurement that can only be approximated from other data sets. The increased power in the mascon solutions relative to the postprocessed SH solutions is attributed to greater signal loss in the SH processing. Analysis of TWSA trends shows greater signals from the JPL and CSR mascons than for gridded SH (GSH) data, particularly in drought regions in the United States and groundwater depleted regions in India [*Save et al*., 2016]. Annual amplitudes in the 50 largest hydrologic basins globally are higher in the JPL mascon results relative to the SH results; however, application of scaling factors to SH improves agreement between mascon and SH results by ~60% [*Watkins et al*., 2015, Figure 12].

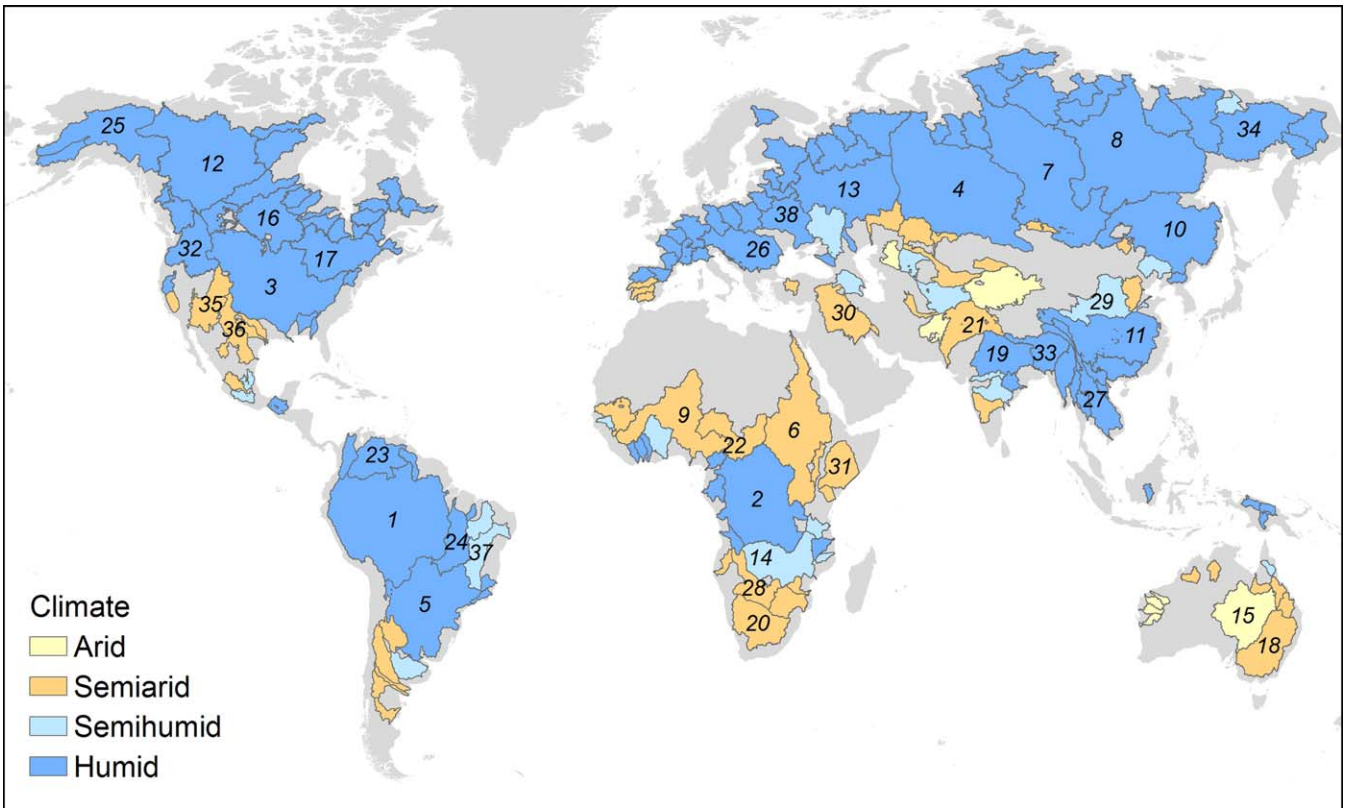


**Figure 2.** Distribution of river basins (176 total) according to climate. Basins with areas ≥500,000 km$^2$ are numbered 1 through 38 in order of decreasing basin area. The climate categories include hyperarid and arid (AI: 0.0–0.2, 8 basins), semiarid (AI: 0.2–0.5, 46 basins), subhumid (AI: 0.5–0.65, 19 basins), and humid (AI > 0.65, 103 basins). The numbered river basins include 1, Amazon; 2, Congo; 3, Mississippi, 4, Ob; 5, Parana; 6, Nile; 7, Yenisei; 8, Lena; 9, Niger; 10, Amur; 11, Yangtze; 12, MacKenzie; 13, Volga; 14, Zambezi; 15, Lake Eyre; 16, Nelson; 17, St. Lawrence; 18, Murray; 19, Ganges; 20, Orange; 21, Indus; 22, Chari; 23, Orinoco; 24, Tocantins; 25, Yukon; 26, Danube; 27, Mekong; 28, Okavango; 29, Huanghe (Yellow River); 30, Euphrates; 31, Jubba; 32, Columbia; 33, Brahmaputra; 34, Kolyma; 35, Colorado; 36, Rio Grande; 37, Sao Francisco; and 38, Dnieper. Basin areas, climate, and irrigation intensity can be found in Table S1.

The primary objective of this study was to compare time series in TWSA from three GRACE solutions: (1) CSR mascons, (2) JPL mascons, and (3) CSR Tellus gridded SH (CSRT-GSH) with and without scaling factor. This is the first comprehensive comparison of GRACE mascon and SH solutions for land hydrology that is based on TWSA from 176 basins globally, representing ~60% of the global land surface (excluding Antarctica and Greenland). The TWSA basin time series were disaggregated into long-term trends, seasonal amplitudes, and residuals. The comparisons of GRACE solutions consider variations in basin size, climate, and irrigation intensity. This comprehensive evaluation of the three GRACE solutions should provide a more thorough understanding of the relative merits of each of the different GRACE solutions and provide the data for hydrologists to select appropriate solutions for different applications.

## 2. Data and Analysis

The following sections provide details of the data used and analysis performed on GRACE TWSA data. A more detailed discussion of the methodology can be found in the supporting information [*Geruo et al.*, 2013; *Cheng and Tapley*, 2004; *Lawrence et al.*, 2011].

### 2.1. Selected River Basins

River basins were selected, totaling 176, to represent a range of climate, land cover, and irrigation intensities (Figures 2 and S4). Basin areas, totaling 79 × 10$^6$ km$^2$, were derived from the Total Runoff Integrating Pathway (TRIP) database [*Oki and Sud*, 1998]. The river basins were subdivided into size classes, including large (38 basins, ≥ 500,000 km$^2$ area; radius ~700 km), medium (85 basins, 100,000–500,000 km$^2$), and small (53 basins, 40,000–100,000 km$^2$, 200–300 km) basins. The climate of the selected basins is based on mean

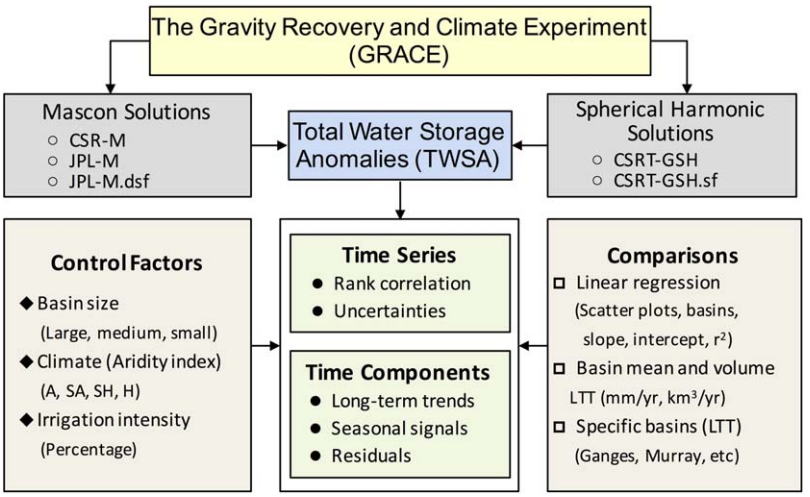


**Figure 3.** Schematic describing GRACE solutions, comparisons of TWSA monthly time series and related uncertainties (section 3.1a and 3.1b), disaggregation into time components (long-term trends [LTT] [3.2], seasonal amplitudes [3.3], and residuals [3.4]), comparisons using linear regression, basin mean trends (mm/yr [3.2a] and $km^3$/yr [3.2b]), and comparison of specific basins with data from the literature (3.2c). Control factors evaluated include basin size, climate, and irrigation intensity.

annual aridity index (AI, mean annual precipitation divided by mean annual potential evapotranspiration, Figure 2) [*Trabucco and Zomer*, 2009], including hyperarid and arid (AI: 0.0–0.2, 8 basins), semiarid (AI: 0.2–0.5, 46 basins), subhumid (AI: 0.5–0.65, 19 basins), and humid (AI $>$ 0.65, 103 basins). The percent of land area equipped for irrigation ranges from 0.00% (51 basins), 0.01–1.0% (59 basins), and 1.0–31% (66 basins) (Figure S4).

### 2.2. GRACE Data

A schematic shows the different GRACE data evaluated in this analysis, including GRACE mascons and spherical harmonics (Figure 3). The time series of TWSA data from different solutions (mascons and SH) were compared using rank correlation. Uncertainties in the TWSA data were quantified. The time series were disaggregated into long-term trends, seasonal (annual and semiannual) amplitudes, and residuals. Linear regression was used to compare the TWSA data from different GRACE solutions to CSR-M data. The long-term trends were examined for well-known basins and compared with trends from the literature. Many of the comparisons examined the impacts of basin size, climate, and irrigation intensity on the relationships among GRACE solutions.

All GRACE data are based on the latest release (RL05) from each of the processing centers, including CSR and JPL. The period analyzed ranges from April 2002 through March 2015, 13 years of data. The data sources and download information are provided in section S1 in the Supporting Information.

#### 2.2.1. GRACE Mascon Solutions

Mascon solutions fall into two basic categories with intersatellite range rate measurements related to:

1. mascons through SH expansion truncated at a finite degree and order, as in CSR mascon solution [*Save et al.*, 2016] and NASA Goddard Space Flight Center (GSFC) [*Rowlands et al.*, 2010; *Luthcke et al.*, 2013, 2015], or
2. mascons through analytical partial derivatives, as in the JPL mascon solution [*Watkins et al.*, 2015]

Although mascons are sometimes fit to SH gravity coefficients derived from GRACE data [*Schrama et al.*, 2014; *Velicogna et al.*, 2014], this approach is essentially a form of Level-3 postprocessing, rather than a Level-2 processing. In this study, we focus only on mascon solutions from CSR and JPL, as these were available to us during the time of this analysis. Future work will incorporate solutions from GSFC into the analysis.

Detailed descriptions of the CSR and JPL mascons data processing are provided in *Save et al.* [2016] and *Watkins et al.* [2015], respectively. CSR-M solutions are computed on an equal area geodesic grid comprised of hexagonal tiles approximately 120 km wide ($\sim 1^\circ \times 1^\circ$ at the Equator). Each mascon cell is related to the

range-rate observations via partials with respect to SH expansion, truncated to degree and order 120. Mass anomalies in each mascon tile are computed from satellite range-rate observations via their partial derivatives [*Save et al*., 2016]. The CSR mascon solutions are produced in two steps:

Step 1: isolation of the land signal from the ocean signal to reduce leakage which uses information from the regularized SH solutions [*Save et al*., 2012] to define constraints, resulting in an intermediate mascon solution, and
Step 2: use of the intermediate solutions generated in step 1 and the regularized SH solutions [*Save et al*., 2012] to generate time variable regularization for the final solutions.

The resulting mascon solutions are then provided on a 0.5° × 0.5° grid. No additional spatial or temporal constraints beyond regularization are applied to the CSR mascon data. The reader is encouraged to refer to *Save et al*. [2016] for the processing details of the mascon solutions developed at CSR.

NASA JPL version 2 mascon data (referred to as JPL-M) were used in this study. The constraints applied in the JPL mascon processing are based on forward models and altimetry data for mass changes over land, oceans, ice, inland seas, earthquake areas, and areas affected by glacial isostatic adjustment. Example models include Global Land Data Assimilation System (GLDAS) LSMs for land [*Rodell et al*., 2004] and ECCO2 for the ocean [*Menemenlis et al*., 2008]. Time correlation is introduced using a sequential Kalman filter to account for trends in the data that are not included in the models. The smoothing associated with the coarse 3° × 3° mascons is downscaled to 1° × 1° using downscaling factors (dsf) calculated with the Community Land Model (CLM ver. 4.0) and resampled finally at 0.5° × 0.5° grid cells. These downscaling factors are mostly close to 1 (Figure S3b). In an additional postprocessing step, a Coastline Resolution Improvement (CRI) filter is applied to mascons that straddle the land–ocean boundary to separate the land and ocean portions of mass. For further information on the CRI filter, downscaling factors, and leakage errors for the JPL-M solution, the reader is referred to *Wiese et al*. [2016].

#### 2.2.2. Tellus Gridded GRACE Spherical Harmonic Solutions

The 1° × 1° GRACE-gridded SH (GSH) data developed by *Landerer and Swenson* [2012] from CSR data were used for comparison with the CSR and JPL mascon solutions. These data are referred to as the CSRT-GSH data and were downloaded from the JPL Tellus website (section S1). Gridded SH data are also available for JPL and GFZ (GeoForschungsZentrum) GRACE data; however, gridded SH data from these other centers are similar to those from CSR [*Sakumura et al*., 2014; *Long et al*., 2015]. Reported anomalies are relative to a mean time baseline from 2004 through 2009. Scale factors are provided separately to account for signal loss during processing related to truncation to degree and order 60 and application of a 300 km Gaussian smoothing filter. The unscaled (GSH) and rescaled (GSH.sf) data, where subscript sf refers to scaling factor, are evaluated here.

#### 2.2.3. Time Series Decomposition

TWSA from each of the GRACE solutions were decomposed into the following time components:

$$S_{total} = S_{long-term} + S_{annual} + S_{semiannual} + \text{Residuals} \quad (1)$$

where the original signal ($S_{total}$) is decomposed into long-term trends, and annual and semiannual amplitudes by simultaneously fitting harmonic functions (sines and cosines) to the data using standard linear least squares regression weighted based on TWSA uncertainties, similar to *Wahr et al*. [2006]. An example is provided for the Ganges basin (Figure S5). The residuals are calculated by subtracting all the other terms from $S_{total}$ and may include interannual and subseasonal signals and noise. The significance of the long-term trends was determined using a modified Mann-Kendall trend test after the annual and semiannual signals were removed [*Hirsch and Slack*, 1984].

Long-term trends were calculated using linear regression applied to the entire basin TWSA time series, expressed as basin wide mean (mm/yr) and as water volume ($km^3$/yr); however, these trends do not represent a monotonic change in TWSA because TWSA can increase and decrease within the time series. Scatter plots were developed based on time components sorted according to basin size, with a slope of 1 and a coefficient of determination ($r^2$) of 1 indicating perfect agreement. The slope of the linear regression indicates the bias relative to CSR-M, with a slope of $<1$ indicating underestimation and $>1$ indicating overestimation. In addition to basin size, the effects of climate and irrigation intensity on long-term trends in TWSA were also examined.

### 2.3. Uncertainty Estimates

TWSA uncertainties for CSR-M were calculated from the residuals after removing long-term trends, and interannual, annual, and semiannual amplitudes as described in more detail in section S1. Interannual signals were estimated by fitting 13 month moving averages to the residuals in equation (1). These uncertainties may overestimate actual uncertainties because the residuals may contain subseasonal signal in addition to noise.

Uncertainties in JPL-M were derived by considering both measurement and leakage errors. Measurement errors are provided on the website and are given by the formal measurement uncertainties (diagonal elements from the posteriori covariance matrix from the gravity inversion), scaled by a factor of 2. The factor of 2 is empirically derived to roughly match the magnitude of the residuals with respect to fitting a trend, annual, and semiannual (equation (2)) to each mascon, as described in *Wahr et al.* [2006]. Leakage errors consider how the shape of each basin conforms to the placement of the mascons, convolved with the spatial distribution of mass at submascon spatial scales according to geophysical models. Leakage errors were calculated using synthetic simulations specifically for the 176 basins, as described in *Wiese et al.* [2016].

Uncertainties in the GRACE SH solutions include GRACE Level 1 measurement errors in the monthly gravity field solutions [*Wahr et al.*, 2006] and leakage errors from the differences between the filtered TWSA applied with the scaling factors and original TWSA from the CLM4 LSM [*Landerer and Swenson*, 2012].

Basin scale uncertainties are less than those calculated by averaging the grid cell uncertainties for all solutions because of spatial correlation of grid cell errors and were calculated according to the procedure described in *Landerer and Swenson* [2012] and summarized in section S1.

At present, there is no direct approach for estimating uncertainties in long-term trends; therefore, these uncertainties for each basin were mathematically derived using weighted linear least squares regression when resolving equation (1).

## 3. Results and Discussion

### 3.1 Comparison of GRACE Monthly Time Series in Total Water Storage Anomalies

#### 3.1.1. GRACE Total Water Storage Anomalies

There is generally good agreement in GRACE TWSA ($S_{total}$ in equation (1)) between JPL and CSR mascons and between rescaled CSRT-GSH.sf and CSR mascon solutions with median rank correlation coefficients ($\rho$ values) for the 176 basins ranging from 0.94 to 0.95 (Figure 4). Data comparisons were conducted relative to CSR-M, not because CSR-M is considered truth but because it has the highest grid resolution and does not require any postprocessing. Application of gridded scaling factors to SH data had little impact on the agreement, resulting in $<$5% differences in correlation coefficients in individual basins. Median $\rho$ values are slightly higher (0.97) for the large basin class ($>$500,000 km$^2$) than the small basin class (0.93; $<$100,000 km$^2$) and basin outliers are more prevalent in the small basin class. There is no systematic variation in $\rho$ values according to climate and irrigation intensity.

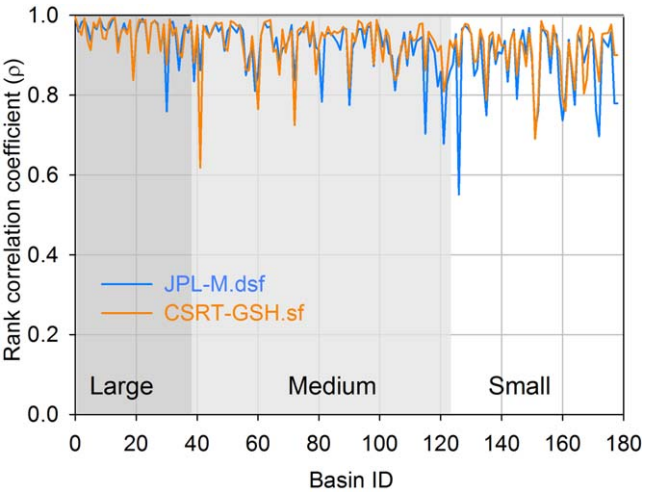


**Figure 4.** Rank correlation coefficients (Spearman's $\rho$) between downscaled JPL mascons (JPL-M.dsf) and CSR Tellus gridded spherical harmonics rescaled (scaling factor, subscript sf, CSRT-GSH.sf) relative to CSR mascons (CSR-M) for 176 basins. Basins are ordered by basin size with large basins (Basin IDs 1–38) $\geq$500,000 km$^2$ (Basin IDs $\leq$ 38), medium basins 100,000 – 500,000 km$^2$ (Basin IDs, 39–123), and small basins $\leq$100,000 km$^2$ (Basin IDs 124–176). Note larger number of outliers in smaller basin class.

#### 3.1.2. Uncertainties in Total Water Storage Anomalies

Uncertainties in GRACE TWSA are similar or slightly lower for GRACE mascon

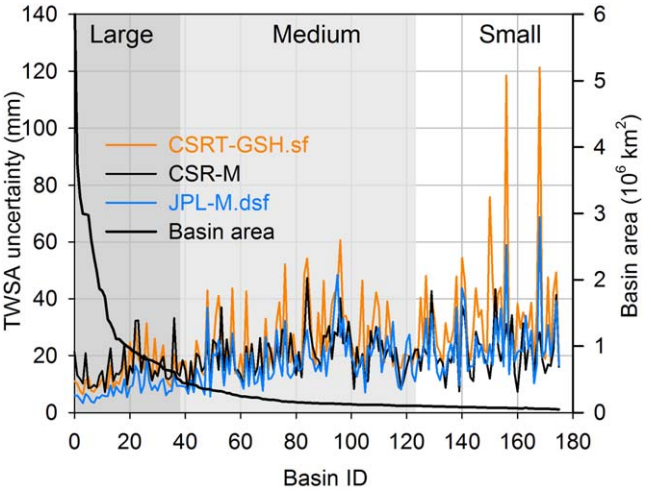


**Figure 5.** Uncertainties in TWSA based on spherical harmonics (CSRT-GSH.sf) and mascon (CSR-M and JPL-M.dsf) solutions for 176 basins ordered by decreasing size. Median values of uncertainty in TWSA according to basin size, climate, and irrigation intensive are provided in Table S2.

solutions relative to SH solutions (Figure 5 and Table S2a). The median error in TWSA is lowest for the large basin size class and increases by up to a factor of 3 for the small basin size class (Table S2a). CSR-M has slightly higher errors than all the other solutions for the large basin class. The errors for the different GRACE solutions are based on different methodologies (Table S2a). However, if we estimate the errors for all GRACE solutions using the same approach as that used for CSR-M, we obtain similar results for mascons and spherical harmonics (Table S2b). Therefore, we suspect that some of the nonseasonal signal retained in the residuals is contributing to the higher uncertainty for CSR-M, especially for the large basins. However, there is uncertainty in the uncertainty estimates because of the lack of independent measurements of TWSA. There is no systematic variation in uncertainties by climate or by irrigation intensity (Table S2).

### 3.2. Long-Term Trends in Total Water Storage Anomalies

#### 3.2.1. Basin-Wide Mean Long-Term Trends in TWSA (mm/yr)

Long-term trends are generally of most interest to hydrologists. Long-term trends for different GRACE solutions are statistically significant in $\geq$70% of all basins and up to 90% for the large basin size class based on the modified Mann-Kendall test. The magnitude of the long-term trends, when expressed in EWT (mm/yr), is generally low, mostly within $\pm$20 mm/yr (Figures 6 and S6). Qualitative comparisons of the trends based on global maps show general agreement in terms of increasing and decreasing trends in TWSA among the CSR and JPL mascons and rescaled SH (CSRT-GSH.sf) solutions (Figures 6, 7, S6 and S7). Differences in long-term trends among the three GRACE solutions generally increase with decreasing basin size (Figure 7a) and uncertainties are highest for the smaller basins ($\leq$100,000 km$^2$ area) (Figure 7b).

Scatter plots show the relationship between long-term trends among GRACE solutions and basin size (delineated with symbol size, Figures 8 and S8; Table 1). Long-term trends from the different GRACE solutions are generally highly correlated with CSR-M ($r^2$: 0.77–0.86 for all basins; Table 1). The level of agreement decreases slightly with decreasing basin size ($r^2$: 0.63–0.80 for small basins). Long-term trends from JPL-M.dsf average 31% higher (slope 1.31) than those for CSR-M (Figure 8a and Table 1). Basin size has little effect on the differences in TWSA between JPL and CSR mascons (Figure S8). Downscaling has little effect on JPL-M long-term trends for all basins (slopes 1.31 versus 1.27) or for different basin size classes (Figure S8 and Table 1).

Long-term trends in TWSA from CSRT-GSH.sf average 4% higher than those from CSR-M (slope 1.04) with no differences for all basin size classes (Figure 8a and Table 1). Scaling brings CSRT-GSH data into alignment with the CSR-M data, increasing the slope from 0.85 to $\sim$1.04 (Figure 8 and Table 1), especially the slope for the small basin size group (0.76–1.04). However, the rescaled SH data have slightly more scatter in the medium and lower basin size groups as seen from the $\sim$10–20% reduction in $r^2$ values, representing increased interbasin variability. While many studies suggest that scaling factors developed for SH solutions may not apply to long-term trends [*Landerer and Swenson*, 2012] because the LSMs used to develop the scaling factors do not simulate some of the long-term processes, such as irrigation, this analysis suggests that the scaling factors improve TWSA long-term trends on average.

Median long-term trends generally become more negative with the transition from humid basins ($-$0.1 to $-$1.3 mm/yr) to arid basins ($-$0.9 to $-$3.4 mm/yr) for each of the GRACE TWSA solutions (Table S3). Long-

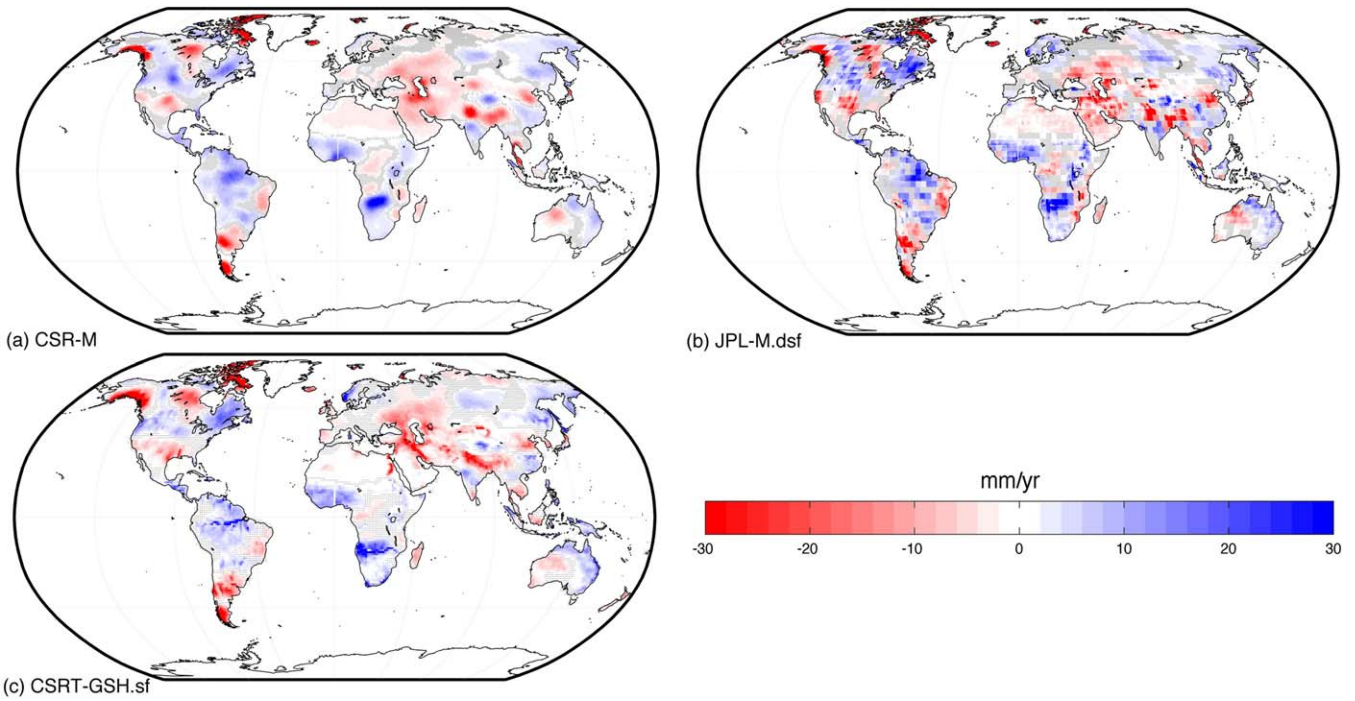


**Figure 6.** Global maps of long-term trends in TWSA in terms of equivalent water thickness (EWT) in mm/yr for (a) CSR-M, (b) JPL-M.dsf, and (c) CSRT-GSH.sf. For larger versions of these maps, see Figure S6. Gray areas denote regions where the long-term trends are not statistically significant based on Mann-Kendall test. Trend calculated for time period April 2002 to March 2015. Figures 6a and 6b are mapped with grid cells at 0.5° × 0.5°, and Figure 6c is at 1.0° × 1.0°.

term trends also vary with irrigation: basins with irrigation ranging from 1 to 31% of the area had the most negative long-term declines in TWSA (median −2.0 to −2.6 mm/yr).

#### 3.2.2. Long-Term Trends in Basin Volumes (km$^3$/yr)

The previous section compared basin-wide mean long-term trends; however, small trends in large basins can impact water volumes as much as or more than large trends in smaller basins (Figures 7c, 7d). Global maps of basin volumetric trends show general similarity among solutions with some notable differences (Figure 9). Basins were sorted according to maximum decreasing ($\leq -2$ km$^3$/yr) and increasing ($\geq 2$ km$^3$/yr) trends using CSR-M data (Table 2. S4a) and corresponding uncertainties are provided in Table S4. Large reductions are found in the Ganges, Brahmaputra, and Indus basins, and also in the Euphrates and Volga basins. Large rises were recorded in the Amazon and in Southern and West Africa (Zambezi, Okavango, and Niger). However, GRACE solutions for some basins have opposite trends: trends from CSRT-GSH.sf are negative (−1.1 to −2.7 km$^3$/yr) for the Mississippi, Parana, and Nile basins but are positive from CSR-M (2.9–9.0 km$^3$/yr) with even higher positive trends from JPL-M (6.0–12.0 km$^3$/yr) (Table 2 and Table S4). We also computed the net volumetric long-term trend (km$^3$/yr) by summing the trends in all 176 basins (Table 3). CSR and JPL mascons give similar net long-term trends (66–69 km$^3$/yr) (Table 3). Net long-term trends from mascons are much higher than those for SH (14 km$^3$/yr). Most of the differences in trends can be accounted for by differences in the large basin class (82–89 km$^3$/yr for CSR and JPL mascons versus 28–32 km$^3$/yr for SH; Table 3). Differences between mascon and SH solutions are greatest in humid basins (56–66 km$^3$/yr for CSR and JPL mascons versus 16–18 km$^3$yr for SH). There are no systematic differences in volumetric trends among solutions for irrigation levels; however, net depletions are highest in the highest irrigation category (−41 to −69 km$^3$/yr).

#### 3.2.3. Comparison of Basin Specific Long-Term Trends With Previous Studies

Long-term trends from different GRACE solutions were compared for large basins that have been studied previously. The well-known drying trend in the vicinity of Northwest India is evident in the decreasing TWSA trends in the Indus, Ganges, and Brahmaputra basins (−6.9 to −17.5 km$^3$/yr; Figure 10a), consistent with previously published results from *Rodell et al.* [2009], *Chen et al.* [2014], and *Long et al.* [2016]. Long-term trends for the Ganges differ by 40% between CSR-M (−12.7 km$^3$/yr) and JPL-M (−17.5 km$^3$/yr), which is surprising because the Ganges is a large basin ($\sim$ 1 million km$^2$) and trends would be expected to

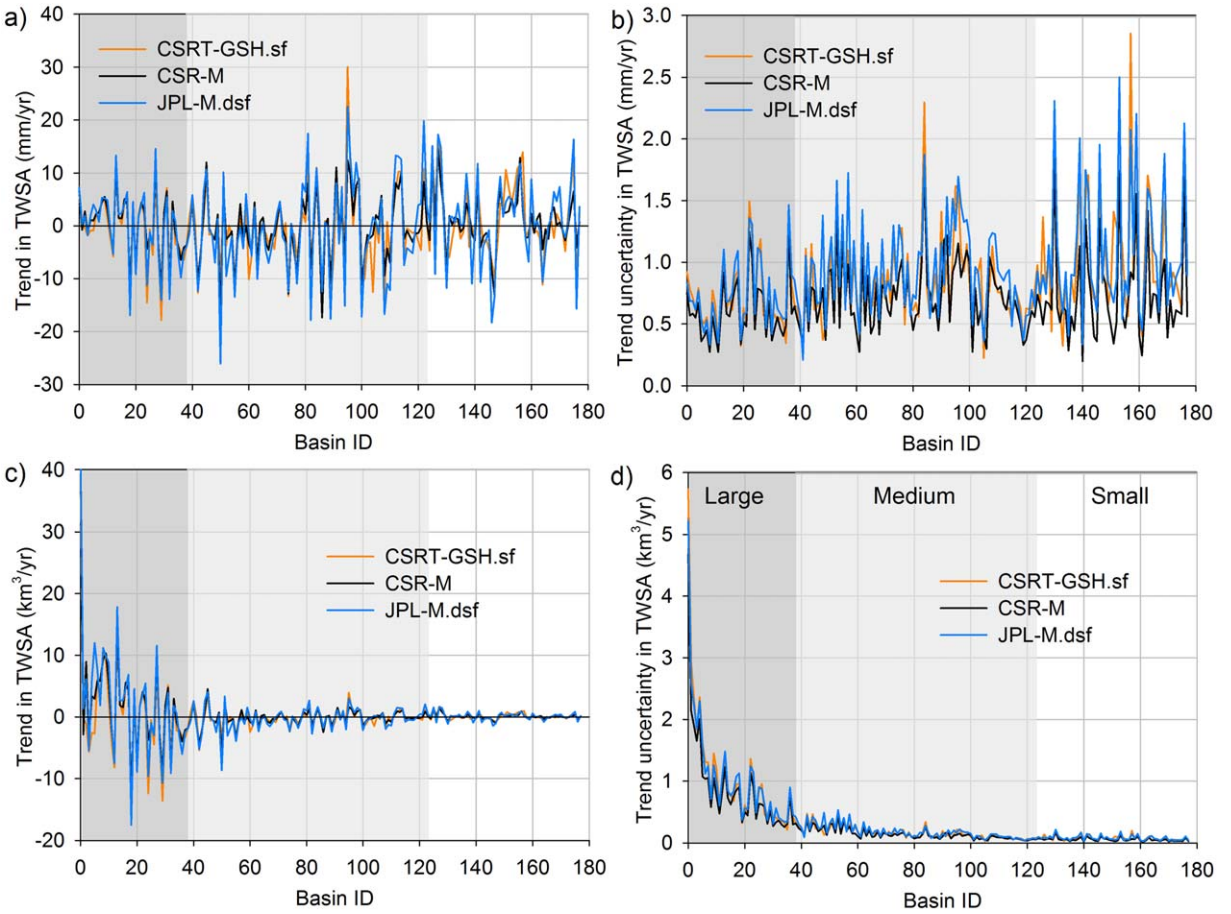


**Figure 7.** Long-term trend in TWSA for 176 river basins in terms of (a) equivalent water thickness in mm/yr and (c) water volume in $km^3$/yr based on CSRT-GSH.sf, CSR-M, and JPL-M.dsf. Related uncertainties are shown in (b) and (d) corresponding to Figures 7a–7c. Basins are arranged according to decreasing basin size. Median values of long-term trends in TWSA according to basin size, climate, and irrigation intensity are provided in Table S3.

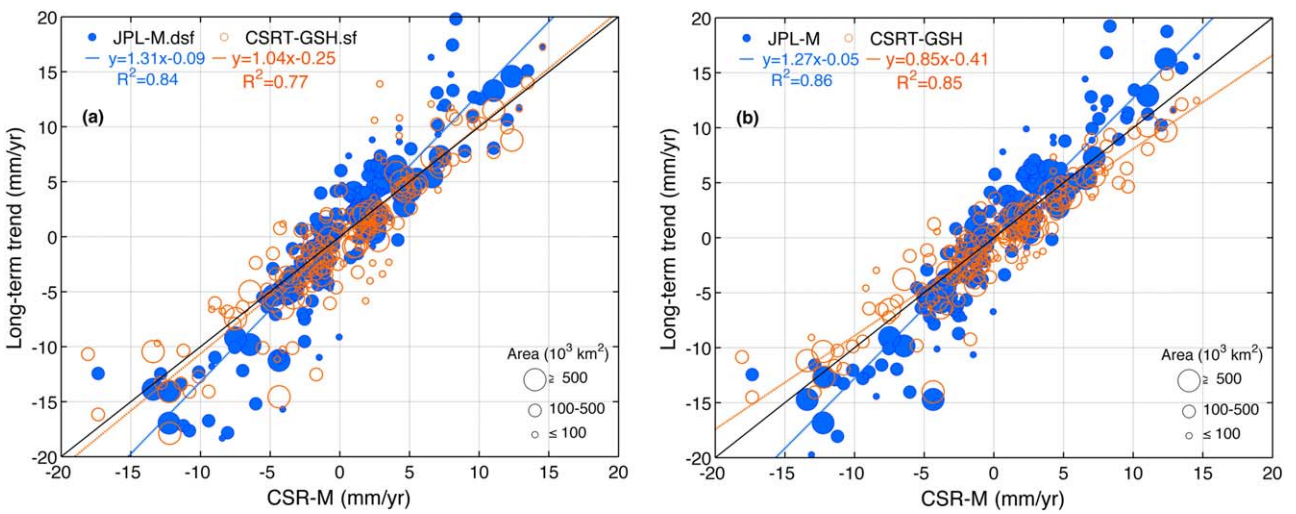


**Figure 8.** Scatter plot of long-term trends in TWSA for 176 river basins based on (a) JPL-M.dsf (downscaled to 0.5°, blue), CSRT-GSH.sf (orange), and CSR-M, and (b) JPL-M (original 3° data) and CSRT-GSH without scaling relative to CSR-M. The 1:1 line is shown in solid black and represents perfect agreement between the selected solutions and CSR-M. Colored lines represent linear regressions and equations are provided along with $R^2$ values. Regression equations are also shown in Table 1. Separate scatter plots for large (38 basins) and small (53) basins are also included in Figure S8. Regression parameters are provided in Table 1.

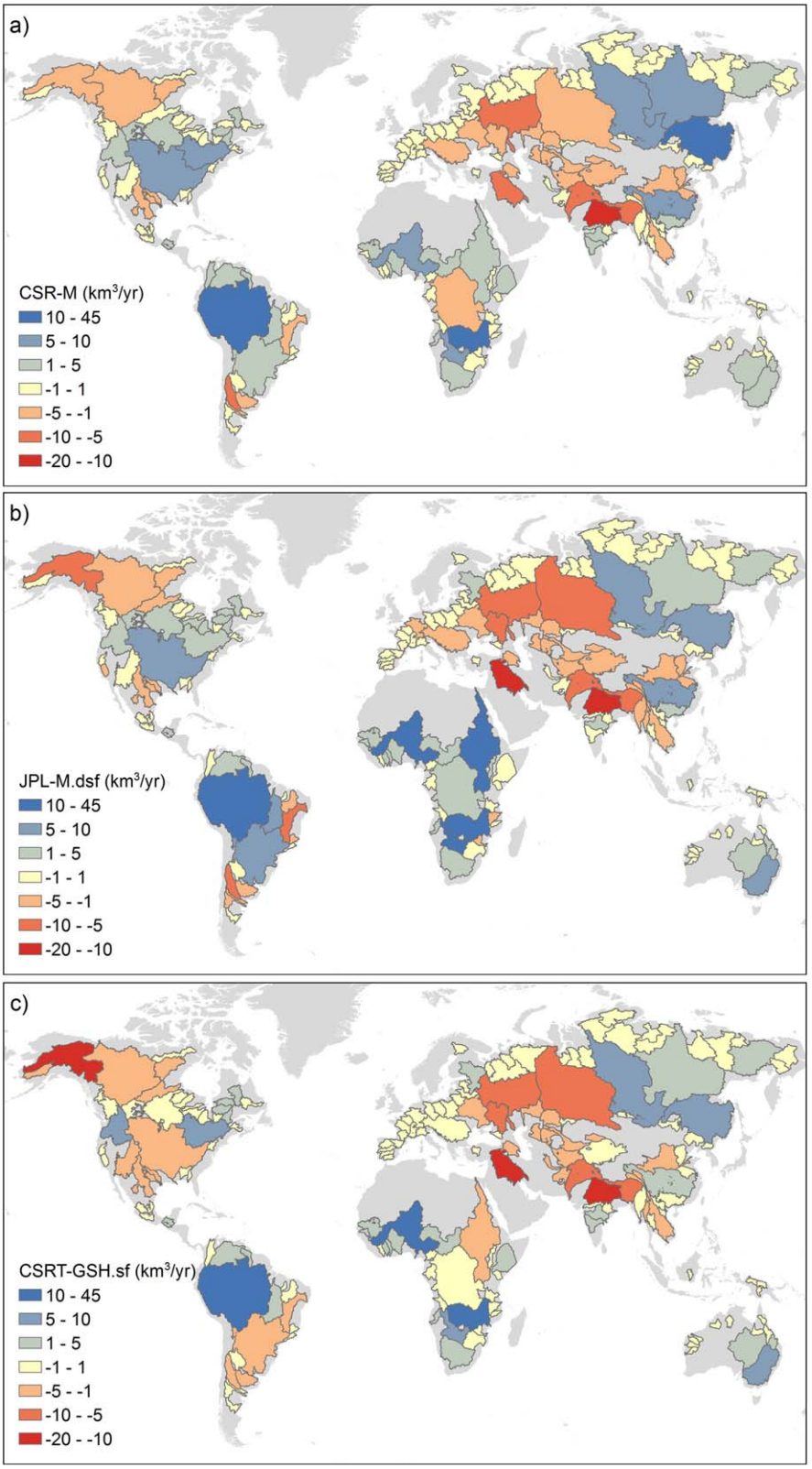


**Figure 9.** Global maps of long-term trends in TWSA in terms of volume (km$^3$/yr) by basin for (a) CSR-M, (b) JPL-M.dsf, and (c) CSRT-GSH.sf. Basin names for large basins can be found in Figure 1 and for remaining basins in Table S1. Tabulated trends for selected basins can be found in Table 2. Maps in basin-wide mean (mm/yr) are provided in Figure S7.

**Table 1.** Linear Regression Relationship Between Different GRACE Solutions and CSR Mascon Solution for 176 River Basins: Large (38), Medium (85), and Small (53)[a]

Linear Regression Parameters Relative to CSR-M

| | Global (176) | | | Large (38) | | | Medium (85) | | | Small (53) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Solutions | *b* | *c* | $r^2$ | *b* | *c* | $r^2$ | *b* | *c* | $r^2$ | *b* | *c* | $r^2$ |
| Long-term trend (mm/yr) | | | | | | | | | | | | |
| JPL-M.dsf | 1.31 | −0.09 | 0.84 | 1.18 | −0.25 | 0.94 | 1.32 | 0.05 | 0.85 | 1.43 | −0.29 | 0.76 |
| CSRT-GSH.sf | 1.04 | −0.25 | 0.77 | 1.04 | −0.95 | 0.88 | 1.03 | −0.20 | 0.80 | 1.04 | 0.17 | 0.63 |
| JPL-M | 1.27 | −0.05 | 0.86 | 1.20 | −0.32 | 0.91 | 1.26 | 0.17 | 0.88 | 1.41 | −0.32 | 0.80 |
| CSRT-GSH | 0.85 | −0.41 | 0.85 | 0.91 | −0.80 | 0.89 | 0.85 | −0.43 | 0.86 | 0.76 | −0.02 | 0.77 |
| Annual amplitude (mm) | | | | | | | | | | | | |
| JPL-M.dsf | 1.08 | −5.13 | 0.93 | 1.06 | −0.60 | 0.99 | 1.12 | −7.96 | 0.93 | 1.06 | −4.99 | 0.88 |
| CSRT-GSH.sf | 0.99 | 5.37 | 0.83 | 1.04 | 4.15 | 0.98 | 0.97 | 6.96 | 0.80 | 0.98 | 4.20 | 0.75 |
| JPL-M | 1.04 | −3.46 | 0.93 | 1.04 | 0.22 | 0.99 | 1.11 | −8.30 | 0.95 | 0.94 | 0.80 | 0.85 |
| CSRT-GSH | 0.70 | 8.61 | 0.83 | 0.81 | 9.34 | 0.94 | 0.73 | 8.23 | 0.86 | 0.60 | 8.18 | 0.77 |
| Semiannual amplitude (mm) | | | | | | | | | | | | |
| JPL-M.dsf | 1.03 | 0.44 | 0.83 | 1.06 | −0.43 | 0.94 | 1.01 | −0.30 | 0.89 | 1.16 | 0.89 | 0.66 |
| CSRT-GSH.sf | 0.97 | 0.90 | 0.74 | 1.00 | 0.60 | 0.94 | 0.96 | 0.51 | 0.87 | 0.98 | 1.38 | 0.38 |
| JPL-M | 0.97 | 0.72 | 0.87 | 1.05 | −0.39 | 0.95 | 0.95 | 0.15 | 0.91 | 1.03 | 1.55 | 0.70 |
| CSRT-GSH | 0.70 | 1.24 | 0.83 | 0.80 | 1.36 | 0.92 | 0.67 | 1.58 | 0.87 | 0.67 | 0.51 | 0.65 |
| Root mean square of residual (mm) | | | | | | | | | | | | |
| JPL-M.dsf | 1.21 | 3.26 | 0.79 | 1.16 | 1.64 | 0.94 | 1.11 | 5.17 | 0.73 | 1.25 | 6.47 | 0.82 |
| CSRT-GSH.sf | 0.90 | 10.54 | 0.52 | 1.08 | 2.87 | 0.85 | 0.92 | 7.57 | 0.59 | 0.78 | 20.74 | 0.40 |
| JPL-M | 1.17 | 3.42 | 0.86 | 1.14 | 2.02 | 0.95 | 1.09 | 4.98 | 0.80 | 1.20 | 5.68 | 0.90 |
| CSRT-GSH | 0.53 | 13.01 | 0.72 | 0.71 | 8.52 | 0.87 | 0.55 | 12.19 | 0.75 | 0.47 | 14.75 | 0.65 |

[a]Slope and intercept and coefficient of determination are provided for long-term trend (April 2002 to March 2015); annual amplitudes (1 cycle/yr), semiannual amplitudes (2 cycles/yr), and residuals are provided.

be similar for large basins. Both CSR and JPL mascon solutions give similar long-term trends as unscaled CSRT GSH results for the Euphrates basin in the Middle East (∼−9.3 to −10.7 $km^3$/yr) (Figure 10b); however, applying a scaling factor to SH (CSRT-GSH.sf, −13.6 $km^3$/yr) increases the trend by 46% relative to CSR-M (−9.3 $km^3$/yr) (Figure S9). This example indicates that sometimes application of scaling factors to SH can result in larger differences relative to mascons trends at the basin scale. Long-term declines in TWSA in

**Table 2.** Basins With Significant Long-Term Trends in TWSA in Total Volume ($km^3$/yr) Sorted in Terms of Decreasing and Increasing Trends for CSR-M, JPL-M.dsf, and CSRT-GSH.sf Solutions, Sorted Based on CSR-M Data[a]

| Decreasing Long-Term Trends ($km^3$/yr) | | | | | Increasing Long-Term Trends ($km^3$/yr) | | | | |
|---|---|---|---|---|---|---|---|---|---|
| ID | River | CSR-M | JPL-M.dsf | CSRT-GSH.sf | ID | River | CSR-M | JPL-M.dsf | CSRT-GSH.sf |
| 19 | Ganges | −12.7 | −17.5 | −14.5 | 1 | Amazon | 44.8 | 45.4 | 39.5 |
| 30 | Euphrates | −9.3 | −10.7 | −13.6 | 14 | Zambezi | 14.8 | 17.8 | 15.5 |
| 33 | Brahmaputra | −8.8 | −9.1 | −6.9 | 10 | Amur | 10.3 | 9.8 | 9.0 |
| 21 | Indus | −7.3 | −8.9 | −7.2 | 28 | Okavango | 9.8 | 11.6 | 7.0 |
| 13 | Volga | −6.7 | −7.5 | −8.2 | 9 | Niger | 9.4 | 11.2 | 10.7 |
| 54 | Salado[b] | −6.0 | −8.6 | −3.5 | 3 | Mississippi | **9.0** | 6.1 | **−1.1** |
| 43 | Don | −4.3 | −5.2 | −5.4 | 8 | Lena | 5.9 | 1.9 | 3.1 |
| 37 | Sao Francisco | −4.0 | −6.0 | −3.1 | 7 | Yenisei | 5.9 | 7.6 | 6.1 |
| 4 | Ob | −3.8 | −5.4 | −5.6 | 11 | Yangtze | 5.7 | 8.9 | 2.7 |
| 25 | Yukon | −3.7 | −9.6 | −12.4 | 17 | St. Lawrence | 5.6 | 4.6 | 5.0 |
| 62 | Salado[c] | −3.3 | −3.9 | −3.9 | 32 | Columbia | 4.8 | 3.9 | 5.2 |
| 2 | Congo | **−2.8** | **1.4** | **−0.3** | 47 | Volta | 4.5 | 4.0 | 3.8 |
| 29 | Huanghe | −2.8 | −4.0 | −3.2 | 18 | Murray | 4.3 | 6.9 | 6.3 |
| 27 | Mekong | −2.7 | −2.9 | −4.5 | 23 | Orinoco | 4.1 | 3.7 | 4.0 |
| 100 | Telon | −2.5 | −1.8 | −2.3 | 5 | Parana | **3.4** | 6.0 | **−2.6** |
| 36 | Rio Grande | −2.4 | −3.2 | −2.3 | 34 | Kolyma | 3.0 | 1.8 | 2.7 |
| 84 | Kura | −2.3 | −2.2 | −2.4 | 6 | Nile | **2.9** | **12.0** | **−2.7** |
| 44 | Syr Darya | −2.2 | −2.7 | −2.0 | 20 | Orange | 2.7 | 4.6 | 2.3 |
| 38 | Dnieper | −2.1 | −2.5 | −3.3 | 24 | Tocantins | 2.5 | 5.5 | 1.9 |
| 12 | MacKenzie | −2.1 | −1.5 | −4.2 | 55 | Godavari | 2.3 | 3.3 | 2.1 |
| | | | | | 41 | Senegal | 2.1 | 2.6 | 2.6 |

[a]The corresponding trends in mm/yr and uncertainties in long-term trends are provided in Table S4. Basins with opposing trends are bolded.
[b]La Pampa.
[c]Atlantic Ocean.

**Table 3.** Net Long-Term Trends (km$^3$/yr) Summed Over 176 Basins Based on Mascon and Spherical Harmonic Solutions According to Basin Size, Climate, and Irrigation

| Solutions | CSR-M | JPL-M | JPL-M.dsf | CSRT-GSH | CSRT-GSH.sf |
|---|---|---|---|---|---|
| Basin area | | | | | |
| Global (176) | 69.2 | 65.6 | 68.1 | 13.8 | 14.3 |
| Large (38) | 82.3 | 85.2 | 88.8 | 32.0 | 28.3 |
| Medium (85) | −17.6 | −24.6 | −25.8 | −21.6 | −19.2 |
| Small (53) | 4.6 | 4.9 | 5.1 | 3.4 | 5.2 |
| Climate (AI) | | | | | |
| Arid | −2.2 | −2.2 | −1.2 | −2.4 | −2.9 |
| Semiarid | 1.5 | 9.5 | 5.6 | −2.0 | −5.4 |
| Subhumid | 4.4 | 2.4 | 2.0 | 2.2 | 4.5 |
| Humid | 65.5 | 55.9 | 61.7 | 15.9 | 18.1 |
| Irrigation (%) | | | | | |
| 0.0–0.01 | 8.8 | 2.1 | 4.5 | −9.4 | −7.5 |
| 0.01–1.0 | 101.4 | 115.4 | 116.1 | 82.2 | 90.6 |
| 1.0–31.0 | −41.0 | −52.0 | −52.5 | −59.0 | −68.8 |

South America occur primarily in the Salado Basin within the Lower La Plata Basin (Figure 10c) consistent with the trends reported previously related to drought [*Chen et al*., 2010a; *Abelen et al*., 2015]. However, there are large differences in the magnitude of the trends among the GRACE solutions (−3.5 to −8.6 km$^3$/yr) attributed to proximity to the mountain glacier.

Increases in TWSA in various basins are generally consistent with previous studies. Increases in TWSA in the Amazon show good agreement between SH and mascon solutions (40–45 km$^3$/yr). The increases in TWSA from GRACE are consistent with the findings from previous studies showing that the Amazon has been subjected to wet and dry periods [*Chen et al*., 2010b; *Xavier et al*., 2010]. Many of the areas of increasing TWSA are found in Africa: west Africa (Volta and Niger, ∼4 km$^3$/yr and 9.4–11.2 km$^3$/yr) with generally good agreement between SH and mascon solutions and consistent with GRACE results from *Ramillien et al*. [2014]. Increasing TWSA in the Okavango (7.0–11.6 km$^3$/yr, higher for mascons relative to SH) and Zambezi (14.8–17.8 km$^3$/yr, similar between mascons and SH) are similar to findings from *Milzow et al*. [2009] and *Ramillien et al*. [2014]. The increasing trend in the Murray Basin in Australia (4.3–6.9 km$^3$/yr; Figure 10d) is generally similar among GRACE solutions and is attributed to recovery from a decadal-scale drought [*Chen et al*., 2016]. These comparisons indicate that each of the GRACE solutions provides the general trends in these large basins; however, the magnitude of the trends can differ among the various solutions and use of scaling factors with SH does not always bring them into closer agreement with mascon solutions.

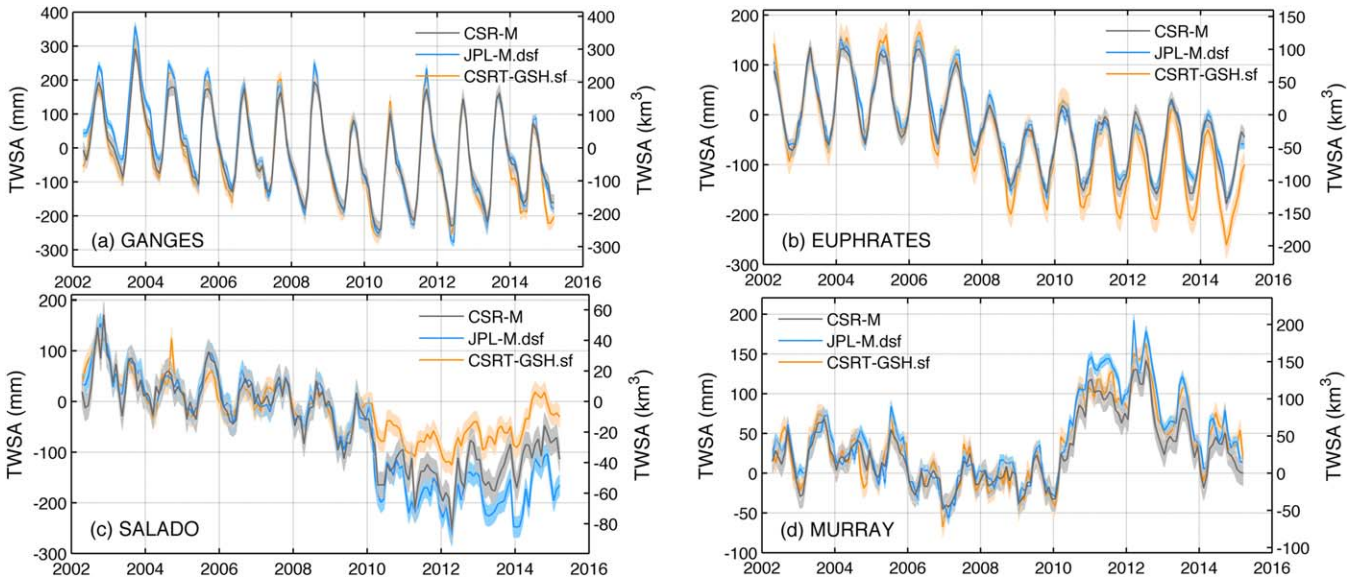


**Figure 10.** Time series of TWSA derived from CSR-M, JPL-M.dsf, and CSR-GSH.sf solutions for (a) Ganges, (b) Euphrates, (c) Salado (La Pampa), and (d) Murray basins. Shaded areas are estimated uncertainties as shown in Figure 5. Corresponding long-term trends in TWSA in the four basins are provided in Table 2.

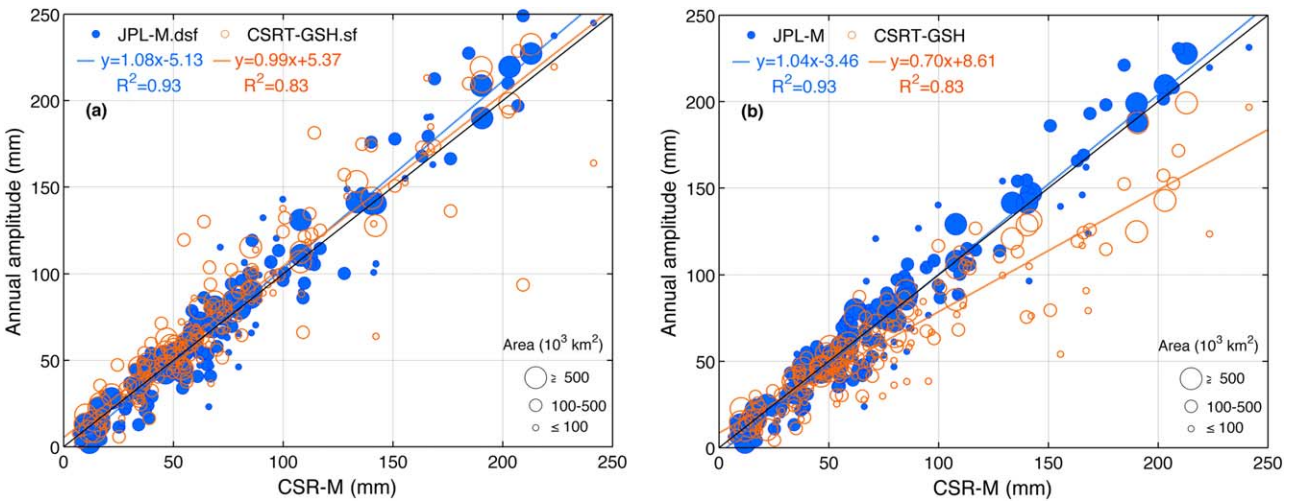


**Figure 11.** Scatter plot of annual amplitudes in TWSA for 176 river basins based on (a) JPL-M.dsf (downscaled to 0.5°, blue), CSRT-GSH.sf (orange), and CSR-M, and (b) JPL-M (original 3° data) and CSRT-GSH without scaling relative to CSR-M. The 1:1 line is shown in solid black and represents perfect agreement between the selected solutions and CSR-M. Colored lines represent linear regressions and equations are provided along with $R^2$ values. Separate scatter plots for large (38 basins) and small (53 basins) also included in Figure S8. Regression parameters are provided in Table 1.

### 3.3. Seasonal Fluctuations in Total Water Storage Anomalies

The amplitude of the annual signal is up to an order of magnitude higher than the long-term trends (up to ~250 mm relative to ±20 mm/yr) (Figures 11, 12, 13, S10 and S11). The annual amplitudes vary with latitude and are highest in equatorial and subtropical regions with magnitudes ranging from 190 to 250 mm in basins such as the Amazon, Tocantins, Orinoco, and Irrawaddy (Figures 12, 13, and S11). There is generally good correspondence between mean annual amplitudes for JPL-M.dsf, CSR-M, and CSRT GSH.sf, as shown by slopes close to 1.0 and high $r^2$ values 0.83–0.93 (Figures 11 and S12, Table 1). Interbasin variability increases slightly with decreasing basin size, as shown by the reduction in $r^2$ (Table 1). The similarity in annual amplitudes for all three GRACE solutions (CSR-M, JPL-M, and CSRT-GSH.sf) provides confidence in GRACE results (Figure 11a and Table 1). While the unscaled CSRT-GSH data have lower annual amplitudes by ~30–40% relative to mascon solutions for the different basin size classes (Figure 11b), rescaling aligns the mean SH results with the mascons results (Figure 11a) and slightly decreases the scatter in the data, as seen in the ~10% increase in $r^2$ values between unscaled and rescaled SH data (Table 1).

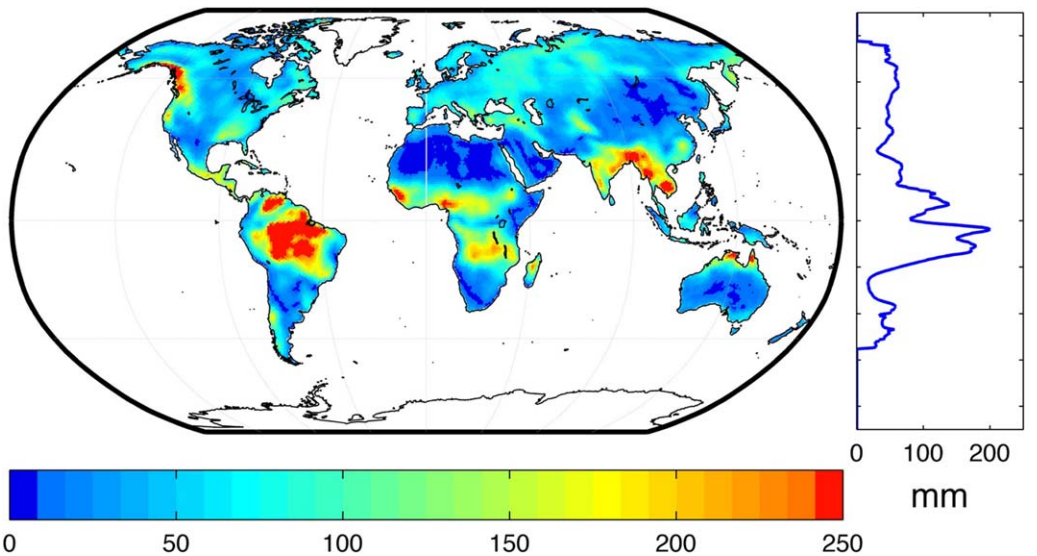


**Figure 12.** Global map of mean annual amplitude (mm) in TWSA based on CSR-M solution and variations with latitude shown on right-hand side. Global maps of mean annual amplitudes for JPL-M and CSRT-GSH.sf are also shown in Figure S10. (Map with grid cells at 0.5° × 0.5°.)

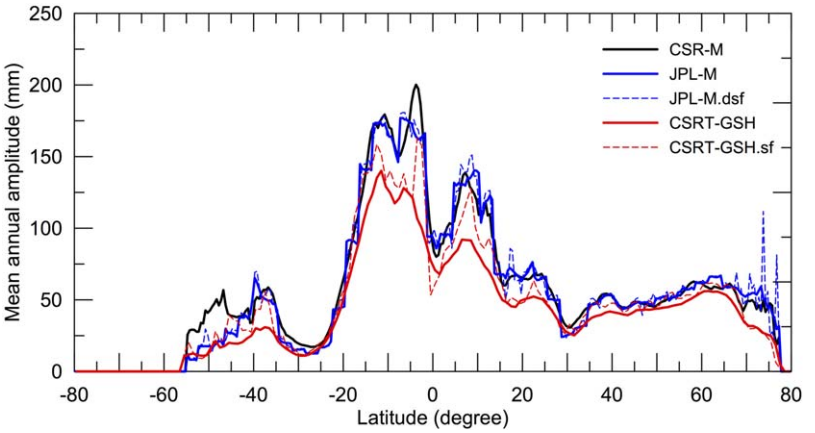


**Figure 13.** Mean annual amplitude relative to latitude from South to North. Root mean square of annual amplitude versus latitude is shown in Figure S11.

The amplitudes of the semiannual GRACE TWSA signals (up to 70 mm) data are about 3 times less than the maximum annual amplitudes on average; however, the latitudinal pattern is similar to that of the annual amplitudes (Figures S13 and S14). There is generally good agreement in semiannual amplitudes among the three GRACE solutions (JPL-M, CSR-M, and CSRT-GSH.sf), with similar slopes to those for the annual amplitudes (Figure S15 and Table 1). Rescaling SH is very important for bringing TWSA into alignment with mascon solutions and increasing the signal amplitude by 20–30% for all basin size classes (Table 1).

### 3.4. Residuals

The residuals (equation (1)) may reflect interannual signals or signals at higher frequency than semiannual, and also noise (Figure 14 and S16). Root mean square of the residuals for JPL-M.dsf exceed those from CSR-M on average by 21% (slope of JPL-M.dsf in 176 basins, 1.21), essentially independent of basin size (Figures S17a, S17b, and Table 1). Rescaled CSRT-GSH residuals (slope, 0.90) average 10% less than those from CSR-M for all basins, with differences between the two varying according to basin size, lower differences for large basins (Figure S17). Application of scaling factors to SH increases slopes by up to 30–40% relative to CSR-M for all basin size classes but discrepancies between SH and mascons are still large for the small basin size class (Figures S17e, S17f and Table 1).

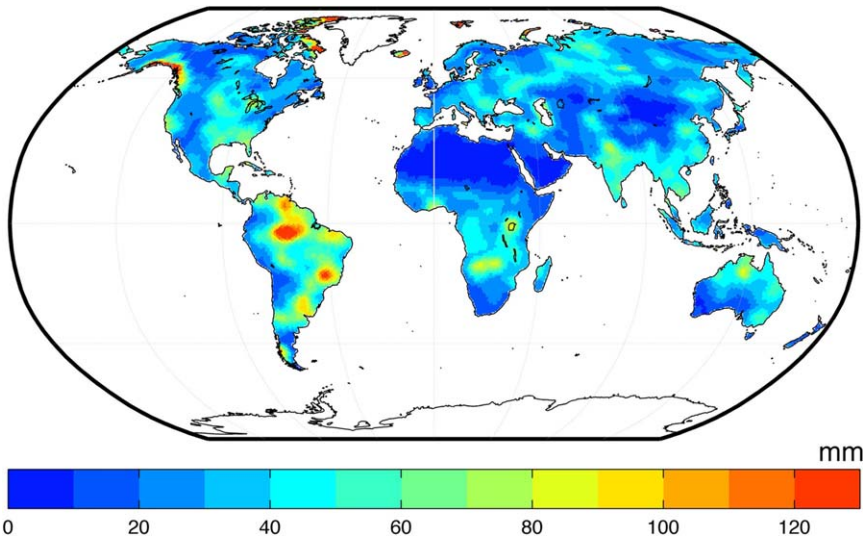


**Figure 14.** Root mean square (RMS) of the residuals in TWSA after the long-term trend and seasonal amplitudes have been removed. The map is based on the CSR-M solution (mm). RMS of the residuals for JPL-M and CSRT GSH are also shown in Figure S16. (Map with grid cells at 0.5° × 0.5°.)

### 3.5. Recommendations for Hydrologic Applications of GRACE Data

Recommendations for GRACE solutions vary with the application, such as global analysis versus single basin analysis and evaluation of long-term trends versus seasonal amplitudes. TWSA from the three GRACE solutions (CSR and JPL mascons and CSR spherical harmonics) are generally highly correlated with spearman's $\rho$ values mostly $>$0.9. Uncertainties in GRACE solutions are similar or lower for GRACE mascons relative to spherical harmonics. Uncertainties increase by up to a factor of $\sim$3 from large to small basins (Table S2a). However, it is difficult to determine actual uncertainties for GRACE data because of the lack of independent estimates of TWSA. The variability among GRACE solutions may provide an alternative indication of uncertainties in GRACE output [*Sakumura et al*., 2014].

Evaluation of long-term trends from GRACE seems to be of most interest to hydrologists. Long-term trends are generally $\leq\pm$20 mm/yr. The level of agreement or consistency in long-term trends among GRACE solutions decreases with decreasing basin size. Therefore, for single basin analysis, it is important to consider the variability among GRACE solutions for long-term trends, particularly for small basins ($\leq$100,000 km$^2$). Long-term volumetric trends summed over all 176 basins result in similar net changes in TWSA for mascon solutions (66–69 km$^3$/yr) which are much greater than trends in SH ($\sim$ 14 km$^3$/yr). Rescaling realigns SH results with mascon results. Therefore, it is important to examine the range of GRACE solutions for long-term trends. The impacts of rescaling SH output should also be examined because analysis of individual basins shows that rescaling of SH results in greater discrepancies with mascon solutions in some basins.

The GRACE processing approach is not as critical for analysis of seasonal fluctuations because the signals are generally much larger than the long-term trend signals, up to 250 mm relative to $\pm$20 mm/yr. Consistency in seasonal signals for the different solutions (CSR and JPL mascons and rescaled SH) is much higher for seasonal amplitudes than for long-term trends. Although rescaled SH are comparable with the mascon solutions, this is clearly not the case for the nonscaled SH, which highlights the importance of the scaling factors for seasonal amplitudes.

Future studies should examine residuals for interannual and subseasonal signals. A recent study lag correlated subseasonal signals in unscaled SH data with precipitation and temperature data related to droughts and floods, showing that the residuals may contain important signals [*Humphrey et al*., 2016].

### 3.6. Future Outlook for GRACE Data

The results of this study highlight many of the advantages of the mascons products for hydrologic studies. In addition, a recently published study by *Sakumura et al*. [2016] describes a regularized sliding window mascon solution based on CSR-M that solves for daily TWSA representing the GRACE data for the window surrounding that day. Daily data greatly enhance hydrologic applications of GRACE TWSA, particularly for floods and other short-lived events.

The current GRACE mission is winding down; however, it has greatly outlived the original expectations of a five year lifespan. The GRACE Follow-On (GRACE-FO), set for launch in late 2017, is the successor to the current GRACE mission and is also jointly implemented in a U.S.-German partnership. GRACE-FO is designed to be almost identical to the current GRACE mission in most respects with the main goal of providing continuous global solutions of the static and time variable gravity fields.

Relative to the current GRACE data, new products from GRACE-FO that should be available include inertial measurement unit Level-1 data, low latency Level 2–3 and final Level-3 data. The lower latency products provide valuable input for hydrologic studies related to regional drought and flood monitoring and prediction. A laser-ranging interferometer will be added to GRACE-FO to further improve the precision of the current measurement system.

## 4. Summary

Changes in TWSA monitored by GRACE satellites provide a unique measurement that does not have a direct ground-based analog for independent validation. However, variability in TWSA data from different processing approaches provides an indication of TWSA uncertainty. The newly released CSR and JPL mascon solutions offer several advantages relative to traditional SH processing, including reduced leakage resulting in

increased signal amplitude, and constraints applied during processing requiring little or no postprocessing. Gridded SH solutions were also developed with no postprocessing requirements to increase applications by hydrologists. CSR and JPL mascons and SH processing were applied to 176 river basins covering almost 60% of the global land surface to assess variability among GRACE solutions.

Time series of GRACE TWSA from three GRACE solutions (CSR and JPL mascons and CSR spherical harmonics) are highly correlated (median $\rho$ 0.97–0.99). Long-term trends in TWSA are of greatest interest to hydrologists but generally have the lowest signal, mostly within $\pm$20 mm/yr. Basin-wide average long-term trends for JPL-M are 27–31% higher than those of CSR-M whereas trends for CSRT-GSH average 15% less than CSR-M. Scaling factors derived from land surface models bring the SH long-term trends into alignment with the CSR-M results, but interbasin variability slightly increases. Differences in long-term trends among GRACE solutions increase with decreasing basin size by up to a factor of 3 when going from basins $\geq$500,000 km$^2$ to $\leq$100,000 km$^2$, indicating that the processing approach may be more critical for small basins. TWSA long-term trends become more negative with the transition from humid to arid basins for each of the GRACE TWSA solutions. River basins with irrigation intensities ranging from 1 to 31% have the most negative TWSA long-term trends. Global maps of basin volumetric long-term trends show general similarity among basins with some large discrepancies in some basins (e.g., Mississippi, Parana, and Nile basins). Comparison of basin long-term trends for many well-known basins shows variable correspondence among the GRACE mascon and SH solutions depending on the basin, with generally consistent trends in TWSA with previously published studies for these basins.

The amplitudes of the GRACE annual TWSA signals are up to an order of magnitude higher than those of long-term trends, mostly near the Equator and subtropical regions. There is good agreement in TWSA annual amplitudes among CSR and JPL mascons and rescaled CSRT SH solutions, increasing confidence in seasonal fluctuations in GRACE TWSA. Rescaling improves correspondence between SH and mascon solutions on average. Applying rescaling for determination of TWSA long-term trends in individual basins still needs to be carefully evaluated.

The many advantages of the new GRACE CSR and JPL mascon solutions relative to traditional SH solutions make it much easier for non-geodesists to apply GRACE data to hydrologic problems. Because of the variability among GRACE solutions, basin-scale analysis of TWSA trends should examine the variety of solutions available to constrain uncertainties, particularly for long-term trends.

**Acknowledgments**
We would like to acknowledge financial support for the Senior Author from the Jackson School of Geosciences. Additional data are provided in supporting information. The work performed by D. N. Wiese and F. W. Landerer was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. GRACE Tellus land grids are available at http://grace.jpl.nasa.gov, supported by the Jackson School of Geosciences, University of Texas at Austin. The CSR GRACE RL05 mascons solutions were downloaded from http://www.csr.utexas.edu/grace.

## References

Abelen, S., F. Seitz, R. Abarca-del-Rio, and A. Guentner (2015), Droughts and floods in the La Plata Basin in soil moisture data and GRACE, *Remote Sens.*, *7*(6), 7324–7349.

Bettadpur, S. (2012), *CSR Level-2 Processing Standards Document for Product Release 05*, *GRACE 327-742*, The GRACE Project, Cent. for Space Res., Univ. of Tex., Austin, Tex.

Brena-Naranjo, J. A., A. D. Kendall, and D. W. Hyndman (2014), Improved methods for satellite-based groundwater storage estimates: A decade of monitoring the high plains aquifer from space and ground observations, *Geophys. Res. Lett.*, *41*, 6167–6173, doi:10.1002/2014GL061213.

Chen, J. L., C. R. Wilson, B. D. Tapley, L. Longuevergne, Z. L. Yang, and B. R. Scanlon (2010a), Recent La Plata basin drought conditions observed by satellite gravimetry, *J. Geophys. Res.*, *115*, D22108, doi:10.1029/22010JD014689.

Chen, J. L., C. R. Wilson, and B. D. Tapley (2010b), The 2009 exceptional Amazon flood and interannual terrestrial water storage change observed by GRACE, *Water Resour. Res.*, *46*, W12526, doi:10.1029/2010WR009383.

Chen, J. L., J. Li, Z. Z. Zhang, and S. N. Ni (2014), Long-term groundwater variations in Northwest India from satellite gravity measurements, *Global Planet. Change*, *116*, 130–138.

Chen, J. L., C. R. Wilson, B. D. Tapley, B. R. Scanlon, and A. Guentner (2016), Long-term groundwater storage change in Victoria, Australia from satellite gravity and in situ observations, *Global Planet. Change*, *139*, 56–65.

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

Famiglietti, J. S., M. Lo, S. L. Ho, J. Bethune, K. J. Anderson, T. H. Syed, S. C. Swenson, C. R. de Linage, and M. Rodell (2011), Satellites measure recent rates of groundwater depletion in California's Central Valley, *Geophys. Res. Lett.*, *38*, L03403, doi:10.1029/2010GL046442.

Feng, W., M. Zhong, J. M. Lemoine, R. Biancale, H. T. Hsu, and J. Xia (2013), Evaluation of groundwater depletion in North China using the Gravity Recovery and Climate Experiment (GRACE) data and ground-based measurements, *Water Resour. Res.*, *49*, 2110–2118, doi:10.1002/wrcr.20192.

Geruo, A., J. Wahr, and S. J. Zhong (2013), Computations of the viscoelastic response of a 3-D compressible Earth to surface loading: An application to glacial isostatic adjustment in Antarctica and Canada, *Geophys. J. Int.*, *192*(2), 557–572, doi:10.1093/gji/ggs030.

Hirsch, R. M., and J. R. Slack (1984), A nonparametric trend test for seasonal data with serial dependence, *Water Resour. Res.*, *20*, 727–732.

Humphrey, V., L. Gudmundsson, and S. I. Seneviratne (2016), Assessing global water storage variability from GRACE: Trends, seasonal cycle, subseasonal anomalies and extremes, *Surv. Geophys.*, *37*(2), 357–395.

Landerer, F. W., and S. C. Swenson (2012), Accuracy of scaled GRACE terrestrial water storage estimates, *Water Resour. Res.*, *48*, W04531, doi:10.1029/2011WR011453.

Lawrence, D. M., et al. (2011), Parameterization improvements and functional and structural advances in version 4 of the community land model, *J Adv. Model. Earth Syst.*, *3*. M03001, doi:10.1029/2011MS000045.
Leblanc, M. J., P. Tregoning, G. Ramillien, S. O. Tweed, and A. Fakes (2009), Basin-scale, integrated observations of the early 21st century multiyear drought in southeast Australia, *Water Resour. Res.*, *45*, W04408, doi:10.1029/2008WR007333.
Long, D., B. R. Scanlon, L. Longuevergne, A. Y. Sun, D. N. Fernando, and H. Save (2013), GRACE satellite monitoring of large depletion in water storage in response to the 2011 drought in Texas, *Geophys. Res. Lett.*, *40*, 3395–3401, doi:10.1002/grl.50655.
Long, D., L. Longuevergne, and B. R. Scanlon (2015), Global analysis of approaches for deriving total water storage changes from GRACE satellites, *Water Resour. Res.*, *51*, 2574–2594, doi:10.1002/2014WR016853.
Long, D., et al. (2016), Have GRACE satellites overestimated groundwater depletion in the Northwest India Aquifer?, *Sci. Rep.*, *6*, 24398.
Longuevergne, L., B. R. Scanlon, and C. R. Wilson (2010), GRACE hydrological estimates for small basins: Evaluating processing approaches on the High Plains Aquifer, USA, *Water Resour. Res.*, *46*, W11517, doi:10.1029/2009WR008564.
Luthcke, S. B., T. J. Sabaka, B. D. Loomis, A. A. Arendt, J. J. McCarthy, and J. Camp (2013), Antarctica, Greenland and Gulf of Alaska land-ice evolution from an iterated GRACE global mascon solution, *J. Glaciol.*, *59*(216), 613–631.
Luthcke, S. B., D. D. Rowlands, T. J. Sabaka, B. D. Loomis, M. Horwath, and A. A. Arendt (2015), Gravimetry measurements from space, in *Remote Sensing of the Cryosphere*, edited by M. Tedesco, pp. 231–247, Wiley-Blackwell, Oxford.
Menemenlis, D., J.-M. Campin, P. Heimbach, C. Hill, T. Lee, A. Nguyen, M. Schodlok, and H. Zhang (2008), ECCO2: High resolution global ocean and sea ice data synthesis, *Mercator Ocean Quart. Newsl.*, *31*, 13–21.
Milzow, C., L. Kgotlhang, P. Bauer-Gottwein, P. Meier, and W. Kinzelbach (2009), Regional review: The hydrology of the Okavango Delta, Botswana-processes, data and modelling, *Hydrogeol. J.*, *17*(6), 1297–1328.
Oki, T., and Y. C. Sud (1998), Design of Total Runoff Integrating Pathways (TRIP): A global river channel network, *Earth Interact.*, *2*, 1–37.
Ramillien, G., F. Frappart, and L. Seoane (2014), Application of the regional water mass variations from GRACE satellite gravimetry to large-scale water management in Africa, *Remote Sens.*, *6*(8), 7379–7405.
Reager, J. T., and J. S. Famiglietti (2013), Characteristic mega-basin water storage behavior using GRACE, *Water Resour. Res.*, *49*, 3314–3329, doi:10.1002/wrcr.20264.
Richey, A. S., B. F. Thomas, M.-H. Lo, J. T. Reager, J. S. Famiglietti, K. Voss, S. Swenson, and M. Rodell (2015), Quantifying renewable groundwater stress with GRACE, *Water Resour. Res.*, *51*, 5217–5238, doi:10.1002/2015WR017349.
Rodell, M., et al. (2004), The global land data assimilation system, *Bull. Am. Meteorol. Soc.*, *85*(3), 381–394.
Rodell, M., I. Velicogna, and J. S. Famiglietti (2009), Satellite-based estimates of groundwater depletion in India, *Nature*, *460*(7258), 999–1002.
Rowlands, D. D., S. B. Luthcke, J. J. McCarthy, S. M. Klosko, D. S. Chinn, F. G. Lemoine, J. P. Boy, and T. J. Sabaka (2010), Global mass flux solutions from GRACE: A comparison of parameter estimation strategies-Mass concentrations versus Stokes coefficients, *J. Geophys. Res.*, *115*, B01403, doi:10.1029/2009JB006546.
Sakumura, C., S. Bettadpur, and S. Bruinsma (2014), Ensemble prediction and intercomparison analysis of GRACE time-variable gravity field models, *Geophys. Res. Lett.*, *41*, 1389–1397, doi:10.1002/2013GL058632.
Sakumura, C., S. Bettadpur, H. M. Save, and C. McCullough (2016), High frequency terrestrial water storage signal capture via a regularized sliding window mascon product from GRACE, *J. Geophys. Res. Solid Earth*, *121*, 4014–4030, doi:10.1002/2016JB012843.
Save, H., S. Bettadpur, and B. D. Tapley (2012), Reducing errors in the GRACE gravity solutions using regularization, *J. Geod.*, *86*(9), 695–711.
Save, H., S. Bettadpur, and B. D. Tapley (2016), High resolution CSR GRACE RL05 mascons, *J. Geophys. Res. Solid Earth*, *121*, 7547–7569, doi: 10.1002/2016JB013007.
Scanlon, B. R., L. Longuevergne, and D. Long (2012), Ground referencing GRACE satellite estimates of groundwater storage changes in the California Central Valley, USA, *Water Resour. Res.*, *48*, W04520, doi:10.1029/2011WR011312.
Schrama, E. J. O., B. Wouters, and R. Rietbroek (2014), A mascon approach to assess ice sheet and glacier mass balances and their uncertainties from GRACE data, *J. Geophys. Res. Solid Earth*, *119*, 6048–6066, doi:10.1002/2013JB010923.
Strassberg, G., B. R. Scanlon, and M. Rodell (2007), Comparison of seasonal terrestrial water storage variations from GRACE with groundwater-level measurements from the High Plains Aquifer (USA), *Geophys. Res. Lett.*, *34*, L14402, doi:10.1029/2007GL030139.
Swenson, S., and J. Wahr (2006), Post-processing removal of correlated errors in GRACE data, *Geophys. Res. Lett.*, *33*, L08402, doi: 10.1029/2005GL025285.
Swenson, S., J. Famiglietti, J. Basara, and J. Wahr (2008), Estimating profile soil moisture and groundwater variations using GRACE and Oklahoma Mesonet soil moisture data, *Water Resour. Res.*, *44*, W01413, doi:10.1029/2007WR006057.
Tapley, B. D., S. Bettadpur, M. Watkins, and C. Reigber (2004), The gravity recovery and climate experiment: Mission overview and early results, *Geophys. Res. Lett.*, *31*, L09607, doi:10.1029/2004GL019920.
Thomas, A. C., J. T. Reager, J. S. Famiglietti, and M. Rodell (2014), A GRACE- based water storage deficit approach for hydrological drought characterization, *Geophys. Res. Lett.*, *41*, 1537–1545, doi:10.1002/2014GL059323.
Trabucco, A., and R. J. Zomer (2009), Global aridity index (global-aridity) and global potential evapotranspiration (global-PET) geospatial database, CGIAR-CSI, Washington, D. C. [Available at http://www.csi.cgiar.org/.]
Velicogna, I., T. C. Sutterley, and M. R. van den Broeke (2014), Regional acceleration in ice mass loss from Greenland and Antarctica using GRACE time-variable gravity data, *Geophys. Res. Lett.*, *41*, 8130–8137, doi:10.1002/2014GL061052.
Wahr, J., M. Molenaar, and F. Bryan (1998), Time variability of the Earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE, *J. Geophys. Res.*, *103*, 30,205–30,229.
Wahr, J., S. Swenson, and I. Velicogna (2006), Accuracy of GRACE mass estimates, *Geophys. Res. Lett.*, *33*, L06401, doi:10.1029/2005GL025305.
Watkins, M. M., D. N. Wiese, D. N. Yuan, C. Boening, and F. W. Landerer (2015), Improved methods for observing Earth's time variable mass distribution with GRACE using spherical cap mascons, *J. Geophys. Res. Solid Earth*, *120*, 2648–2671, doi:10.1002/2014JB011547.
Wiese, D. N., F. W. Landerer, and M. M. Watkins (2016), Quantifying and reducing leakage errors in the JPL RL05M GRACE mascon solution, *Water Resour. Res.*, *52*, 7490–7502, doi:10.1002/2016WR019344.
Xavier, L., M. Becker, A. Cazenave, L. Longuevergne, W. Llovel, and O. C. Rotunno Filho (2010), Interannual variability in water storage over 2003-2008 in the Amazon Basin from GRACE space gravimetry, in situ river level and precipitation data, *Remote Sens. Environ.*, *114*(8), 1629–1637.