**Natural and human-induced terrestrial water storage change: A global analysis using hydrological models and GRACE**

Farshid Felfelani[1,*], Yoshihide Wada[2,3], Laurent Longuevergne[4], Yadu N. Pokhrel[1]

[1] Department of Civil and Environmental Engineering, Michigan State University, East Lansing, MI, USA

[2] International Institute for Applied Systems Analysis, Schlossplatz 1 - A-2361, Laxenburg, Austria

[3] Department of Physical Geography, Utrecht University, Domplein 29, 3512 JE, Utrecht, The Netherlands

[4] Géosciences Rennes, Université de Rennes, Rennes Cedex, France

*Corresponding author: Farshid Felfelani

Email: felfelan@msu.edu

Keywords: Terrestrial water storage; Hydrological models; GRACE; Human impacts

**Abstract**

Hydrological models and the data derived from the Gravity Recovery and Climate Experiment (GRACE) satellite mission have been widely used to study the variations in terrestrial water storage (TWS) over large regions. However, both GRACE products and model results suffer from inherent uncertainties, calling for the need to make a combined use of GRACE and models to examine the variations in total TWS and their individual components, especially in relation to natural and human-induced changes in the terrestrial water cycle. In this study, we use the results from two state-of-the-art hydrological models and different GRACE spherical harmonic products to examine the variations in TWS and its individual components, and to attribute the changes to natural and human-induced factors over global river basins. Analysis of the spatial patterns of the long-term trend in TWS from the two models and GRACE suggests that both models capture the GRACE-measured direction of change, but differ from GRACE as well as each other in terms of the magnitude over different regions. A detailed analysis of the seasonal cycle of TWS variations over 30 river basins shows notable differences not only between models and GRACE but also among different GRACE products and between the two models. Further, it is found that while one model performs well in highly-managed river basins, it fails to reproduce the GRACE-observed signal in snow-dominated regions, and vice versa. The isolation of natural and human-induced changes in TWS in some of the managed basins reveals a consistently declining TWS trend during 2002-2010, however; significant differences are again obvious both between GRACE and models and among different GRACE products and models. Results from the decomposition of the TWS signal into the general trend and seasonality indicate that both models do not adequately capture both the trend and seasonality in the managed or snow-dominated basins implying that the TWS variations from a single model cannot be reliably used for all global regions. It is also found that the uncertainties arising from climate forcing datasets can introduce significant additional uncertainties, making direct comparison of model results and GRACE products even more difficult. Our results highlight the need to further improve the representation of human land-water management and snow processes in large-scale models to enable a reliable use of models and GRACE to study the changes in freshwater systems in all global regions.

## 1. Introduction

The question of how freshwater systems are changing under the dual influence of climate variability and increasing human water exploitation has been a topic of great concern and debate in the face of growing water scarcity around the world (Alley et al., 2002; Famiglietti, 2014; Fan, 2015; Gleeson et al., 2012). Ground-based monitoring of surface water and groundwater (GW) systems suggests profound changes in surface water flows and GW storages globally due to accelerating human alteration of land and water systems (Giordano, 2009; Scanlon et al., 2012a) which can be both direct, e.g., flow regulation and groundwater pumping and indirect, e.g., changes in climate forcing, $CO_2$ concentrations and impacts on photosynthetic activities (Trancoso et al., 2017). However, the lack of in-situ observations worldwide limits our understanding of the dynamic relationship between natural climate variability and direct and indirect human impacts (HI) on freshwater systems (Alley et al., 2002; Döll et al., 2016; Taylor et al., 2013). Large-scale hydrological models play an irreplaceable role in filling this data gap and provide an improved understanding of the changes in the water cycle, which is crucial for the accurate assessment and realistic prediction of water availability and use. In recent years, satellite-based observations of water flows and storages have substantially advanced our ability to better monitor the changing water systems at the global scale. In particular, the combined use of the satellite data and hydrological models has revolutionized the way we study global freshwater systems (Dijk and Renzullo, 2011; Famiglietti et al., 2015).

Large-scale hydrological models have been widely used to study global freshwater systems and human water use (Nazemi and Wheater, 2015; Pokhrel et al., 2016). These models can be classified into two general types: (i) land surface models (LSMs) and (ii) global hydrological models (GHMs) (Haddeland et al., 2011). LSMs, such as the MATSIRO (Takata et al., 2003) and CLM (Lawrence et al., 2011), are designed to simulate the land hydrology within the general circulation models (GCMs) and Earth system models (ESMs), but GHMs, such as the WaterGAP (Alcamo et al., 2003; Döll et al., 2003) and PCR-GLOBWB (van Beek et al., 2011; Wada et al., 2010), have been traditionally developed as stand-alone models for offline water resource assessment. While LSMs simulate various hydrological processes on a physical basis and solve both surface water and energy balances at the land surface, GHMs simulate these processes using relatively simple and conceptual approaches even though they are more comprehensive in simulating human land-water management practices (Pokhrel et al., 2016). As such, LSMs and

GHMs have certain limitations in simulating the natural or human-induced changes in various branches of the water cycle. In particular, despite noteworthy progress that has been made in model improvements over the years (Overgaard et al., 2006; Pitman, 2003; Sellers et al., 1997), water table dynamics and GW pumping still remain largely ignored or poorly simulated (Nazemi and Wheater, 2015; Pokhrel et al., 2016), making the models incapable of accurately capturing subsurface water flows and storages in general, and the human-induced GW storage depletion in particular. While the hydrological fluxes such as river discharge can be simulated with relatively high accuracy either by calibrating the model with observations (Döll et al., 2003) and/or by employing lumped routing schemes to explicitly simulate shallow GW flows (Kim et al., 2009), these approaches do not guarantee the correct simulation of soil moisture and GW storage. Moreover, the uncertainties arising from these deficiencies in model parameterizations can be further amplified by the uncertainties in meteorological forcing datasets used to drive these models (Decharme and Douville, 2006).

Advances in satellite observations have enabled us to address some of the challenges in using hydrological models for large-scale hydrological studies (Pail et al., 2015). For example, the assimilation of terrestrial water storage (TWS) derived from the Gravity Recovery and Climate Experiment (GRACE) satellite mission into LSMs has been used to improve global simulation of TWS and its components by model calibration and assimilation techniques (Chen et al., 2017; Eicker et al., 2014; Girotto et al., 2016; Houborg et al., 2012; Li et al., 2012; Li and Rodell, 2015; Zaitchik et al., 2008) and to quantify the changes in certain variables that are not explicitly simulated by the models (e.g., GW storage) (Castellazzi et al., 2016; Famiglietti et al., 2011; Feng et al., 2013; Jin and Feng, 2013; Long et al., 2016; Nanteza et al., 2016; Rodell et al., 2009; Scanlon et al., 2012b). GRACE data has also been extensively used to benchmark the accuracy of hydrological model simulations (Alkama et al., 2010; Decharme et al., 2010; Döll et al., 2014; Eicker et al., 2016; Freedman et al., 2014; Grippa et al., 2011; Landerer et al., 2010, 2013; Rosenberg et al., 2013; Swenson and Lawrence, 2015; Xie et al., 2012; Yang et al., 2011); conversely, LSMs have also proved useful to evaluate the performance of different GRACE products and processing methods (Klees et al., 2008; Werth et al., 2009) and used as a priori information to restore signal attenuation and leakage errors arising from the low spatial resolution of GRACE (Landerer and Swenson, 2012; Long et al., 2015a, 2015b).

The GRACE and hydrological models complement each other to better constrain the different components on the water cycle; however, GRACE products are affected by various limitations and uncertainties. First, it provides a large-scale estimate of vertically integrated water storage variations, limiting safe interpretation to relatively large regions (>200,000 km2) (Longuevergne et al., 2010). Second, GRACE products are affected by latitude-dependent uncertainties with higher uncertainties in mid and low latitudes compared to the poles (Wahr et al., 2006). Moreover, varying uncertainties can be found even among different GRACE solutions i.e., spherical harmonic (SH) products and mascons (Long et al., 2017; Scanlon et al., 2016; Watkins et al., 2015) which vary across different global regions.

GRACE measures the vertically integrated TWS variations caused by both natural and anthropogenic drivers. Therefore, hydrological models or other supplementary data are required to disintegrate the total TWS into separate components and to partition it into the natural and human-induced changes. For example, Human-induced TWS variations are estimated by computing the difference between GRACE that includes the human factors and hydrological models that simulate only the natural part of the water cycle (Huang et al., 2015; Pan et al., 2016). Some other studies have used GRACE-based TWS variations and observed or simulated surface water storage variations to derive GW storage change in depleted aquifer systems where in some cases, the GRACE-detected TWS signature is mostly due to human-induced GW storage change (Famiglietti et al., 2011; Rodell et al., 2009; Scanlon et al., 2012b) and in some cases it is due to specific climatic events such as climate variability or droughts (Russo and Lall, 2017; Scanlon et al., 2015). Although these approaches are useful for extracting human-induced TWS variations from models that do not account for human activities, they can involve significant uncertainties arising from the errors and uncertainties in two independent products (GRACE and models). The recent advancements in representing human activities in models (e.g., Pokhrel et al., 2016) provide the opportunity to directly isolate the human-induced TWS variations from models (e.g., Pokhrel et al., 2017) and compare the results with GRACE-based approaches.

Given the above background, we use multiple GRACE SH products and results from two hydrological models (one LSM and one GHM) to examine the spatio-temporal patterns of TWS variations and the uncertainties arising from the use of different GRACE products and hydrological models. To limit the propagation of some GRACE errors, we use the strategy to filter model output as GRACE before performing a comparison. Both models explicitly simulate the human-induced changes in TWS, including the changes in GW storage due to pumping, making the results directly comparable with GRACE. A detailed analysis is presented for the selected river basins located in different geographic regions and having different extent of human alterations in terms of flow regulation and GW use. Results from the simulation with natural settings (without considering human factors) are then used in conjunction with GRACE data to isolate the human-induced TWS variations from the total TWS change measured by GRACE. Our specific objectives are to: (1) examine the global spatial patterns in TWS variations over different river basins, especially by quantifying the contribution of different components to the total TWS variations; (2) carry out a temporal comparison among multiple GRACE SH products and two models and attribute the TWS variations to climate and human-induced factors in the basins where human land-water management has largely altered the terrestrial water balance; and (3) quantify the uncertainties in simulated TWS caused by the use of different sets of meteorological forcing data. These objectives provide the structural sub-headings used in the Methods, Results, and Discussion sections.

## 2. Models and Data

### 2.1 Models

We use two state-of-the-art hydrological models, namely the HiGW-MAT, a LSM (Pokhrel et al., 2015) and the PCR-GLOBWB, a GHM (Wada et al., 2014) to simulate the global terrestrial water fluxes and storages (excluding Antarctica and Greenland). Both models simulate the natural and human-induced changes in flows and storage of water, explicitly taking into account GW abstractions and the resulting changes in subsurface storage, which is crucial to realistically simulate the variations of TWS in regions with intensive GW mining. However, the two models use different GW representations; while PCR-GLOBWB simulates the GW storage as a linear reservoir model without explicitly representing water table dynamics, HiGW-MAT uses a more sophisticated GW scheme that explicitly simulates the water table dynamics. A detailed description of both models can be found in our earlier works (Pokhrel et al., 2015; Wada et al., 2014) but for completeness, we provide a brief summary of the models below.

The HiGW-MAT model is based on the Minimal Advanced Treatment of Surface Interactions and Runoff (MATSIRO) (Takata et al., 2003) LSM. In MATSIRO, effects of vegetation on the surface energy balance are calculated on the basis of the multi-layer canopy model of Watanabe (1994) and the photosynthesis-stomatal conductance model of Collatz et al. (1991). The vertical movement of soil moisture is estimated by numerically solving the Richards equation (Richards, 1931) for the soil layers in the unsaturated zone. Surface and subsurface runoff parameterizations are based on the simplified TOPMODEL (Beven and Kirkby, 1979; Stieglitz et al., 1997). In our recent studies, we enhanced MATSIRO by first representing HI schemes such as reservoir operation and irrigation (Pokhrel et al., 2012a, 2012b) and then GW pumping (Pokhrel et al., 2015), resulting in the latest development called the HiGW-MAT.

In HiGW-MAT, irrigation is simulated by using a soil moisture deficit based scheme described in Pokhrel et al. (2012a). Gridded irrigated areas are based on the Pokhrel et al. (2012a). The pumping scheme described in Pokhrel et al. (2015) explicitly simulates the amount of water withdrawn from aquifer and the associated changes in GW storage. The water table dynamics is simulated by using the scheme of Koirala et al. (2014). All soil and vegetation parameters and land cover data are prescribed based on the Global Soil Wetness Project 2 (GSWP2) (Dirmeyer et al., 2006). Subgrid variability of vegetation is represented by partitioning each grid cell into two tiles: natural vegetation and irrigated cropland. The crop growth module, based on the crop vegetation formulations and parameters of the Soil and Water Integrated Model (SWIM) (Krysanova et al., 1998), estimates the growing period necessary to obtain mature and optimal total plant biomass for 18 different crop types. The leaf area index (LAI) is resolved according to Hirabayashi et al. (2005). Surface runoff is routed through the river network using the Total Runoff Integrating Pathways (TRIP) (Oki and Sud, 1998). The reservoir operation is based on Hanasaki et al. (2006). Data for large and medium-sized reservoirs are same as in Pokhrel et al. (2012a), which account for the majority of dams having a height of 15m or more.

The original MATSIRO and the HI schemes in HiGW-MAT have been extensively validated using observed river discharge, TWS, irrigation water withdrawals, GW pumping, and water table depth (Koirala et al., 2014; Pokhrel et al., 2012a, 2012b, 2015; Zhao et al., 2017). The results of evapotranspiration (ET) have not been validated due to the lack of reliable global ET products, but as in any typical global model, the underlying assumption is that since the models are forced by observed meteorological data and they perform reasonably well in reproducing river flow, ET simulations are also reasonable.

PCR-GLOBWB is an offline GHM that simulates the interaction of surface water and subsurface water through the atmosphere, land surface, two vertically stacked soil layers and an explicit underlying GW reservoir that is represented as a linear reservoir model (Kraijenhoff Van De Leur, 1958). PCR-GLOBWB explicitly simulates the water demands for agriculture, industry and households, and associated use from different water sources. The irrigation water requirement including the losses is calculated for paddy and nonpaddy crops based on the MIRCA2000 dataset (Portmann et al., 2010). The irrigation scheme is dynamically linked to the surface and subsurface hydrology schemes to provide a more realistic soil moisture content and ET over irrigated croplands (Wada et al., 2014). Other water demands including livestock, industry and domestic are calculated based on various available socio-economic data and country statistics including livestock densities, GDP, electricity production, energy consumption, and population (Wada et al., 2014).

The vegetation and land cover are parameterized according to the Global Land Cover Characteristics Data Base version 2.0 (GLCC 2.0; https://lta.cr.usgs.gov/glcc/globdoc2_0#avhrr) and the Land Surface Parameter dataset (LSP2) (Hagemann, 2002). Soil properties are obtained from the vector-based FAO Digital Soil Map of the World (DSMW) (FAO, 2003) and the ISRIC-WISE global dataset of derived soil properties (Batjes, 2005). Using Simulated Topological Network (STN30) (Vörösmarty et al., 2000), surface and subsurface runoff are routed along the river network. The Global Reservoir and Dam database (GRanD) (Lehner et al., 2011) is used to locate the reservoirs on the river network based on the construction year. Reservoir regulation and release is simulated based on Hanasaki et al. (2006) and van Beek et al. (2011) to satisfy downstream water demands (Wada et al., 2010, 2014). The PCR-GLOBWB model is also validated with the observations of river discharge and runoff, TWS, irrigation water requirement, and GW withdrawal (van Beek et al., 2011; Wada et al., 2014).

### 2.2 Climate Forcing

We use forcing data from multiple sources. HiGW-MAT is driven by three forcing datasets: (1) the WFDEI (WATCH Forcing Data methodology applied to ERA-Interim reanalysis data) (Weedon et al., 2014), (2) the forcing data from Princeton University (Sheffield et al., 2006), and (3) the JRA-25 atmospheric reanalysis data provided by Japanese Meteorological Agency (JMA) Climate Data Assimilation System (JCDAS) (Kim et al., 2009; Onogi et al., 2007). The results from the third forcing data, which are validated in our previous studies, are used for the analysis of TWS, and the other two datasets are used to examine the uncertainty arising from the climate forcing data (see Section 3.3). PCR-GLOBWB is forced only by WFDEI data and is not considered for uncertainty analysis.

### 2.3 GRACE Data

The GRACE data along with model results are used to analyze the TWS variations. We use different level-3 SH-based GRACE products of equivalent water height (EWH) from three processing centers, namely: (i) the Center for Space Research (CSR) at University of Texas at Austin, (ii) Jet Propulsion Laboratory (JPL) at California Institute of Technology, and (iii) the German Research Center for Geoscience (GFZ) (available for download from JPL website; http://grace.jpl.nasa.gov/data/get-data/) for model evaluation and to characterize the uncertainty within the three GRACE products. In general, while the three official products (CSR, JPL, and GFZ) underestimate GRACE uncertainties (Sakumura et al., 2014), they provide a fair estimate to evaluate hydrological models. The GRACE satellite level 2 data processing delivers the dimensionless Stokes' coefficients ($C_{lm}$ and $S_{lm}$) complete to degree and order 96 ($l = m = 96$). Corrections and adjustments are needed to reduce noises and isolate the TWS changes from other signals visible in GRACE. The GRACE data from aforementioned sources already carry corrections and filtering including atmospheric mass changes removal, glacial isostatic adjustment (GIA), truncation of SH coefficients at degree 60, and application of destriping filter alongside with a 300-km Gaussian smoother.

**Deleted:** solutions

It is important to consider observational errors when using GRACE data to evaluate models. The GRACE error budget can be separated into three types (Longuevergne et al., 2010): (1) errors associated with fundamental GRACE measurements satellite to satellite range rate (~5 mm EWH), (2) errors in atmospheric and oceanic corrections (~10 to 20 mm EWH) and (3) bias and leakage correction errors which can be the largest depending on basin area and context (~30 mm EWH for a 200,000 km² basin). In this work, rescaling factors are not used and the model results are filtered as GRACE to compare at an equivalent resolution and avoid type (3) errors. This method has been highlighted as a robust approach for model evaluation (Güntner, 2008; Xie et al., 2012).

## 3. Methods

### 3.1. Spatial Patterns in TWS Variations and Contribution of Different Components

We use the results from the fully coupled versions of both models (i.e., by activating all human impacts schemes) to evaluate the model performance in capturing the spatial variability in TWS rates measured by GRACE. For consistent comparison with GRACE data, the spatial map of simulated TWS rates from both models is transformed into SH domain, truncated at degree and order 60, and smoothed by the 300-km Gaussian filter, following Wahr et al. (1998). The spatial filtering process reduces the errors and noises together with the true signals. Different approaches (e.g., scaling factor approach and the additive correction approach) have been proposed to restore the true signal losses (Landerer and Swenson, 2012; Long et al., 2015a, 2015b). Using the same filtering processes for model outputs, as used for GRACE products, offsets the necessity for reconstructing the attenuated signals when directly comparing the GRACE and simulated TWS (Landerer and Swenson, 2012).

Deleted: have been

Additionally, understanding how different storage compartments (i.e., snow and ice, soil water, river water, and GW) contribute to the variations of total TWS is crucial to investigate how the changes in these individual compartments can potentially affect the availability and utilization of water resources. Isolation of the individual components also provides key insights on the interactions and feedback among different components under changing hydrologic regime. Here, we use a dimensionless metric called the component contribution ratio (CCR) proposed by Kim et al. (2009) to determine the role of different TWS components in modulating the total TWS variations in river basins from different climate regions. The ratio is calculated as:

$$CCR = \frac{MAD}{TV} \qquad (1)$$

where MAD is the mean absolute deviation of a TWS component ($\frac{1}{N}\sum_t^N |S_t - \bar{S}|$, $S_t$ is the value of component $S$ at time $t$ and $N$ is the number of months), TV is the total variability and is calculated as summation of all components MADs ($\sum_{i=S}^{components} MAD_i$). The CCR values are calculated by using HiGW-MAT model results.

### 3.2. Temporal Variability of TWS in Global Basins: Human-induced TWS Change

We make an integrated use of GRACE data and models to examine the temporal variability of TWS over the selected global river basins, and isolate the human-induced TWS change. To estimate basin-scale water storage, a simple basin function (which has the value 1 for inside the basin and 0 outside) is used. The function is then multiplied by different model and GRACE signals to form the basin scale water storage. Since the data are in 1 degree resolution with varying grid cell area, an area-weighted arithmetic mean is finally calculated as:

$$H(x,t) = \frac{\sum_{i=1}^{n} S_i(x,t)}{A}, \;\; S_i(x) = \begin{cases} 1 \times s \times a_i & \textit{inside the basin} \\ 0 & \textit{outside the basin} \end{cases} \qquad (2)$$

where $s$ is the LSM or GRACE estimate, $a_i$ is the cell area, $S_i$ is the weighted estimate for each cell inside the basin, $n$ is the number of cells in a basin, $A$ is the total area of the basin, and $H(x,t)$ represents the estimate of water storage for basin at time $t$.

We quantify the human-induced TWS change using GRACE and hydrological models in some of the basins affected by human activities. First, we estimate the long-term linear trend in TWS from GRACE observations, PCR-GLOBWB, and HiGW-MAT (simulations with HI). Then, we estimate the similar trend using the model results from the simulation with natural setting in which all HI schemes are deactivated. We then calculate the difference between the two trends as an estimate of the direct human-induced changes in TWS. To estimate the variations in monthly TWS from model results, we use two different approaches. First, for simulations with HI, we directly integrate the individual TWS components (i.e., snow water, canopy water, river water, soil moisture, and groundwater). Due to explicit representations of human activities in both HiGW-MAT and PCR-GLOBWB, all TWS components are explicitly simulated, also taking into account the impacts of human activities. In this approach, the vertically integrated TWS is expressed as:

$$TWS = SW + SnW + SM + GW + CW \qquad (3)$$

where, $SW$, $SnW$, $SM$, $GW$, and $CW$ denote surface water, snow water, soil moisture, groundwater, and canopy water storages (all terms have the dimension $[L]$), respectively. The changes in storage terms (Equation 3) include GW storage and water table changes due to pumping; changes in surface water reservoirs, and changes in soil moisture due to human water management (e.g., irrigation).

Second, for the simulation with natural setting, we use the water balance approach (Famiglietti et al., 2011; Nanteza et al., 2016; Rodell et al., 2004; Syed et al., 2008; Zeng et al., 2008) in which the TWS change is deduced from monthly precipitation (P), evapotranspiration (ET), and runoff (R) as:

$$\frac{dTWS}{dt} = P - ET - R \qquad (4)$$

where, $P$ is the observed precipitation, $ET$ is the simulated actual evapotranspiration, and $R$ is the simulated runoff (all terms have the dimension $[LT^{-1}]$). Equation 4 can be used over large river basins and long-term simulation period with the assumption of no lateral GW fluxes in the boundaries (Long et al., 2017). However, we use the water balance method only for the simulation with natural setting (and not for HI simulations) due to high uncertainties in flux variables, particularly in $ET$ and $R$ (Long et al., 2014, 2017; Wang et al., 2015b) that are strongly influenced by HI such as irrigation, surface water flow regulation, and $GW$ storage change due to pumping. While we use Equation 3 to derive the TWS from model simulations with all HI schemes activated which is used for model evaluation with GRACE, the TWS estimated by using Equation 4 (based on HiGW-MAT model) is combined with GRACE data to isolate the human-induced TWS variations in the highly-managed river basins.

To better investigate the performance of models in TWS simulations, we decompose the observation data and simulated time series into general trend and seasonality using moving averages and applying convolution filter. In the decomposition progress, the data ($Y[t]$) is disaggregated into general trend ($T[t]$), seasonality ($S[t]$), and residuals ($e[t]$) to form the additive model: $Y(t) = T(t) + S(t) + e(t)$.

### 3.3. The Uncertainty from Climate Forcing Data

We examine the uncertainty in the simulated TWS by using different forcing datasets listed in Section 2.2. For this purpose, we use only the HiGW-MAT model which is driven by the three forcing datasets. Among the three datasets, we use the data from Kim et al. (2009) to derive the TWS used for the spatio-temporal analysis, including the comparison with the results from PCR-GLOBWB model which is driven by the WFDEI data, and the estimation of CCR because the same data has been used in our previous model validation studies (Pokhrel et al., 2012a, 2012b, 2015). The other two datasets are then used to examine the uncertainties in simulated TWS that are caused by the use of different forcing data. We did so to ensure that the HiGW-MAT simulations used to derive the key conclusion are well-validated before.

The results from the uncertainty analysis are not directly compared with GRACE and so, we present the gridded scaling factors to account for the signal loss caused by filters and smoothers. The scaling factors that also referred as multiplicative factors are derived from the least squares

Deleted:

fit (Equation 5) between the gridded filtered and unfiltered TWS changes from the HiGW-MAT model (see Landerer and Swenson, 2012 and Long et al., 2015a for details).

$$M = \sum_T (S_t - kS_f)^2 \qquad (5)$$

where, $M$ is the objective function to be minimized, $S_t$ is the true signal (model output), $S_f$ is the filtered signal, $T$ is the time steps (here, months in 2002-2008), and $k$ is the scaling factor.

## 4. Results

### 4.1. Spatial Patterns in TWS Variations and Contribution of Different Components

We first evaluate the spatial variability of the long-term trend in total TWS variations simulated by the two models with GRACE (the mean of CSR, JPL, and GFZ) TWS trend (Figure 1). Due to high susceptibility of the linear trend to the selection of time window, we use the 2002-2008 period that represents high diversity in signal patterns with relatively distinct spatial variations in positive and negative trends among natural and human-affected global regions, especially the downward TWS trends due to GW depletion. Overall, a good agreement can be seen between GRACE (Figure 1a), and both HiGW-MAT (Figure 1b), and PCR-GLOBWB (Figure 1c) models in terms of the direction of change; however, significant discrepancies are also apparent in terms of the magnitude. For example, the global hotspots of GW depletion such as the northwestern India and parts of Pakistan, the North China Plain, and parts of Middle East (where the changes in total TWS are known to be dominated by GW storage change) are detected in both GRACE and models but the magnitude of changes varies largely among the three estimates. In northwest India, clear differences can be seen; while GRACE data suggest a small downward trend, HiGW-MAT suggests a much larger TWS depletion and PCR-GLOBWB shows little change. In California Central Valley, HiGW-MAT simulates a larger decrease in TWS compared to the other two estimates, which is likely due to overestimation of GW pumping as suggested by Pokhrel et al. (2015). The performance of PCR-GLOBWB is generally good in many of these regions that are affected by human activities but it doesn't reproduce the GRACE-detected negative trends in parts of southeastern Australia and northeastern China.

In some of the regions with relatively low human influence such as the Amazon, Orinoco, and Parana river basins in South America and southern parts of Africa, significant variations are obvious among the models and GRACE both in the sign and magnitude. In the Amazon and

**Deleted:** northeastern China

Orinoco, the HiGW-MAT model captures the GRACE trend reasonably well while the PCR-GLOBWB shows a larger deviation. On the contrary, in the southern parts of Africa HiGW-MAT simulates a large positive trend while PCR-GLOBWB simulates a milder trend, consistent with GRACE. In the river basins in the northern high latitude such as the Yukon, GRACE detects a large negative TWS trend during 2002-2008 which has been suggested to be due to glacier melts, permafrost thaw, and snow cover shrinkage (Ge et al., 2013; Spence, 2002; St. Jacques and Sauchyn, 2009; Wang et al., 2015a), processes that are not explicitly simulated by both models.

# Figure 1 to be inserted here

The contribution of the individual storage components to total TWS is quantified for 30 river basins. The river basins are selected considering: (a) a wide coverage over different climatic regions and continents, and (b) a good balance between natural and human-affected regions. Figure 2 depicts the river basins along with the CCR calculated by using HiGW-MAT model results. The size of the circles is proportional to the seasonal amplitude of the total TWS variation, with the largest amplitude being 500 mm in the Orinoco river basin. Both models used in the study do not explicitly simulate glacier processes, so the surface water component includes only snow and river water. As expected, in the northern high latitudes and polar regions snow storage component dominates the TWS. The highest contribution of snow is found in the Yenisey (61%), Mackenzie (60%), Yukon (59%), Lena (54%), and OB (54%) river basins. Moving toward the mid-latitudes and the subtropical area, high snow storage is substituted by surface and subsurface storages. The highest contribution of surface water storage can be seen in the Yangtze (33%), Brahmaputra (28%), and Ganges (20%), all located in the subtropics and managed by large number of reservoirs (Lehner et al., 2011). Subsurface water storage dominatingly modulates the total TWS variations in the temperate and tropical regions such as the Niger (97%), Parana (90%), Tocantins (90%), and Congo (89%) river basins, and also in river basins with semi-arid climates such as the Murray–Darling (95%) and Euphrates (88%) basins. The contribution of subsurface water storage is also found to be large in the river basins with strong human influence, particularly in regions where excessive GW is used for irrigation (e.g., the Indus, Huang-He, Euphrates, and Murray-Darling basins).

# Figure 2 to be inserted here

**4.2. Temporal Variability of TWS in Global Basins: Human-induced TWS Change**

Figure 3 presents the seasonal cycle of TWS variations from GRACE, HiGW-MAT, and PCR-GLOBWB for the selected basins. We present the range of variations among the three SH solutions (CSR, JPL, and GFZ) as the gray-shaded band. In this figure, the basins have been classified into three categories, namely the natural, managed, and snow-dominated which are shown with white, yellow, and light-blue background, respectively. Similar to the spatial patterns of the long-term trend (Figure 1), a generally good agreement can be seen between GRACE products and models, especially in the basins with less human influence and snow contribution (white background). In some of the managed and snow-dominated basins such as the Huang-He (Yellow river), Amur, Murray-Darling, and Yukon the GRACE-model agreement is generally poor for both models. In the basins such as the Huang-He, Indus, Amur, Lena, Mackenzie, and Yukon notable difference between the two models are also obvious both in terms of the seasonal amplitude and timing of peak.

Also shown in Figure 3 are the individual TWS components (i.e., snow, river, soil, and GW storages) to scrutinize how different storage compartments modulate the total TWS signal in different geographic and climatic regions. For clarity of view we present these details only from the HiGW-MAT model. In many of the selected basins where the contribution of snow is relatively small, the seasonal TWS signal is strongly modulated by the variations in subsurface storage, which is governed by the inverse relationship between soil moisture and GW. These two components compete for the same storage space and thus evolve over time in opposite phase (Duffy, 1996; Pokhrel et al., 2013). Note that in HiGW-MAT, the soil moisture and GW are estimated as water stored above and below the water table depth, respectively, which is different than in typical global LSMs and GHMs that consider soil moisture to be the water stored within the fixed soil depth (typically top 1-2m) resulting in the same-phase relationship between soil moisture and groundwater storages, but with certain time lag. The dominance of surface water can be seen in basins such as the Ganges, Brahmaputra, and Mekong where the seasonal flood pulse transports large volume of water during the monsoon season. In snow-dominated basins such as the Mackenzie, Yenisey, and Yukon a strong seasonal signal of snow accumulation can be seen during the boreal spring which is followed by an increase in river water arising from snowmelt.

# Figure 3 to be inserted here

In figure 4, we provide further details on the inter-annual variability of TWS from different GRACE solutions (shown as shaded range) and both models along with the individual components from HiGW-MAT. All results are shown as anomalies relative to the 2004-2009 time-mean baseline to be consistent with GRACE. The simulated TWS from both expansions (Equation 3 and Equation 4) is truncated at degree and order 60 and smoothed by the 300-km Gaussian filter in all figures corresponding to GRACE products. In figure 4, the slopes of the trend lines from GRACE, models (with activated HI modules), and the water balance analysis (i.e., the simulation without human activities) are shown at the bottom of each panel. The $p$-value approach is used to measure the statistical significance of linear trends from GRACE and model outputs, i.e., to determine the probability of whether the simulated trends are non-zero and that is statistically significant (Zhou et al., 2014). Results indicate that the TWS trend in natural simulation, which is mostly close to zero, is not statistically significant ($p$ values > $0.05$) in most of the managed basins. Further, the $p$ values indicate that the PCR-GLOBWB trend for Euphrates, Indus, Murray-Darling, and Volga basins, the GRACE trend for Brahmaputra, Euphrates, Ganges, Indus, and Volga basins, and the HiGW-MAT trend for most of the managed basins are statistically significant ($p$ values < $0.05$).

For most of the managed river basins (except for the Colorado and Murray-Darling), the long-term negative trend in the total TWS is larger in GRACE solutions than in the results from water balance, suggesting that these basins experienced certain loss of water during the analysis period. The PCR-GLOBWB model mostly follows the GRACE trends in most river basins but the HiGW-MAT model suggests a substantially larger negative trend in TWS in the managed basins that is primarily due to the decline in GW storage (noticeable in the Indus and Huang-He basins). This also implies that the pumping scheme in HiGW-MAT may have overestimated GW pumping as discussed earlier in Figure 1. Colorado and Murray-Darling, show unexpected increase in GRACE TWS that represents smaller deficit rate than in the natural simulation. The positive trend in GRACE data in these basins is primarily due to some wet cycles (e.g., year 2005 and year 2010) in their long-term inter-annual variability of TWS. For instance, the precipitation increase in the wet year of 2010 in Murray-Darling basin and also the snow amount rise that is followed by two wet cycles around the years 2005 and 2010 in the Colorado basin resulted in such positive overall trends during 2002-2010. As such, if the wet cycles of 2005 and 2010 are excluded from the analysis, Murray-Darling and Colorado basins also show a significant TWS loss.

The largest difference between GRACE and natural trends can be seen in the Euphrates, a transboundary river basin between Iraq, Turkey, Jordan, and Saudi Arabia. While GRACE TWS regression line drops at rate of 2.13 $cm/yr$, only 0.06 $cm/yr$ of that is caused by natural variability, and the rest (2.07 $cm/yr$) is caused by direct HI. The Ganges river basin with the second largest divergence between the natural and GRACE trend lines also experiences a 1.99 $cm/yr$ human-induced TWS loss. For this basin, HiGW-MAT performs well especially in simulating the drought years (negative peaks). In the Indus, despite a relatively constant and positive precipitation trend as well as a small negative P-ET-R trend (0.01 $cm/yr$ of water storage loss), GRACE shows a larger drop in TWS that is 0.82 $cm/yr$. Clearly, this huge difference is due to the widely reported depletion of groundwater resources in part of the basin (Rodell et al., 2009; Tiwari et al., 2009). For river basins with considerable snow water component (distinguished by light blue background color), HiGW-MAT performs better. In particular, HiGW-MAT shows the seasonal variations consistent with GRACE (Figures 3 and 4) likely due to advanced energy balance scheme. In other basins that represent low human influence and small contribution from snow (e.g., Amazon, Danube, and Niger), both models simulate TWS variability and seasonal cycle well.

<u># Figure 4 to be inserted here</u>

To provide further insights, we present a decomposition of the TWS signal into the general trend and seasonality for two selected river basins, namely the Indus (managed) and the Lena (snow-dominated). As shown in Figure 5, for the Indus while the PCR-GLOBWB simulates both the trend and seasonality in line with GRACE, HiGW-MAT doesn't capture the long-term trend despite simulating the seasonality relatively well. This further confirms that the issue in HiGW-MAT could be the overestimation of GW pumping that results in a larger depletion rate even though the model simulates the seasonal dynamics of the various land surface hydrologic processes as well as water table dynamics. The results for the Lena are contrasting. Here, both models capture the general trend rather accurately but the PCR-GLOBWB fails to simulate the seasonality and timing of TWS anomaly. Analysis of the results for other basins such as the

Amudarya, Colorado, and Euphrates (not shown) suggests that the performance of HiGW-MAT in these basins is similar to that in the Indus but it performs relatively well in the Brahmaputra, Ganges, and Volga basins. The performance of PCR-GLOBWB in most of the other snow-dominated basins is similar to that in the Lena.

# Figure 5 to be inserted here

### 4.3. The Uncertainty Arising from the Climate Forcing Data

The standard deviation of 2002-2008 trend map from three climate forcing datasets illustrates high uncertainty in the order of 10 $cm/yr$ (Figure 6a), highlighting the significant impact of forcing data selection in model results. The standard deviation map of TWS trend drawn from the filtered simulations needs the spatial distribution of scaling factors (Figure 6b) to provide more realistic assessment of existent uncertainties originate from the forcing data. Considering the scaling factors, the restored TWS trend compared to filtered one can be of the order of 2-3 times larger in some grid cells (e.g., northwestern India). The spatial pattern of standard deviation in TWS trend using three different forcing datasets (Figure 6) in comparison with the discrepancies between the spatial pattern of TWS trend from GRACE and HiGW-MAT (Figure 1a vs 1b) notes that the discrepancies between model results and GRACE could partly be contributed by high uncertainties arising from forcing datasets. Furthermore, high standard deviation is particularly obvious over the human affected areas comprising northwest of India, northeastern China, southern Australia, Argentina, central US, and west regions of the Caspian Sea. This is reasonable because the forcing datasets are based on reanalysis (e.g., Onogi et al., 2007), which are produced by assimilating the available observations with the results from atmospheric models that typically do not account for human activities. That is, the forcing datasets, particularly precipitation, may have relatively larger biases in the highly-managed regions.

# Figure 6 to be inserted here

## 5. Discussion

### 5.1. Spatial Patterns in TWS Variations and Contribution of Different Components

The spatial patterns of the long-term trend in total TWS from models show a generally good agreement with GRACE in capturing the direction of change; however, significant differences are found in the magnitude of TWS signal between the two models and GRACE as well as

**Deleted:** Comparing t

**Deleted:** to

between the two models. These differences are highly pronounced especially in the global hotspots of GW overexploitation identified by various previous studies. This is found to be caused partly by the overestimation of groundwater abstraction and the associated change in subsurface storage in the HiGW-MAT model. In other regions, such as the northern high latitudes where the TWS variations are largely modulated by snow water storage, the HiGW-MAT model generally captures the GRACE-based TWS trend but the PCR-GLOBWB model shows a larger deviation from the GRACE trend. The differences between GRACE and models in the high latitudes is likely due to glacier melts, permafrost thaw, and snow cover shrinkage processes that are not explicitly represented in the models as in any other current-generation LSMs and GHMs (Chen et al., 2017; Long et al., 2017). In most of the regions with relatively less human influence and snow contribution (e.g., parts of Europe, western Australia, central Asia and northern Africa) both models perform relatively well, suggesting higher reliability of model results in these areas.

These analyses contribute to the discussion on how the two models that include HI representations regenerate the spatial patterns of the long-term trend in TWS observed by GRACE. Our results corroborate the findings of previous studies that have reported certain discrepancies between GRACE and models in some of the river basins studied here by using other GHMs and LSMs such as the CLM (Swenson and Lawrence, 2015), WaterGAP model (Döll et al., 2014), and GLDAS (Jin and Feng, 2013) models. Together, these findings suggest that a single model cannot be identified as the best model over all global regions, implying that an ensemble model mean could provide a better estimate of TWS variations.

### 5.2. Temporal Variability of TWS in Global Basins: Human-induced TWS Change

An in-depth analysis of the seasonal cycle of TWS variations further suggests that the PCR-GLOBWB tends to perform better in some of the managed basins (e.g., the Indus), in line with studies such as Wada et al. (2014). However, it is found that both models do not accurately capture the seasonal dynamics of TWS in some of these managed basins such as the Huang-He and Murray-Darling. It is also evident from the results that while one model captures the amplitude of the positive seasonal anomaly accurately, it fails to reproduce the negative seasonal anomaly with similar accuracy, and this applies to both models (see Huang-He, Indus, Murray-Darling basins). This implies that while certain human water management practices such as reservoir operation may have been well simulated, the model may have failed to accurately simulate other processes such as GW dynamics that can act as a buffer during high and low flow seasons. It is also important to note that there are differences among the GRACE products in some of these basins making it difficult to evaluate the model performance with high confidence. In the snow-dominated basins (e.g., the Lena, Amur, Mackenzie, and Yukon), the performance of HiGW-MAT is relatively good likely due to its relatively robust and physically-based snow melt scheme which is based on multi-layer snow energy balance (Takata et al., 2003).

The partitioning of inter-annual TWS changes into natural and human components in the highly-managed basins such as the Indus, Amudarya, Ganges, Brahmaputra, Euphrates, and Volga suggests a large deviation in the natural trend from the trend in GRACE data, indicating an expansion of human influence in these basins during 2002-2010. It is worth noting that the rates of TWS change from HI simulations are remarkably different from GRACE observations in many basins, which highlights the uncertainties in simulated trends. The GW extraction scheme in HiGW-MAT tends to consistently overestimate GW withdrawals in some of the human affected basins such as Amudarya, Colorado, Euphrates, Huang-He, and Indus, causing larger TWS decline compared with both GRACE and the PCR-GLOBWB model. However, in other basins such as the Brahmaputra, Ganges, Mekong, and Volga, which also include some managed agricultural regions, no such overestimation of GW depletion is found. The varying performance of HiGW-MAT in the managed basins is likely owing to the use of inaccurate parameters such as the specific yield or overestimation of agricultural demands caused by overestimated irrigated areas (Giordano, 2009; Pokhrel et al., 2015). Similar to the results for the spatial variability, the PCR-GLOBWB performs relatively better in the managed basins but simulates large deviations from both GRACE and HiGW-MAT in the snow-dominated basins such as the Amur, Lena, and Yukon.

Further, the analysis of the general trend and seasonal variability in the Indus and Lena river basins shows that while one model captures the general trend in one basin the other model performs better in capturing the seasonal variability. These large differences in capturing different aspects of the TWS variations in river basins located in different regions again suggest that a single model cannot be used with high reliability in all global regions or to simulate all aspects of TWS variations.

### 5.3. The Uncertainty Arising from the Climate Forcing Data

Results from the HiGW-MAT TWS simulations with three different meteorological forcing datasets reveal that, in some regions, the uncertainties in TWS trends due to the uncertainty in forcing datasets are as high as the differences among different models, or among different models and GRACE data. The forcing uncertainties are particularly pronounced in the highly-managed regions, possibly due to the large uncertainties in the reanalysis products in which results from models without HI are assimilated. The spatial distribution of gain factors derived from the HiGW-MAT model is comparable with gridded scaling factors obtained from other LSMs (Landerer and Swenson, 2012; Long et al., 2015a) and suggesting even larger uncertainties over some grid cells. Such large uncertainties arising from forcing datasets suggest that the model results of TWS based on one particular forcing data need to be interpreted with enough caution, which is especially important when using the model results to evaluate the disagreements among different GRACE solutions and the performance of various filtering and other post-processing techniques applied to GRACE solutions.

## 6. Conclusions

This study quantifies the impacts of human activities (e.g., irrigation, reservoir operation, and GW extraction) on TWS variations over global regions by using multiple GRACE SH products and results from two different hydrological models. Two state-of-the-art models are used, namely the HiGW-MAT LSM and PCR-GLOBWB GHM, both simulate the natural as well as anthropogenic flow of water, also taking into account groundwater abstractions and associated changes in subsurface water storage. We find that despite noteworthy progress that has been made in incorporating human factors in global-scale LSMs and GHMs, significant limitations still remain in accurately simulating the spatial patters and temporal variations in TWS over all global regions. In particular, results indicate that while one model performs better in the highly-managed river basins, it fails to reproduce the GRACE-observed signal in snow-dominated regions, and vice versa. Further, in some regions the uncertainties in TWS trends due to the uncertainties in forcing datasets underscore the need to consider forcing data uncertainties when evaluating the disagreements among different model results and GRACE. Our results from the partitioning of total TWS into natural and human-induced components suggest a continuing decline in TWS through 2002-2010 in the Euphrates, Ganges, Brahmaputra, Volga, and Indus river basins, which is largely human-induced. Overall, our results highlight the need to improve model parameterizations for the simulation of human water management and snow physics (e.g., glacier melts, permafrost thaw, and snow cover shrinkage) to reliably simulate the spatial and temporal variability in TWS over all global regions.

## 7. Acknowledgment

We would like to thank the Editor, Associate Editor, and two anonymous reviewers for providing highly constructive comments that helped tremendously in improving the quality of the paper.

## References


Alcamo, J., Döll, P., Henrichs, T., Kaspar, F., Lehner, B., Rösch, T., Siebert, S., 2003. Development and testing of the WaterGAP 2 global model of water use and availability. Hydrol. Sci. J. 48, 317–337. doi:10.1623/hysj.48.3.317.45290

Alkama, R., Decharme, B., Douville, H., Becker, M., Cazenave, A., Sheffield, J., Voldoire, A., Tyteca, S., Le Moigne, P., 2010. Global Evaluation of the ISBA-TRIP Continental Hydrological System. Part I: Comparison to GRACE Terrestrial Water Storage Estimates and In Situ River Discharges. J. Hydrometeorol. 11, 583–600. doi:10.1175/2010JHM1211.1

Alley, W.M., Healy, R.W., LaBaugh, J.W., Reilly, T.E., 2002. Flow and Storage in Groundwater Systems. Science 296, 1985–1990. doi:10.1126/science.1067123

Batjes, N.H., 2005. ISRIC-WISE global data set of derived soil properties on a 0.5 by 0.5 degree grid (version 3.0, with data set). ISRIC, World Soil Inf. Ctr., Wageningen, the Netherlands.

Beven, K.J., Kirkby, M.J., 1979. A physically based, variable contributing area model of basin hydrology / Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. Hydrol. Sci. Bull. 24, 43–69. doi:10.1080/02626667909491834

Castellazzi, P., Martel, R., Rivera, A., Huang, J., Pavlic, G., Calderhead, A.I., Chaussard, E., Garfias, J., Salas, J., 2016. Groundwater depletion in Central Mexico: Use of GRACE and InSAR to support water resources management. Water Resour. Res. 52, 5985–6003. doi:10.1002/2015WR018211

Chen, X., Long, D., Hong, Y., Zeng, C., Yan, D., 2017. Improved modeling of snow and glacier melting by a progressive two-stage calibration strategy with GRACE and multisource data: How snow and glacier meltwater contributes to the runoff of the Upper Brahmaputra River basin? Water Resour. Res. 53, 2431–2466. doi:10.1002/2016WR019656

Collatz, G.J., Ball, J.T., Grivet, C., Berry, J.A., 1991. Physiological and environmental regulation of stomatal conductance, photosynthesis and transpiration: a model that includes a laminar boundary layer. Agric. For. Meteorol. 54, 107–136. doi:10.1016/0168-1923(91)90002-8

Decharme, B., Alkama, R., Douville, H., Becker, M., Cazenave, A., 2010. Global Evaluation of the ISBA-TRIP Continental Hydrological System. Part II: Uncertainties in River Routing Simulation Related to Flow Velocity and Groundwater Storage. J. Hydrometeorol. 11, 601–617. doi:10.1175/2010JHM1212.1

Decharme, B., Douville, H., 2006. Uncertainties in the GSWP-2 precipitation forcing and their impacts on regional and global hydrological simulations. Clim. Dyn. 27, 695–713. doi:10.1007/s00382-006-0160-6

Dijk, A.I.J.M. van, Renzullo, L.J., 2011. Water resource monitoring systems and the role of satellite observations. Hydrol. Earth Syst. Sci. 15, 39–55. doi:10.5194/hess-15-39-2011

Dirmeyer, P.A., Gao, X., Zhao, M., Guo, Z., Oki, T., Hanasaki, N., 2006. GSWP-2: Multimodel Analysis and Implications for Our Perception of the Land Surface. Bull. Am. Meteorol. Soc. 87, 1381–1397. doi:10.1175/BAMS-87-10-1381

Döll, P., Douville, H., Güntner, A., Schmied, H.M., Wada, Y., 2016. Modelling Freshwater Resources at the Global Scale: Challenges and Prospects, in: Cazenave, A., Champollion, N., Benveniste, J., Chen, J. (Eds.), Remote Sensing and Water Resources, Space Sciences Series of ISSI. Springer International Publishing, pp. 5–31. doi:10.1007/978-3-319-32449-4_2

Döll, P., Kaspar, F., Lehner, B., 2003. A global hydrological model for deriving water availability indicators: model tuning and validation. J. Hydrol. 270, 105–134. doi:10.1016/S0022-1694(02)00283-4

Döll, P., Müller Schmied, H., Schuh, C., Portmann, F.T., Eicker, A., 2014. Global-scale assessment of groundwater depletion and related groundwater abstractions: Combining hydrological modeling with information from well observations and GRACE satellites. Water Resour. Res. 50, 5698–5720. doi:10.1002/2014WR015595

Duffy, C.J., 1996. A Two-State Integral-Balance Model for Soil Moisture and Groundwater Dynamics in Complex Terrain. Water Resour. Res. 32, 2421–2434. doi:10.1029/96WR01049

Eicker, A., Forootan, E., Springer, A., Longuevergne, L., Kusche, J., 2016. Does GRACE see the terrestrial water cycle “intensifying”? J. Geophys. Res. Atmospheres 121, 2015JD023808. doi:10.1002/2015JD023808

Eicker, A., Schumacher, M., Kusche, J., Döll, P., Schmied, H.M., 2014. Calibration/Data Assimilation Approach for Integrating GRACE Data into the WaterGAP Global Hydrology Model (WGHM) Using an Ensemble Kalman Filter: First Results. Surv. Geophys. 35, 1285–1309. doi:10.1007/s10712-014-9309-8

Famiglietti, J.S., 2014. The global groundwater crisis. Nat. Clim. Change 4, 945–948. doi:10.1038/nclimate2425

Famiglietti, J.S., Cazenave, A., Eicker, A., Reager, J.T., Rodell, M., Velicogna, I., 2015. Satellites provide the big picture. Science 349, 684–685. doi:10.1126/science.aac9238

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

Fan, Y., 2015. Groundwater in the Earth’s critical zone: Relevance to large-scale patterns and processes. Water Resour. Res. n/a-n/a. doi:10.1002/2015WR017037

FAO, 2003. Food and Agriculture Organization of the United Nations (FAO) (2003), Digital Soil Map of the World, Version 3.6. Rome, Italy.

Feng, W., Zhong, M., Lemoine, J.-M., Biancale, R., Hsu, H.-T., Xia, J., 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

Freedman, F.R., Pitts, K.L., Bridger, A.F.C., 2014. Evaluation of CMIP climate model hydrological output for the Mississippi River Basin using GRACE satellite observations. J. Hydrol. 519, Part D, 3566–3577. doi:10.1016/j.jhydrol.2014.10.036

Ge, S., Yang, D., Kane, D.L., 2013. Yukon River Basin long-term (1977–2006) hydrologic and climatic analysis. Hydrol. Process. 27, 2475–2484. doi:10.1002/hyp.9282

Giordano, M., 2009. Global Groundwater? Issues and Solutions. Annu. Rev. Environ. Resour. 34, 153–178. doi:10.1146/annurev.environ.030308.100251

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

Gleeson, T., Wada, Y., Bierkens, M.F.P., van Beek, L.P.H., 2012. Water balance of global aquifers revealed by groundwater footprint. Nature 488, 197–200. doi:10.1038/nature11295

Grippa, M., Kergoat, L., Frappart, F., Araud, Q., Boone, A., de Rosnay, P., Lemoine, J.-M., Gascoin, S., Balsamo, G., Ottlé, C., Decharme, B., Saux-Picart, S., Ramillien, G., 2011. Land water storage variability over West Africa estimated by Gravity Recovery and Climate Experiment (GRACE) and land surface models. Water Resour. Res. 47, W05549. doi:10.1029/2009WR008856

Güntner, A., 2008. Improvement of Global Hydrological Models Using GRACE Data. Surv. Geophys. 29, 375–397. doi:10.1007/s10712-008-9038-y

Haddeland, I., Clark, D.B., Franssen, W., Ludwig, F., Voß, F., Arnell, N.W., Bertrand, N., Best, M., Folwell, S., Gerten, D., Gomes, S., Gosling, S.N., Hagemann, S., Hanasaki, N., Harding, R., Heinke, J., Kabat, P., Koirala, S., Oki, T., Polcher, J., Stacke, T., Viterbo, P., Weedon, G.P., Yeh, P., 2011. Multimodel Estimate of the Global Terrestrial Water Balance: Setup and First Results. J. Hydrometeorol. 12, 869–884. doi:10.1175/2011JHM1324.1

Hagemann, S., 2002. An improved land surface parameter dataset for global and regional climate models, Report / Max-Planck-Institut für Meteorologie. Max-Planck-Institut für Meteorologie, Hamburg.

Hanasaki, N., Kanae, S., Oki, T., 2006. A reservoir operation scheme for global river routing models. J. Hydrol. 327, 22–41. doi:10.1016/j.jhydrol.2005.11.011

Hirabayashi, Y., Kanae, S., Struthers, I., Oki, T., 2005. A 100-year (1901–2000) global retrospective estimation of the terrestrial water cycle. J. Geophys. Res. Atmospheres 110, D19101. doi:10.1029/2004JD005492

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

Huang, Y., Salama, M.S., Krol, M.S., Su, Z., Hoekstra, A.Y., Zeng, Y., Zhou, Y., 2015. Estimation of human-induced changes in terrestrial water storage through integration of GRACE satellite detection and hydrological modeling: A case study of the Yangtze River basin. Water Resour. Res. 51, 8494–8516. doi:10.1002/2015WR016923

Jin, S., Feng, G., 2013. Large-scale variations of global groundwater from satellite gravimetry and hydrological models, 2002–2012. Glob. Planet. Change 106, 20–30. doi:10.1016/j.gloplacha.2013.02.008

Kim, H., Yeh, P.J.-F., Oki, T., Kanae, S., 2009. Role of rivers in the seasonal variations of terrestrial water storage over global basins. Geophys. Res. Lett. 36, L17402. doi:10.1029/2009GL039006

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

Koirala, S., Yeh, P.J.-F., Hirabayashi, Y., Kanae, S., Oki, T., 2014. Global-scale land surface hydrologic modeling with the representation of water table dynamics. J. Geophys. Res. Atmospheres 119, 2013JD020398. doi:10.1002/2013JD020398

Kraijenhoff Van De Leur, D.A., 1958. A study of non-steady groundwater flow with special reference to a reservoir-coefficient. Ing. 70, 87–94.

Krysanova, V., Müller-Wohlfeil, D.-I., Becker, A., 1998. Development and test of a spatially distributed hydrological/water quality model for mesoscale watersheds. Ecol. Model. 106, 261–289. doi:10.1016/S0304-3800(97)00204-4

Landerer, F.W., Dickey, J.O., Güntner, A., 2010. Terrestrial water budget of the Eurasian pan-Arctic from GRACE satellite measurements during 2003–2009. J. Geophys. Res. Atmospheres 115, D23115. doi:10.1029/2010JD014584

Landerer, F.W., Gleckler, P.J., Lee, T., 2013. Evaluation of CMIP5 dynamic sea surface height multi-model simulations against satellite observations. Clim. Dyn. 43, 1271–1283. doi:10.1007/s00382-013-1939-x

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

Lawrence, D.M., Oleson, K.W., Flanner, M.G., Thornton, P.E., Swenson, S.C., Lawrence, P.J., Zeng, X., Yang, Z.-L., Levis, S., Sakaguchi, K., Bonan, G.B., Slater, A.G., 2011. Parameterization improvements and functional and structural advances in Version 4 of the Community Land Model. J. Adv. Model. Earth Syst. 3, M03001. doi:10.1029/2011MS00045

Lehner, B., Liermann, C.R., Revenga, C., Vörösmarty, C., Fekete, B., Crouzet, P., Döll, P., Endejan, M., Frenken, K., Magome, J., Nilsson, C., Robertson, J.C., Rödel, R., Sindorf, N., Wisser, D., 2011. High-resolution mapping of the world’s reservoirs and dams for sustainable river-flow management. Front. Ecol. Environ. 9, 494–502. doi:10.1890/100125

Li, B., Rodell, M., 2015. Evaluation of a model-based groundwater drought indicator in the conterminous U.S. J. Hydrol., Drought processes, modeling, and mitigation 526, 78–88. doi:10.1016/j.jhydrol.2014.09.027

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

Long, D., Chen, X., Scanlon, B.R., Wada, Y., Hong, Y., Singh, V.P., Chen, Y., Wang, C., Han, Z., Yang, W., 2016. Have GRACE satellites overestimated groundwater depletion in the Northwest India Aquifer? Sci. Rep. 6, 24398. doi:10.1038/srep24398

Long, D., Longuevergne, L., Scanlon, B.R., 2015a. 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., Longuevergne, L., Scanlon, B.R., 2014. Uncertainty in evapotranspiration from land surface modeling, remote sensing, and GRACE satellites. Water Resour. Res. 50, 1131–1151. doi:10.1002/2013WR014581

Long, D., Pan, Y., Zhou, J., Chen, Y., Hou, X., Hong, Y., Scanlon, B.R., Longuevergne, L., 2017. Global analysis of spatiotemporal variability in merged total water storage changes using multiple GRACE products and global hydrological models. Remote Sens. Environ. 192, 198–216. doi:10.1016/j.rse.2017.02.011

Long, D., Yang, Y., Wada, Y., Hong, Y., Liang, W., Chen, Y., Yong, B., Hou, A., Wei, J., Chen, L., 2015b. Deriving scaling factors using a global hydrological model to restore GRACE total water storage changes for China’s Yangtze River Basin. Remote Sens. Environ. 168, 177–193. doi:10.1016/j.rse.2015.07.003

Longuevergne, L., Scanlon, B.R., Wilson, C.R., 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

Nanteza, J., de Linage, C.R., Thomas, B.F., Famiglietti, J.S., 2016. Monitoring groundwater storage changes in complex basement aquifers: An evaluation of the GRACE satellites over East Africa. Water Resour. Res. 52, 9542–9564. doi:10.1002/2016WR018846

Nazemi, A., Wheater, H.S., 2015. On inclusion of water resource management in Earth system models – Part 1: Problem definition and representation of water demand. Hydrol Earth Syst Sci 19, 33–61. doi:10.5194/hess-19-33-2015

Oki, T., Sud, Y.C., 1998. Design of Total Runoff Integrating Pathways (TRIP)—A Global River Channel Network. Earth Interact. 2, 1–37. doi:10.1175/1087-3562(1998)002<0001:DOTRIP>2.3.CO;2

Onogi, K., Tsutsui, J., Koide, H., Sakamoto, M., Kobayashi, S., Hatsushika, H., Matsumoto, T., Yamazaki, N., Kamahori, H., Takahashi, K., Kadokura, S., Wada, K., Kato, K., Oyama, R., Ose, T., Mannoji, N., Taira, R., 2007. The JRA-25 Reanalysis. J. Meteorol. Soc. Jpn. Ser II 85, 369–432. doi:10.2151/jmsj.85.369

Overgaard, J., Rosbjerg, D., Butts, M.B., 2006. Land-surface modelling in hydrological perspective – a review. Biogeosciences 3, 229–241. doi:10.5194/bg-3-229-2006

Pail, R., Bingham, R., Braitenberg, C., Dobslaw, H., Eicker, A., Güntner, A., Horwath, M., Ivins, E., Longuevergne, L., Panet, I., Wouters, B., Panel, I.E., 2015. Science and User Needs for Observing Global Mass Transport to Understand Global Change and to Benefit Society. Surv. Geophys. 36, 743–772. doi:10.1007/s10712-015-9348-9

Pan, Y., Zhang, C., Gong, H., Yeh, P.J.-F., Shen, Y., Guo, Y., Huang, Z., Li, X., 2016. Detection of human-induced evapotranspiration using GRACE satellite observations in the Haihe River basin of China. Geophys. Res. Lett. 2016GL071287. doi:10.1002/2016GL071287

Pitman, A.J., 2003. The evolution of, and revolution in, land surface schemes designed for climate models. Int. J. Climatol. 23, 479–510. doi:10.1002/joc.893

Pokhrel, Y.N., Fan, Y., Miguez-Macho, G., Yeh, P.J.-F., Han, S.-C., 2013. The role of groundwater in the Amazon water cycle: 3. Influence on terrestrial water storage computations and comparison with GRACE. J. Geophys. Res. Atmospheres 118, 3233–3244. doi:10.1002/jgrd.50335

Pokhrel, Y.N., Felfelani, F., Shin, S., Yamada, T.J., Satoh, Y., 2017. Modeling large-scale human alteration of land surface hydrology and climate. Geosci. Lett. 4, 10. doi:10.1186/s40562-017-0076-5

Pokhrel, Y.N., Hanasaki, N., Koirala, S., Cho, J., Yeh, P.J.-F., Kim, H., Kanae, S., Oki, T., 2012a. Incorporating Anthropogenic Water Regulation Modules into a Land Surface Model. J. Hydrometeorol. 13, 255–269. doi:10.1175/JHM-D-11-013.1

Pokhrel, Y.N., Hanasaki, N., Wada, Y., Kim, H., 2016. Recent progresses in incorporating human land–water management into global land surface models toward their integration into Earth system models. Wiley Interdiscip. Rev. Water 3, 548–574. doi:10.1002/wat2.1150

Pokhrel, Y.N., Hanasaki, N., Yeh, P.J.-F., Yamada, T.J., Kanae, S., Oki, T., 2012b. Model estimates of sea-level change due to anthropogenic impacts on terrestrial water storage. Nat. Geosci. 5, 389–392. doi:10.1038/ngeo1476

Pokhrel, Y.N., Koirala, S., Yeh, P.J.-F., Hanasaki, N., Longuevergne, L., Kanae, S., Oki, T., 2015. Incorporation of groundwater pumping in a global Land Surface Model with the representation of human impacts. Water Resour. Res. 51, 78–96. doi:10.1002/2014WR015602

Portmann, F.T., Siebert, S., Döll, P., 2010. MIRCA2000—Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling. Glob. Biogeochem. Cycles 24, GB1011. doi:10.1029/2008GB003435

Richards, L.A., 1931. CAPILLARY CONDUCTION OF LIQUIDS THROUGH POROUS MEDIUMS. J. Appl. Phys. 1, 318–333. doi:10.1063/1.1745010

Rodell, M., Famiglietti, J.S., Chen, J., Seneviratne, S.I., Viterbo, P., Holl, S., Wilson, C.R., 2004. Basin scale estimates of evapotranspiration using GRACE and other observations. Geophys. Res. Lett. 31, L20504. doi:10.1029/2004GL020873

Rodell, M., Velicogna, I., Famiglietti, J.S., 2009. Satellite-based estimates of groundwater depletion in India. Nature 460, 999–1002. doi:10.1038/nature08238

Rosenberg, E.A., Clark, E.A., Steinemann, A.C., Lettenmaier, D.P., 2013. On the contribution of groundwater storage to interannual streamflow anomalies in the Colorado River basin. Hydrol Earth Syst Sci 17, 1475–1491. doi:10.5194/hess-17-1475-2013

Russo, T.A., Lall, U., 2017. Depletion and response of deep groundwater to climate-induced pumping variability. Nat. Geosci. 10, 105–108. doi:10.1038/ngeo2883

Sakumura, C., Bettadpur, S., Bruinsma, S., 2014. Ensemble prediction and intercomparison analysis of GRACE time-variable gravity field models. Geophys. Res. Lett. 41, 1389–1397. doi:10.1002/2013GL058632

Scanlon, B.R., Faunt, C.C., Longuevergne, L., Reedy, R.C., Alley, W.M., McGuire, V.L., McMahon, P.B., 2012a. Groundwater depletion and sustainability of irrigation in the US High Plains and Central Valley. Proc. Natl. Acad. Sci. 109, 9320–9325. doi:10.1073/pnas.1200311109

Scanlon, B.R., Longuevergne, L., Long, D., 2012b. Ground referencing GRACE satellite estimates of groundwater storage changes in the California Central Valley, USA. Water Resour. Res. 48, W04520. doi:10.1029/2011WR011312

Scanlon, B.R., Zhang, Z., Reedy, R.C., Pool, D.R., Save, H., Long, D., Chen, J., Wolock, D.M., Conway, B.D., Winester, D., 2015. Hydrologic implications of GRACE satellite data in the Colorado River Basin. Water Resour. Res. 51, 9891–9903. doi:10.1002/2015WR018090

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

Sellers, P.J., Dickinson, R.E., Randall, D.A., Betts, A.K., Hall, F.G., Berry, J.A., Collatz, G.J., Denning, A.S., Mooney, H.A., Nobre, C.A., Sato, N., Field, C.B., Henderson-Sellers, A., 1997. Modeling the Exchanges of Energy, Water, and Carbon Between Continents and the Atmosphere. Science 275, 502–509. doi:10.1126/science.275.5299.502

Sheffield, J., Goteti, G., Wood, E.F., 2006. Development of a 50-Year High-Resolution Global Dataset of Meteorological Forcings for Land Surface Modeling. J. Clim. 19, 3088–3111. doi:10.1175/JCLI3790.1

Spence, C., 2002. Streamflow Variability (1965 to 1998) in Five Northwest Territories and Nunavut Rivers. Can. Water Resour. J. Rev. Can. Ressour. Hydr. 27, 135–154. doi:10.4296/cwrj2702135

St. Jacques, J.-M., Sauchyn, D.J., 2009. Increasing winter baseflow and mean annual streamflow from possible permafrost thawing in the Northwest Territories, Canada. Geophys. Res. Lett. 36, L01401. doi:10.1029/2008GL035822

Stieglitz, M., Rind, D., Famiglietti, J., Rosenzweig, C., 1997. An Efficient Approach to Modeling the Topographic Control of Surface Hydrology for Regional and Global Climate Modeling. J. Clim. 10, 118–137. doi:10.1175/1520-0442(1997)010<0118:AEATMT>2.0.CO;2

Swenson, S.C., Lawrence, D.M., 2015. A GRACE-based assessment of interannual groundwater dynamics in the Community Land Model. Water Resour. Res. 51, 8817–8833. doi:10.1002/2015WR017582

Syed, T.H., Famiglietti, J.S., Rodell, M., Chen, J., Wilson, C.R., 2008. Analysis of terrestrial water storage changes from GRACE and GLDAS. Water Resour. Res. 44, W02433. doi:10.1029/2006WR005779

Takata, K., Emori, S., Watanabe, T., 2003. Development of the minimal advanced treatments of surface interaction and runoff. Glob. Planet. Change 38, 209–222. doi:10.1016/S0921-8181(03)00030-4

Taylor, R.G., Scanlon, B., Döll, P., Rodell, M., Beek, R. van, Wada, Y., Longuevergne, L., Leblanc, M., Famiglietti, J.S., Edmunds, M., Konikow, L., Green, T.R., Chen, J., Taniguchi, M., Bierkens, M.F.P., MacDonald, A., Fan, Y., Maxwell, R.M., Yechieli, Y., Gurdak, J.J., Allen, D.M., Shamsudduha, M., Hiscock, K., Yeh, P.J.-F., Holman, I., Treidel, H., 2013. Ground water and climate change. Nat. Clim. Change 3, 322–329. doi:10.1038/nclimate1744

Tiwari, V.M., Wahr, J., Swenson, S., 2009. Dwindling groundwater resources in northern India, from satellite gravity observations. Geophys. Res. Lett. 36, L18401. doi:10.1029/2009GL039401

Trancoso, R., Larsen, J.R., McVicar, T.R., Phinn, S.R., McAlpine, C.A., 2017. CO2-vegetation feedbacks and other climate changes implicated in reducing base flow. Geophys. Res. Lett. 44, 2017GL072759. doi:10.1002/2017GL072759

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

Vörösmarty, C.J., Fekete, B.M., Meybeck, M., Lammers, R., 2000. A simulated topological network representing the global system of rivers at 30-minute spatial resolution (STN-30). Global Biogeo. Cy. 14, 599–621.

Wada, Y., van Beek, L.P.H., van Kempen, C.M., Reckman, J.W.T.M., Vasak, S., Bierkens, M.F.P., 2010. Global depletion of groundwater resources. Geophys. Res. Lett. 37, L20402. doi:10.1029/2010GL044571

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

Wahr, J., Molenaar, M., Bryan, F., 1998. Time variability of the Earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE. J. Geophys. Res. Solid Earth 103, 30205–30229. doi:10.1029/98JB02844

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

Wang, S., Huang, J., Yang, D., Pavlic, G., Li, J., 2015a. Long-term water budget imbalances and error sources for cold region drainage basins. Hydrol. Process. 29, 2125–2136. doi:10.1002/hyp.10343

Wang, S., Pan, M., Mu, Q., Shi, X., Mao, J., Brümmer, C., Jassal, R.S., Krishnan, P., Li, J., Black, T.A., 2015b. Comparing Evapotranspiration from Eddy Covariance Measurements, Water Budgets, Remote Sensing, and Land Surface Models over Canada. J. Hydrometeorol. 16, 1540–1560. doi:10.1175/JHM-D-14-0189.1

Watanabe, T., 1994. Bulk parameterization for a vegetated surface and its application to a simulation of nocturnal drainage flow. Bound.-Layer Meteorol. 70, 13–35. doi:10.1007/BF00712521

Watkins, M.M., Wiese, D.N., Yuan, D.-N., Boening, C., Landerer, F.W., 2015. Improved methods for observing Earth's time variable mass distribution with GRACE using spherical cap mascons. J. Geophys. Res. Solid Earth 120, 2014JB011547. doi:10.1002/2014JB011547

Weedon, G.P., Balsamo, G., Bellouin, N., Gomes, S., Best, M.J., Viterbo, P., 2014. The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data. Water Resour. Res. 50, 7505–7514. doi:10.1002/2014WR015638

Werth, S., Güntner, A., Schmidt, R., Kusche, J., 2009. Evaluation of GRACE filter tools from a hydrological perspective. Geophys. J. Int. 179, 1499–1515. doi:10.1111/j.1365-246X.2009.04355.x

Xie, H., Longuevergne, L., Ringler, C., Scanlon, B.R., 2012. Calibration and evaluation of a semi-distributed watershed model of Sub-Saharan Africa using GRACE data. Hydrol. Earth Syst. Sci. 16, 3083–3099. doi:10.5194/hess-16-3083-2012

Yang, Z.-L., Niu, G.-Y., Mitchell, K.E., Chen, F., Ek, M.B., Barlage, M., Longuevergne, L., Manning, K., Niyogi, D., Tewari, M., Xia, Y., 2011. The community Noah land surface model with multiparameterization options (Noah-MP): 2. Evaluation over global river basins. J. Geophys. Res. Atmospheres 116, D12110. doi:10.1029/2010JD015140

Zaitchik, B.F., Rodell, M., Reichle, R.H., 2008. Assimilation of GRACE Terrestrial Water Storage Data into a Land Surface Model: Results for the Mississippi River Basin. J. Hydrometeorol. 9, 535–548. doi:10.1175/2007JHM951.1

Zeng, N., Yoon, J.-H., Mariotti, A., Swenson, S., 2008. Variability of Basin-Scale Terrestrial Water Storage from a PER Water Budget Method: The Amazon and the Mississippi. J. Clim. 21, 248–265. doi:10.1175/2007JCLI1639.1

Zhao, F., Veldkamp, T., Frieler, K., Schewe, J., Ostberg, S., Willner, S., Schauberger, B., Gosling, S., Müller Schmied, H., Portmann, F.T., Leng, G., Huang, M., Liu, X., Tang, Q., Hanasaki, N., Biemans, H., Gerten, D., Satoh, Y., Pokhrel, Y., Stacke, T., Ciais, P., Chang, J., Ducharne, A., Guimberteau, M., Wada, Y., Kim, H., Yamazaki, D., 2017. Choice of routing scheme considerably influences peak river discharge simulation in global hydrological models. Environ. Res. Lett.

Zhou, L., Tian, Y., Myneni, R.B., Ciais, P., Saatchi, S., Liu, Y.Y., Piao, S., Chen, H., Vermote, E.F., Song, C., Hwang, T., 2014. Widespread decline of Congo rainforest greenness in the past decade. Nature 509, 86–90. doi:10.1038/nature13265

**Deleted:** Alcamo, J., Döll, P., Henrichs, T., Kaspar, F., Lehner, B., Rösch, T., Siebert, S., 2003. Development and testing of the WaterGAP 2 global model of water use and availability. Hydrol. Sci. J. 48, 317–337. doi:10.1623/hysj.48.3.317.45290¶
Alkama, R., Decharme, B., Douville, H., Becker, M., Cazenave, A., Sheffield, J., Voldoire, A., Tyteca, S., Le Moigne, P., 2010. Global Evaluation of the ISBA-TRIP Continental Hydrological System. Part I: Comparison to GRACE Terrestrial Water Storage Estimates and In Situ River Discharges. J. Hydrometeorol. 11, 583–600. doi:10.1175/2010JHM1211.1¶
Alley, W.M., Healy, R.W., LaBaugh, J.W., Reilly, T.E., 2002. Flow and Storage in Groundwater Systems. Science 296, 1985–1990. doi:10.1126/science.1067123¶
Batjes, N.H., 2005. ISRIC-WISE global data set of derived soil properties on a 0.5 by 0.5 degree grid (version 3.0, with data set). ISRIC, World Soil Inf. Ctr., Wageningen, the Netherlands.¶
Beven, K.J., Kirkby, M.J., 1979. A physically based, variable contributing area model of basin hydrology / Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. Hydrol. Sci. Bull. 24, 43–69. doi:10.1080/02626667909491834¶
Castellazzi, P., Martel, R., Rivera, A., Huang, J., Pavlic, G., Calderhead, A.I., Chaussard, E., Garfias, J., Salas, J., 2016. Groundwater depletion in Central Mexico: Use of GRACE and InSAR to support water resources management. Water Resour. Res. 52, 5985–6003. doi:10.1002/2015WR018211¶
Chen, X., Long, D., Hong, Y., Zeng, C., Yan, D., 2017. Improved modeling of snow and glacier melting by a progressive two-stage calibration strategy with GRACE and multisource data: How snow and glacier meltwater contributes to the runoff of the Upper Brahmaputra River basin? Water Resour. Res. 53, 2431–2466. doi:10.1002/2016WR019656¶
Collatz, G.J., Ball, J.T., Grivet, C., Berry, J.A., 1991. Physiological and environmental regulation of stomatal conductance, photosynthesis and transpiration: a model that includes a laminar boundary layer. Agric. For. Meteorol. 54, 107–136. doi:10.1016/0168-1923(91)90002-8¶
Decharme, B., Alkama, R., Douville, H., Becker, M., Cazenave, A., 2010. Global Evaluation of the ISBA-TRIP Continental Hydrological System. Part II: Uncertainties in River Routing Simulation Related to Flow Velocity and Groundwater Storage. J. Hydrometeorol. 11, 601–617. doi:10.1175/2010JHM1212.1¶
Decharme, B., Douville, H., 2006. Uncertainties in the GSWP-2 precipitation forcing and their impacts on regional and global hydrological simulations. Clim. Dyn. 27, 695–713. doi:10.1007/s00382-006-0160-6¶
Dijk, A.I.J.M. van, Renzullo, L.J., 2011. Water resource monitoring systems and the role of satellite observations. Hydrol. Earth Syst. Sci. 15, 39–55. doi:10.5194/hess-15-39-2011¶
Dirmeyer, P.A., Gao, X., Zhao, M., Guo, Z., Oki, T., Hanasaki, N., 2006. GSWP-2: Multimodel Analysis and Implications for Our Perception of the Land Surface. Bull. Am. Meteorol. Soc. 87, 1381–1397. doi:10.1175/BAMS-87-10-1381¶ ...

**Figure Captions:**

**Figure 1.** Spatial pattern of TWS trend from GRACE, and the two models (HiGW-MAT and PCR-GLOBWB) for 2002-2008. GRACE results are shown as the mean of the solutions from three different processing centers (i.e., CSR, JPL, and GFZ).

**Figure 2.** Map showing the selected 30 river basins with the component contribution ratio (CCR) for snow water, surface water (rivers and reservoirs), and subsurface water (soil moisture and groundwater) storages, shown as pie charts for each of the basins. The CCR values are calculated by using HiGW-MAT model results. The size of pie chart is proportional to the seasonal amplitude of TWS variation, with the largest amplitude being 500 mm in the Orinoco river basin.

**Figure 3.** Seasonal cycle of simulated and observed TWS and components for the selected river basins. Yellow background indicates the region with human impacts and light blue background represents snow-dominated basin. Basins with relatively less human influence and contribution from snow are shown with white background. The thick black line represents the mean of three GRACE products from CSR, JPL, and GFZ and the gray-shaded band shows the range of variations among the three GRACE products. While the simulated total TWS from both models are shown, the individual components (i.e., snow, river and reservoir, soil moisture, and groundwater storages) are shown only from the HiGW-MAT model for clarity of view.

**Figure 4.** Inter-annual variability in TWS from GRACE and the two models. Background colors represent the same as in Figure 3. For the managed basins (top five rows with yellow background), the GRACE data and model results are plotted as line diagram on the top and the results from the water balance analysis using the natural simulations (Equation 4) are shown on the bottom as bars. The gray-shaded range represents the range of variations of the GRACE products (CSR, JPL, and GFZ) along with the thick black line that shows the mean. The individual water storage components are shown only from the HiGW-MAT model for clarity of view.

**Figure 5.** Decomposition of TWS time series into the general trend and seasonality for the Lena (snow-dominated) and Indus (managed) river basins.

**Figure 6.** (a) Standard deviation in TWS trend calculated for 2002-2008 based on the results from HiGW-MAT model simulated by using three different forcing datasets, (b) the spatial distribution of scaling factors derived from the HiGW-MAT model.