Long-term Total Water Storage Change from a SAtellite Water Cycle (SAWC) reconstruction over large south Asian basins

Victor Pellet[1,2,5], Filipe Aires[1,2], Fabrice Papa[3], Simon Munier[4], and Bertrand Decharme[4]

[1]Laboratoire d'Études du Rayonnement et de la Matière en Astrophysique et Atmosphères, Observatoire de Paris, Paris, France.
[2]Estellus, Paris, France.
[3]Laboratoire d'Études en Géophysique et Océanographie Spatiales, IRD, Toulouse, France.
[4]CNRM-Université de Toulouse, Météo-France, CNRS, Toulouse, France
[5]Now at Institute of Industrial Science, University of Tokyo, Tokyo, Japan.

Correspondence: Victor Pellet(victor.pellet@obspm.fr)

Abstract. The Total Water Storage Change (TWSC) over land is a major component of the global water cycle, with a large influence on climate variability, sea level budget and water resources availability for human life. Its first estimates at large-scale were made available with GRACE observations for the 2002-2016 period, followed since 2018 by the launch of GRACE-FO mission. In this paper, using an approach based on the water mass conservation rule, we proposed to merge satellite-based observations of precipitation and evapotranspiration along with *in situ* river discharge measurements to estimate TWSC over longer time periods (typically from 1980 to 2016), compatible with climate studies. We performed this task over five major Asian basins, subject to both large climate variability and strong anthropogenic pressure for water resources, and for which long term record of *in situ* discharge measurements are available. Our SAtellite Water Cycle (SAWC) reconstruction provides TWSC estimates very coherent in terms of seasonal and interannual variations with independent sources of information such as (1) TWSC GRACE-derived observations (over the 2002-2015 period), (2) ISBA-CTRIP model simulations (1980-2015), and (3) multi-satellite inundation extent (1993-2007). This analysis shows the advantages of the use of multiple satellite-derived data sets along with *in situ* data to perform hydrologically coherent reconstruction of missing water component estimate. It provides a new critical source of information for long term monitoring of TWSC and to better understand their critical role in the global and terrestrial water cycle.

1 Introduction

Continental freshwater, excluding ice caps, represents only few percents of the total amount of water on Earth. Nevertheless they have a major impact on terrestrial environment and human life and activities and play a very important role in climate variability. Thus, understanding and predicting continental water storage variations is a topic of great importance for climate research, global water cycle studies (IPCC, 2014) and water resource management. In particular, the Total Water Storage Change (TWSC), comprising of all water mass variations from surface waters (wetland, floodplains, lakes, rivers and man-made reservoirs), soil moisture, snow pack, glaciers and groundwater, is of high interest because it represents a good indicator of potential long-term water cycle (WC) modifications related to natural or anthropogenic factors (Rodell et al., 2018).

Therefore, monitoring long-term spatio-temporal changes in continental freshwater storage has become fundamental. This question is particularly important for regions such as South Asia that experienced drastic changes over the last decades. The region includes some of the worlds largest rivers (Fig. 1), originating in the Himalayas and crossing densely inhabited areas of the Indian subcontinent or South-East Asia, where changes in freshwater availability (Babel and Wahid, 2008) might threaten food and security for more than a billion people (Shamsudduha and Panda, 2019; Wijngaard et al., 2018).

Given the limited availability of *in situ* data in the region, satellite observations are unique to monitor the dynamic of terrestrial waters (Tiwari et al., 2009; Papa et al., 2015; Salameh et al., 2017) and analyzed their recent large-scale changes (Rodell et al., 2009; Asoka et al., 2017; Khaki et al., 2018). In particular, since 2002, the GRACE mission (Tapley et al., 2004) monitors the mass gravity field variation and provides an estimation of the TWSC at the monthly scale (for the period 2002-2016), followed since 2018 by the Gravity Recovery and Climate Experiment Follow-On (GRACE-FO). However, GRACE data time span is still too limited to study the long-term behavior of the WC related to climate changes or to human practices.

If classical approaches to retrieve TWS rely on Land Surface Model (LSM) (Decharme et al., 2019b; Tootchi et al., 2018), studies have recently attempted to use instead statistical models fed with various climate drivers. For instance, Humphrey et al. (2017) have reconstructed the TWSA using a linear regression from precipitation and temperature; and Chen et al. (2019) have use an artificial neural network to reconstruct TWSA based on precipitation, temperature and surface variables (e.g. soil moisture and NDVI). Yang et al. (2018) reviewed and compared several statistical methods (linear, random forest, artificial neural network and support vector machine) to reconstruct TWSA from soil moisture, canopy water and snow water equivalent. These studies have focused on TWSA, without monitoring the whole WC. In fact, if statistical methods offer the opportunity to estimate TWS anomalies at global scale in a simpler way than LSMs, they do not consider the water balance and the related TWS estimate may not be coherent with the other water components.

Several studies have attempted to monitor the WC and provide independent estimates of TWSC using satellite observations (Lawford et al., 2007; Pan et al., 2012; Rodell et al., 2015; Munier and Aires, 2017; Tang et al., 2017). These analyses potentially allow new opportunities for the WC monitoring over long time-records in regions with limited access to *in situ* measurements. The use of satellite data to study the WC is however not straightforward. (1) Various datasets exist for the same geophysical variable and (2) they all have uncertainties (systematic bias and random errors), which lead to (3) the inconsistency between estimate of the same variable or among the variable estimates of the WC (Pellet et al., 2018). No singular estimate can be considered as perfect and many authors preferred to combine various available products (Sheffield et al., 2009; Sahoo et al., 2011; Azarderakhsh et al., 2011; Lorenz et al., 2014). For that purpose, some have focused on the water conservation equation:

$$\Delta S \quad = \quad P - E - D \tag{1}$$

where ΔS is the TWSC, P the Precipitation, E the Evaporation, and D the Discharge (expressed in mm/month, area-normalised). This closure of the WC budget allows to better constrain the integration of the datasets. **For instance, Pan**

et al. (2012) have used an assimilation approach based on a Kalman filtering in the Variable Infiltration Capicity (VIC) model to derive a coherent analysis of the four terrestrial water variables (P, E, D and ΔS) at bassin scale and Zhang et al. (2017) derive the methodology at the 0.5° LSM pixel. Tang et al. (2017) use implicitly the water conservation through the Budyko model to estimate long term annual TWSC based on P, E and R. This approach is not based on the assimilation of satellite observations but rather on the calibration of model parameters to match observed TWSC. Rodell et al. (2015) used variational 3D-VAR strategy to optimize the water cycle estimates at the global and annual scales.

Other approaches perform this integration independently from any model, which allows the integrated datasets being interesting for the calibration/validation of the model (Aires, 2014; Munier et al., 2014; Pellet et al., 2019). Pan and Wood (2006); Aires (2014) have presented several methodologies to integrate coherently different hydrological datasets based on a budget closure. Munier et al. (2014) applied one of them (Aires, 2014) over the Mississippi basin using remote sensing observation for P, E and ΔS and *in situ* measurment for D. The optimal integration is based on, first, a Simple Weighting (SW) average, then, a closure Post-Filtering (PF). The SW+PF method improved the WC components estimate compared to *in situ* observation. The uncertainties of integrated product are reduced compared to the original datasets, the coherency is improved, and the residuals of the WC budget closure are decreased. Furthermore, the authors have develloped a calibration approach based on the integrated product, able to correct each original estimate in an independent way. This calibration led to a significative reduction of the budget residual (see also Pellet et al. (2019)). It was shown in Munier et al. (2014) that when three out of four WC components in Eq. (1) are available, the reconstruction of the missing one can be attempted. This is possible if the signal-to-noise ratio is sufficient: discharge reconstruction was not possible in Munier et al. (2014) but the TWSC could be obtained in a very simple way, with results quality comparable to a complex assimilation into a hydrologic model.

In this study, we propose to use this methodology to reconstruct the long-term evolution of the TWSC over large South Asian basins, based on satellite and *in situ* measurements and no hydrological model. We denote “SAWC" this SAtellite Water Cycle reconstruction. Section 2 introduces the tools used in this study, including a description of the region, the data sets used and the methodology. Section 3 presents the results and evaluations while section 4 draws the conclusions and perspectives.

2 Materials and methods

2.1 Basins

Table 1 lists the basins considered in this study. They are also represented in Fig. 1. They were defined by first choosing river discharge (D) *in situ* measurement stations close to the sea, over the major Himalayan rivers, with a long-enough time record. The HydroSHEDS topography (Lehner et al., 2006) was then used to determine the drainage area and basin delineation. The basins were selected based on: (1) their spatial domain needs to be large enough compared to the spatial resolution of the GRACE instrument, (2) the river discharge measurements need to cover the GRACE period (2002-2015).

Five basins were chosen:

- *Mekong*: The Mekong Delta is one of the largest deltas in the world. It is a vast plain (55000 km^2) mostly lower than 5 m above sea level. Due to the seasonal variation in water level, the area presents extensive wetlands. The Mekong Delta region that represents only 12% of the total Vietnam area, allows 50% of the annual rice production (up to three harvests per year on some provinces), 50% of the fisheries, and 70% of the fruit production. Furthermore, questions related to oceanic water intrusions, change of agriculture practices (e.g. number of rice harvest in one year), dam construction, ground water pumping and resulting land depletion, all have an impact on the TWSC and would therefore benefit from its monitoring.

- *Ganges and Brahmaputra*: The Ganges-Brahmaputra is the major river basin of the Indian Sub-Continent supplying more than 700 millions people. It covers an area of 1.7 million km^2, at the crossroad of Bangladesh, India, China, Bhutan and Nepal and is the third largest freshwater provider to the world's oceans (after the Amazon and the Congo rivers) with a high influence on the regional climate. The basin is seasonally subject to the monsoon and faces strong climate variability between drought and floods periods. Furthermore, water management is an issue because of the increasing needs the population and the demands for the industry and agriculture sectors. Thus, the freshwater supply leads to an over-abstraction of groundwater stock during dry-season, and then to a rapid fall of groundwater tables.

- *Godavari*: The Godavari River is the second longest river of India after Ganges, covering a total drainage area of 312000 km^2 and accounting for nearly 9.5% of the total geographical area of the country. It flows for a length of about 1465 km, from its origin near the Arabian Sea before outfalling in to the Bay of Bengal, crossing several states of India. The basin receives its maximum rainfall during the southwest monsoon, from June to September. The major part of basin is covered with agricultural land accounting up to 60% of the total area, while 3.5% of the basin are covered by water bodies. Godavari basin faces several hydroclimatic problems with a large portion of the basin being prone to drought, while flooding problems are common in its lower reaches and its coastal areas are cyclone-prone.

- *Irrawaddy*: Running over a length of 2100km mainly within the boundaries of Myanmar, the Irrawaddy River is the most important river of the country. The basin takes up the northwestern part of the Indochina Peninsula, with its source on the south slopes of the Himalayas Mountains and emptying into the Andaman Sea of the Indian Ocean. The river basin area covers more than 400000 km^2 and collects 2/3 of the surface water volume of Myanmar. It is subject to a tropical and subequatorial monsoon climate and its hydrological regime, similarly to other large rivers of south Asia, is fed with water on the slopes of the Himalayas Mountains, mainly from rains during the southwest monsoon period and melt water of glaciers. The river is vital for human activities, water supply, and irrigation and hosts a high biodiversity. It is prone to extreme events, such as floods from very heavy monsoon rains or extreme weather events like cyclones and severe droughts and under climate change impacts, the region is facing major challenges for water resources.

2.2 Datasets

2.2.1 Datasets used in the integration

The datasets presented in this section will be used in the integration process to obtain an optimised description of the WC over the Himalayan basins. Most of them are satellite products. Only global satellite products have been considered. In order to integrate them, the datasets have been projected onto a common $0.25°$ spatial resolution grid using a conservative interpolation (Jones, 1999), and re-sampled at the monthly scale.

Precipitation, P - Three datasets based on remote-sensing observations have been selected. All are products calibrated using gauges measurements: the Global Precipitation Climatology Project (GPCP-V2, Adler et al., 2003); the Tropical Rainfall Measuring Mission Multi-satellite Precipitation Analysis (TMPA,3B42-V7, Huffman et al., 2007); and the Multi-Source Weighted-Ensemble Precipitation (MSEWP) dataset Beck et al. (2017). **All these global datasets are widely used in the community. GPCP and TMPA use a same Threshold Matched Precipitation Index (TMPI) algorithm to estimate instantaneous precipitation from multiple satellites by combining high-quality passive micro-wave observations and infrared data: their approach differ only in the use of gauge analyses (GPCC) to obtain calibrated estimates. While TMPA is based on an inverse random-error variance weighting of the gauge data, GPCP assumes that the precipitation distribution estimated from the combined satellite estimates is optimal and uses the gauge observations only for debiasing. The MSWEP dataset merges the highest quality precipitation data sources available as a function of timescale and location. It uses a combination of rain gauge measurements, the two previous satellite datasets, and a reanalysis. These datasets have been compared in terms of uncertainties and performance in Sun et al. (2018). It should be noted that these datasets are not independent from each other but represent the best up-to-date precipitation estimates for hydrological studies.**

Evapotranspiration, E - Three satellite-based products were chosen to describe evapotranspiration over land. All these datasets are assumed to be satellite-based products even if their retrieval algorithms can use auxiliary information and a model. **The Global Land Evaporation Amsterdam Model (GLEAM-V3B, Martens et al., 2016; Miralles et al., 2011) uses Priestley and Taylor (1972) empirical energy-based equation to calculate the reference evapotranspiration and separately estimate the different components of land evaporation: transpiration, bare-soil evaporation, interception loss, open-water evaporation and sublimation. GLEAM uses reanalysis (vA) or satellite (vB) precipitation inputs. The global observation-driven Penman-Monteith-Leuning (CSIRO, Zhang Yongqiang et al., 2016) evapotranspiration introduced by the Commonwealth Scientific and Industrial Research Organisation (CSIRO) and the MODIS Global Evapotranspiration Project (MOD16, Mu et al., 2011) are both evapotranspiration estimates based on Penman-Monteith equations (PENMAN, 1948; Monteith, 1965). We choose these three datasets due to their different equations for the evapotranspiration. Inter-comparison of global evapotranspiration algorithms and datasets can be found in (Michel et al., 2016).**

Total Water Storage Change (TWSC), ΔS - The TWSC estimates are all based on the GRACE satellite measurements (Tapley et al., 2004). These estimates include the surface (wetland, floodplains, lakes, rivers and man-made reservoirs), soil moisture, snow pack, glaciers and groundwater waters. Satellite datasets are based either on the spherical decomposition of GRACE(for instance (JPL, Watkins and Yuan, 2014) or on the "MASCON" solution: the Jet Propulsion Laboratory (JPL, Watkins et al., 2015a, MSC) product that also includes a scaling factor for hydrological coherency. **The CSR offers another MASCON solution. The CSR and JPL MASCON solutions differ in their processing: the JPL solution is based on an explicit estimation of mass anomalies at specific mass concentration block location using the analytical partial derivatives of the inter-satellite range-rate measurements (Watkins et al., 2015b). The CSR MASCON solution is first based on a spherical decomposition of the inter-satellite range-rate measurements that is truncated spatially at the location of mass concentration (Save et al., 2016). The two solutions have been compared to the spherical solutions in terms of uncertainty in both min-max range and trend, in (Scanlon et al., 2016; Save et al., 2016). We choose here the JPL solution because it is more independent of the spherical solution.**

Based on preliminary tests, it was observed that the MASCON solution for TWSC (ΔS) was in better agreement with the three other water component estimate, and in particular over the Irrawaddy basin, compared to the spherical solutions. This could be due to the local inversion in MASCON solution that prevent from the de-striping processing usually done in the spherical decomposition of GRACE. It as been shown that de-striping could limit the capability of spherical solution over particularly South/North oriented basin (Wahr et al., 2006; Rateb et al., 2017). In the following, the MASCON solution from JPL is used. Fig. 2 represents the GRACE TWSC and TWSC anomaly (with respect to averaged season computed over the 2002-2015 period), over the five basins of the previous section. The annual cycle is well pronounced in each basin, showing the strong seasonality of the WC in these regions. The anomalies have strong inter annual variations showing the evolution of the WC along the years.

Discharge, D - The Global Runoff Data Centre (GRDC) gathers discharge measurements at the global scale. However, for large tropical rivers, and more particularly over South Asia, only few stations are available and they are not all updated to recent periods. In particular, among the five considered rivers of this study (Fig. 1), four of them are not available at GRDC and we obtained them instead from personal communication and sharing of local colleagues (Table 1).

In the following, an *a priori* specification of the uncertainties for each one of these satellite sources are required. **Such characterizations are generally product-and-site-specific. Some studies (Pan et al., 2012; Sahoo et al., 2011; Zhang et al., 2017) estimate the *a priori* uncertainty of particular water components based on the spread among the various estimates (taking the spread of estimates as an estimate of the uncertainties can sometimes be dangerous). In our case, this approach would not take into account the fact that the precipitation estimates are not independent. The values used here are derived from (Munier et al., 2014) in which the authors reviewed carefully the literature on this topic. The partitioning of uncertainty between P and E has however been modified to allow larger uncertainty in P since datasets are dependent in our case. As the objective of the current study is to reconstruct GRACE TWSC, the approach assumes lower errors in the GRACE estimate that becomes our reference. For the three precipitation datasets, we specify a 14 mm/month STD (STandard Deviation) error. Similarly, for three evapotranspiration datasets, we specify a 7 mm/month STD. River discharge is an *in situ* measurement so a 3 mm/month STD is chosen. Since the objective is the reconstruction of the GRACE observations over long time series, we specify a small uncertainty (1 mm/month) to avoid changing these values during the integration.**

2.2.2 Datasets used in the evaluation

• *ISBA hydrological model* - To evaluate our reconstruction of the long-term evolution of TWSC over large Himalayan basins, we also use the ISBA-CTRIP numerical land surface system. ISBA-CTRIP is a “state of the art" hydrological system that simulates TWSC at the global scale with an good accuracy as shown in Decharme et al. (2019a). The ISBA-CTRIP TWSC all water mass variation (river water mass and floodplains, snowpack, canopy water, total soil moisture and groundwater storage). The ISBA land surface model explicitely solve the energy and water budgets at the land surface at any time step. The CTRIP river routing model simulates river discharges up to the ocean from the total runoff computed by ISBA. A two-way coupling between ISBA and CTRIP allows to account for, (1) a dynamic river flooding scheme with explicit interactions betwen the floodplains, the soil and the atmosphere (through free-water evaporation, precipitation interception and infiltration) and (2) a two-dimensional diffusive groundwater scheme to represent unconfined aquifers and upward capillarity fluxes into the superficial soil. More details can be found in Decharme et al. (2019a). In this study, we use a product derived from a global offline simulation at 0.5° resolution done with this ISBA-CTRIP configuration and driven at a 3-hourly time scale by the ERA-Interim reanalysis over the 1979-2015 period. To ensure that realistic precipitation are fed to the ISBA-CTRIP system (Szczypta et al., 2014), the ERA-Interim precipitation is, here, hybridized to match the monthly values from the gauge-based Global Precipitation Climatology Center (GPCC) Full Data Product V6 (Schneider et al., 2011, 2014). At each time step, ISBA-CTRIP gives the variation of the total mass of water. The TWSC estimate from ISBA-CTRIP is then the monthly average of this field, which is slightly different than the reconstruction via Eq. (1) (see Appendix A). Since GRACE data are anomalies relative to a reference geoid, the TWSC estimate from ISBA-CTRIP is also calculated in terms of anomaly over the analysed period. To be consistent with the GRACE data, the simulated TWS were smoothed using a 200 km-width Gaussian filter which is quasi-similar to the filter used for the GRACE products (Watkins et al., 2015a). In the following, ISBA-CTRIP is shortened in ISBA.

• *GLDAS hydrological model* - For comparison purpose, we also use the Noah 2.7.1 land hydrology model of the Global Land Data Assimilation System (GLDAS). Its purpose is to ingest satellite- and ground-based observations using advanced land surface modeling and data assimilation techniques, in order to generate optimal fields of land surface states and fluxes. GLDAS is an uncoupled land surface modeling system that drives multiple models runs globally at the resolution of 0.25°, and produces results in near-real time. The GLDAS system is described in (Rodell et al., 2004). GLDAS is a platform of assimilation and differs from hydrological models. In particular, they do not model reservoirs. For our comparison, we use the Land Water Content output of GLDAS.

These two global and well known models have been chosen for comparison even if none of them includes anthropogenic effects on the river discharge and groundwater storage. Significant efforts have been made during the last two decades to incorporate anthropogenic impacts in LSM (Hanasaki et al., 2006; Haddeland et al., 2014) but crucial challenges still remain. Most of these new schemes in LSMs have been developed and used offline for regional scale studies

and without common and standardized framework (Pokhrel et al., 2016; Döll et al., 2016). At global scale, a state of the art does not include the global representation of flow regulation and irrigation water needs.

2.3 Methodology

The notations are presented in this section but more methodological details can be found in Aires (2014),. The last version of the integration methodology is well described in Pellet et al. (2019).

2.3.1 Water cycle budget closure at basin scale

The first step of the integration process consists in merging the various datasets presented in Section 2.2.1. The "Simple Weighting" estimate (i.e. ensemble mean) is used to describe each water component based on all the available datasets (Aires, 2014):

$$P_{SW} = \frac{1}{p-1} \sum_{i=1}^{p} \frac{\sum_{k \neq i} (\sigma_k)^2}{\sum_k (\sigma_k)^2} P_i. \tag{2}$$

This equation is valid when there is no bias error in the P_is (thanks to a preliminary bias correction) and is optimal when the errors ϵ_i are statistically independent from each other. This expression is valid for the other water components. The variance of the P_{SW} uncertainties is then given by:

$$\sigma_{P_{SW}} = \frac{1}{(p-1)^2} \sum_{i=1}^{p} \left(\frac{\sum_{k \neq i} (\sigma_k)^2}{\sum_k (\sigma_k)^2} \right)^2 \sigma_i^2. \tag{3}$$

A similar approach is used for the three estimates of the evapotranspiration. Following the error specification in Section 2.2.1, the uncertainty of the precipitation (resp. evapotranspiration) merged estimate is characterized by a 8 mm/month STD (resp. 4 mm/month STD). Only one discharge dataset is available and only the MASCON-JPL is used for TWSC. We denote $X_{SW} = (P_{SW}, E_{SW}, D_{SW}, \Delta S_{SW})$ where SW stands for "Simple weighting" the results of the merging.

Following Aires (2014), it is then possible to write the conservation of water mass at the basin scale as a constraint applied on the state vector $X = (P, E, D, \Delta S)$. The WC budget constraint is expressed in Eq. (4). A relaxed constraint is considered (Pellet et al., 2019): the WC budget is closed within an error r that follows a normal distribution with specified uncertainty (Yilmaz et al., 2011). The problem can be written in the following way:

$$\begin{aligned} X^t &= (P,\ E,\ D,\ \Delta S) \\ G &= [1,\ -1,\ -1,\ -1] \\ X^t \cdot G^t &= r \text{ with } r \sim \mathcal{N}(0, \Sigma), \end{aligned} \tag{4}$$

where ^{t} is the transpose sign. G the closure operator and Σ the variance of the relaxation r. The optimised solution of this problem can be expressed as:

$$X_{PF} = (I - K_{PF} \cdot G\Sigma^{-1}G^t) \cdot X_{SW}, \tag{5}$$

where $K_{PF} = (B^{-1} + G\Sigma^{-1}G^t)^{-1}$, PF stands for the "Post-Filtering" of the previous solution X_{SW}, and B is the **a priori** error covariance matrix of X_{SW} that is specified here as :

$$B = \begin{pmatrix} 8 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}^2 . \tag{6}$$

This methodology allows obtaining a solution X_{PF} that closes the WC budget (within the relaxation r).

2.3.2 The calibration for the temporal extension of the closure constraint

An important limitation of the closure integration is the need for a common coverage period for all sources of information used in the integration. The optimised dataset cannot be provided for time steps with a missing component. To overcome this issue, a calibration to correct independently each water component towards the closure solution has been introduced (Munier et al., 2014). This calibration is based on a statistical regression between the merged observations X_{SW} and the optimised estimates X_{PF}, assuming that this optimised dataset represents the reference. In Pellet et al. (2019), the calibration is not strictly linear in order to avoid correcting null water fluxes (Munier and Aires, 2017). The following regression is used for P, E and D:

$$X_{CAL} = a \cdot X_{SW} + b \cdot (1 - e^{\frac{-X_{SW}}{c}}) \tag{7}$$

so that X_{CAL} becomes closer to X_{PF}. This calibration is close to a linear calibration, but zero-values are kept unchanged. a, b and c are the calibration parameters. The calibration is performed not only during the GRACE period (2002-2015) but over the complete record of each satellite dataset. It was shown in previous studies that the calibration does not allow for a perfect balance of WC, but it greatly reduces the WC budget residuals compared to the original estimates.

The calibration of Eq. (7) is applied independently on each dataset of Section 2.2.1. Table 2 shows the original (X_{SW}) versus the calibrated (X_{CAL}) Root Mean Squares of the WC budget residuals. It can be seen that the calibration is a significant improvement in each basin, with a decrease of the error from 25 to 54%.

Fig. 3 compares in row the original (X_{SW}) and calibrated (X_{CAL}) estimates of the four water components with the WC budget residuals, for the Ganges and the Brahmaputra basins, over the GRACE period. It can be seen that for the Ganges, the water components are not particularly impacted by the calibration. This is due to the overall coherency of the various water components estimates and the relatively low WC budget residual for this particular basin. Only a small improvement can be noticed in the WC budget residuals. For the Brahmaputra basin, precipitation and evapotranspiration are much more impacted by the calibration. The discharge is slightly changed because we specified a low uncertainty on this *in situ* variable (i.e. STD=1$mm/month$). The resulting WC budget residuals are much smaller for the calibrated solution meaning that this solution is more coherent hydrologically.

2.3.3 SAtellite Water Cycle (SAWC) reconstruction

Similarly to Munier et al. (2014), the three available water components (P, E, D) are used to estimate the fourth one, the TWSC ΔS. This allows extending temporally the monitoring of the TWSC before and after the GRACE period (2002-2015). In addition, GRACE satellite is down for maintenance every six month since 2011. The calibration approach allows filling gaps measurement in GRACE observation with high accurancy (Munier et al., 2014). Following Landerer et al. (2010) and to avoid temporal mismatching between GRACE-derived TWSC and the monthly estimate of the other water components, we use a centred differences of the mean TWS anomalies to compute the TWSC $\Delta S(t)$. The right-hand side of Eq. (1) is therefore computed for each month t as:

$$\Delta S(t) \quad = \tfrac{1}{4}X'_{CAL}(t-1) + \tfrac{1}{2}X'_{CAL}(t) + \tfrac{1}{4}X'_{CAL}(t+1) \tag{8}$$

where $X'_{CAL} = (P_{CAL} - E_{CAL} - D_{CAL})$, and ΔS is the centered mass rates:

$$\Delta S_t \quad = \tfrac{S(t+1)-S(t-1)}{2}. \tag{9}$$

2.3.4 Uncertainty characterisation

Estimating product uncertainty is a valuable information for instance in an assimilation framework. Most uncertainty estimation approaches require defining first a reference: *in situ*, reanalysis, or a consensus of all available data. In this work, the chosen reference is the optimized product (X_{PF}), this is a solution that is hydrologically more coherent and reliable than the original datasets. All original satellite datasets are compared to this reference to compute bias (not shown) and uncertainty (standard deviation) errors. For instance, for precipitation: $\sigma_P^2 = E[(P - P_{PF})^2]$. Such an approach was used in Munier and Aires (2017); Pellet et al. (2019).

Table 3 gathers the *a posteriori* uncertainty estimates **(computed as the distance between the original datasets and the reference)** for all the original satellite datasets, for P, E and D, over the five considered basins. **These *a posteriori* uncertainty estimates are in line with the specifications that were taken *a priori* for each of the datasets (Section 2.2.1)**. It can be seen that the Brahmaputra has higher uncertainties, especially for precipitation. MSEWP appears less reliable than GPCP or TMPA for precipitation; and GLEAM seems more reliable than MOD16 or CSIRO over these five basins.

It is possible to compute the SAWC reconstruction of TWSC based on Eq. (10). Once calibrated, P_{CAL}, E_{CAL} and D_{CAL} estimates are available over a long time period. They can be used to infer ΔS using the WC budget equation (Eq. (1)), before and after the GRACE period:

$$\Delta S_{SAWC} \quad = \quad P_{CAL} - E_{CAL} - D_{CAL}. \tag{10}$$

The reconstruction of TWSC has a different temporal coverage for the five basins because D **(and then D_{CAL})** availability varies (see Table 1). The measurement errors of P, E, and D of Table 3 are assumed to be independent and normally distributed. In this case, the error in the SAWC reconstruction of ΔS is given by:

$$\sigma_{\Delta S}^2 \quad = \quad \sigma_P^2 + \sigma_E^2 + \sigma_D^2, \tag{11}$$

where σ_P, σ_E, σ_D are **the *a posteriori* uncertainties merged estimate using Eq.** (3) **with the values of a posteriori estimate in Table 3**.

3 SAWC reconstruction of TWSC and evaluation

3.1 Evaluation over the GRACE period (2002-2015)

The resulting SAWC time series can be observed (red) and compared to GRACE measurements (blue) over the 2002-2015 period in Fig. 2, over the five basins, for the raw and the anomalies. The ISBA simulation is represented too (green). The seasonality is well represented in every estimate. The specific seasonality of each basin is well characterised by the SAWC reconstruction, see for instance the difference between the Mekong and the Brahmaputra seasons. The SAWC reconstruction uses the GRACE data to calibrate the other datasets, but once the P, E and D calibrations are done, the SAWC data in Eq. (10) does not use the GRACE data anymore. This is a good demonstration that GRACE-compatible TWSC estimates can be obtained from the other water components. The rich inter-annual signal in the anomalies is well captured too by SAWC times series. Some extreme years are well captured: e.g. the high extremes of the 2008 year over the Ganges basin, which are depicted also by the ISBA model.

In order to quantify the agreement of these times series, Fig. 4 represents the correlation, the correlation of the anomalie, and the Root Mean Square of the Difference (RMSD) between the four estimate (GRACE, SAWC, ISBA and GLDAS), over the five Himalayan basins, for the GRACE period (2002-2015). It can be see that the SAWC times series is highly correlated (0.96 on average) to the GRACE data, better than the ISBA (0.94); GLDAS has much lower agreement with GRACE (0.91) because it misses completely the season in the Irrawaddy basin for some years (not shown for clarity in Fig. 2). Again, it is not surprising that SAWC is close to GRACE because it has been designed to do that.

In terms of correlation of anomalies, SAWC estimate is always closer to ISBA than to GRACE, even if SAWC has high correlation of anomalies with GRACE (between 0.69 and 0.79) except over the Brahmaputra basin (0.36). This will be analyzed below. Comparatively, GLDAS estimate is less correlated to GRACE over the four basins (except Brahmaputra basin). The RMSD and RMSD of anomalies show similar pattern that of the correlation values, over all the basins. The RMSD statistics are better for the SAWC (19 mm/month error) than for ISBA (26 mm/month error), but this is no surprise because the season is a large part of these discrepancies. GRACE has a low spatial resolution (300 km^2 at the equator), this can decrease the accuracy of the TWSC anomaly estimates over small basins (e.g. Godavari or Irrawaddy). The smaller the basin is, the larger the gap between SAWC-GRACE and SAWC-ISBA correlation becomes. SAWC estimate is based on precipitation and evapotranspiration obtained at a finner spatial resolution (0.25°) than GRACE. Therefore, SAWC, as ISBA (at the 0.25° spatial resolution) represents better the anomaly over small basins as far as precipitation and evapotranspiration are accurate enough. Over the Brahmaputra basin, the large uncertainty of satellite evapotranspiration products over the mountains (see the impact of the calibration for the evapotranspiration estimate over this basin in Fig. 3) might impact the SAWC TWSC accuracy and explain why GLDAS and ISBA are better over this basin. This assumption is later confirmed in Fig. 6 in which precipitation in ISBA and SAWC are close

but the anomalies of E differ. Overall, it can be concluded that SAWC seems closer to GRACE than ISBA, for some events, as seen in the anomalies, over the Brahmaputra basin. GLDAS has a lower agreement with ISBA, in particular over the Irrawaddy basin. The discrepancy between simulated TWSC from ISBA and GLDAS can be explained by the different representation of aquifers in these two models. While a two-dimensional diffusive groundwater scheme in ISBA represents unconfined aquifer processes (Vergnes and Decharme, 2012; Vergnes et al., 2012), the Noah land model used in the GLDAS simulations did not include surface and groundwater storage. Therefore, the simulated mean seasonal cycle and the inter-annual variability of the TWSC is improved in ISBA (Decharme et al., 2019b). On the contrary, deviations from GRACE TWSC can thus be expected with GLDAS (Syed et al., 2008). Based on these results SAWC solution is compared only to ISBA in the following over the long time period.

3.2 Comparison with ISBA TWSC

In Fig. 5, the SAWC reconstruction is compared to the ISBA simulation over a long time record. ISBA is available from 1980 to 2015, SAWC is available based on the river discharge *in situ* measurements coverage (see Table 1). The first important remark to be done is the very good seasonal cycle agreement between SAWC and ISBA: correlations are larger than 0.93, except for the Brahmaputra with a correlation of 0.88. This basin presents a particular water cycle in 2007 that is analyzed in Appendix (Fig. A3). In the following, the year 2007 has been removed from the comparison analysis. For the Ganges basin, the main difference between the two estimates is that ISBA represents larger negative seasonal peaks and a slight phase in the seasons over the Mekong. In terms of seasonal anomalies, the agreement is also satisfactory with correlations between 0.61 (Godavari and Irrawaddy) to 0.78 (Mekong that is always well represented in our analysis), except again for the Brahmaputra (0.22 correlation). However, based on the 2007 analysis over the Brahmaputra (Fig. A3), we believe that the SAWC anomalies might be more reliable because they use measured *in situ* D compared to the models.

In Fig. 6, we analyse the long-term TWSC time series in terms of anomalies with respect to the climatological season. Furthermore, the times series of these anomalies have been smoothed using a three-year moving window. For instance, a peak value of 10 mm/month means that the time series was on average 10 mm higher than the climatological season, for 3 consecutive years (i.e. 360 accumulated mm in three years).

In general, SAWC reproduces well the long-term anomalies of the MSEWP precipitation dataset. This satellite dataset was used as input with two close other products (TMPA and GPCP calibrated using the same precipitation gauges) for the SAWC reconstruction (while ISBA uses a mix between GPCC and ECMWF reanalysis, see section 2.b.2). In general, ISBA precipitation inputs has some differences with MSEWP during the 80's and in 2010-2015, this requires further investigations beyond the scope of this study. The evapotranspiration anomalies are relatively flat for all basins, except for the Godavari where both SAWC and ISBA are in good agreement. By construction, the discharge D measurements are well reproduced by SAWC, but some significant differences can be observed for the ISBA model. These important temporal variations of the D anomalies have an important impact on the other WC components of the SAWC reconstruction. TWSC anomalies ΔS have a rather constant behaviour in the ISBA analysis, but large variations are present in the SAWC reconstruction. For instance, there is a large water deficit in the 1990-1991 over the Brahmaputra, or over the Ganges in 1985-1887.

From this comparison, the following conclusions can be drawn. When precipitation from SAWC matches well precipitation used to force ISBA, discharge simulated by ISBA is quite close to *in situ* measurements (discharge from SAWC), as for the Mekong and the Godavari basins, which could be seen as an indicator of the good quality of P_{SAWC}. On the contrary, main differences between D_{ISBA} and D_{SAWC} are either due to large differences between precipitations or to the TWSC dynamics. In ISBA, the groundwater storage is a simple delayed reservoir (with a constant delay parameter) which tends to attenuate the river dynamics. It is then not able to simulate long term groundwater dynamics (Pedinotti et al., 2012). Moreover, the ISBA model does not represent anthropogenic factors such as groundwater extraction, river regulation or irrigation, which may significantly impact river discharges. **For instance, in Fig. 6, the Mekong river discharge anomalies show lower min-max range in the observations than in ISBA. Li et al. (2017) highlight the impact of the construction of the Xiaowan and the Nuozhadu dams starting in 1991. The dam reduces the streamflow in particularly wet seasons and increases the streamflow in particularly dry seasons which lowers the anomaly variations. For this basin, D is more correlated to precipitation in ISBA (0.94) than in SAWC (0.63) solutions. This shows that modeled D is more straightforwardly dependent of the precipitation than in observations. On the contrary, TWSC anomaly is less linked to precipitation in the ISBA model than in SAWC where natural recharge is better represented. This difference is also discussed in the Appendix A.** The integration of anthropogenic processes are currently under development, as well as alternatives like data assimilation (Emery et al., 2018; Albergel et al., 2017).

3.3 Indirect evaluation using GIEMS inundation area

An important component of the TWS corresponds to the surface waters. The GIEMS (Global Inundation Extent from Multi-Satellite) database provides an estimation of the inundation extent from 1993 to 2007, at the global scale, on a 0.25° resolution equal-area grid (Prigent et al., 2007). GIEMS was fully assessed over Asian basins, especially using GRACE data (Papa et al., 2008). The SAWC reconstruction of TWSC and the inundation area time series are represented jointly in Fig. 7 to measure their coherency. Since surface waters are only one part of the TWS, we do not expect a perfect match between the two times series. However, the coherency between them is noticeable, correlation ranges from 0.76 (Godavari) to 0.85 (Ganges). Furthermore, the inter-annual variability of both times series can be measured by the seasonal anomalies, their correlations are significative; they oscillate from 0.24 (Irrawaddy) to 0.42 (Ganges) except for the Brahmaputra where problems were already noticed (see Fig. 6). This comparison is not a direct evaluation of the TWSC, but the fact that coherency can be found between two completely independent measurements on the water cycle is a positive point for the evaluation of the SAWC reconstruction.

4 Conclusion and perspectives

The Total Water Storage (TWS) and its Changes (TWSC) is a crucial element of the water cycle because of its impact on water management, and its role of tracer of human activity. The first measurement available to monitor it came from the GRACE instrument in 2002. Longer time records being necessary for climate studies, we proposed here to use satellite observations for precipitation and evapotranspiration with *in situ* river discharge measurements to estimate the TWSC over the period 1980-

2015. Our approach is based on the water conservation rule over each basin. We performed this task over five major Asian basins because their evolution is related to importante questions about water management, climate change, and land-use changes. Our SAWC reconstruction of TWSC has been evaluated using (1) GRACE observations (over the 2002-2007 period), (2) ISBA model simulations (1980-2015), and (3) surface inundation area (1993-2007). The seasonality and inter-annual variability of SAWC's TWS appear coherent with these independent sources of information.

The advantages of the proposed methodology are numerous. It is an integration method that gathers all the observations available, satellite and *in situ* measurements. Contrarily to traditional assimilation, this methodology does not use any land or hydrological model, except the water conservation law. It uses the multiplicity of observations to reduce uncertainties on each one of the water cycle components, and introduces more hydrological coherency among them. If the river discharge measurement is available, it allows to handle a true anthropized discharge (not an idealized natural one, as provided by models in most cases). However, if this discharge is not available, the methodology cannot be applied; and if important uncertainties on discharge measurements are present, these errors will be propagated to all the other water components.

We foresee many perspectives for this work. First, we would like to extend the TWSC estimation to other basins. This work can be done over large basins (compatible with the GRACE spatial resolution) and where the *in situ* river discharge measurements are available. Once this is done over a sufficient number of large basins, the optimized databases can be used as a reference to calibrate the satellite estimates at the global scale. This allows for the use of the satellite observations at the global scale and not only over the basins where the integration was performed (Munier and Aires, 2017). River discharges could potentially be estimated over un-monitored basins, or over longer time series than the monitoring. Total water storage could also be estimated over monitored basins, over longer times series than the GRACE record (as it was done here). When discharge is not monitored, the use of modeled river discharges could be attempted.

Our methodology can also be used to detect incoherencies in our estimations of the water cycle components. For instance, large budget errors could indicate regions where the evapotranspiration is biased (e.g. due to an under- or over-estimation of the irrigation as in the Nile basin). Our approach could detect such problems and propose a bias-correction of the incriminated water component.

Finally, we expect to use a similar methodology over connected sub-basins. It is possible to estimate the surface water storage using water extent and topography (Papa et al., 2015), but the horizontal underground transport of water cannot be measured so far. The difference of total water storage and surface water storage should help us estimate the ground water storage and characterise its horizontal transport. This would be a major achievement as this important process of the global water cycle is largely unknown so far.

Appendix A: Computation of Total Water Storage Change (TWSC)

For a given month, the TWSC corresponds to the variation of the Total Water Storage (TWS) between the first day of the month and the first day of the following month. As show in Eq. (1), TWSC equals the sum of inflows into the domain (P) minus outflows out from the domain ($E+Q$) during the whole month (P, E and Q represent monthly averages). The ISBA land surface model may provide all variables, including TWS, at a daily time step, so that it is possible to compute TWSC as the above-mentioned difference. By construction, the water balance is closed in the ISBA model, and the absolute difference between TWSC and $(P - E - Q)$ does not exceed 10^{-6} mm/month (with a RMS of 10^{-9} mm/month). On the other hand, it is not possible to compute the exact TWSC with GRACE data since TWS at the beginning of each month is not available. Instead, GRACE data correspond to monthly averages of TWS anomalies (temporal mean removed). To approximate TWSC, we used the centred difference from Eq. (9). Yet, this approximation introduces important errors compared to the true TWSC. Fig. A1 (left) shows the impact of this approximation with ISBA outputs (where the true TWSC is represented by $P - E - Q$). To reduce this error, we followed Landerer et al. (2010) by computing $P - E - Q$ using Eq. (6). Fig. A1 (right) shows a better match between the approximated $P - E - Q$ and the centred TWSC. Nevertheless, the reader should keep in mind that both approximations increase final uncertainties of about 5 to10 mm/month. This error has a quite high frequency and is reduced to less than 1 mm/month when using a 3-year moving average as in Fig. 7.

Appendix B: Zoom on 2007 event over Brahmaputra

An extreme inundation has occurred in the Spring of 2007 over the Brahmaputra basin (Gouweleeuw et al., 2018; Islam et al., 2010; Webster et al., 2010). It is interesting here to analyse how this was handled in the SAWC reconstruction and the ISBA model. The Fig. A3 compares them for the four water component estimates. The climatological seasons are also represented (dashed lines) for comparison purpose. Two basins are illustrated: the Ganges and the Brahmaputra.

In this sample, it can be seen that the model follows a classical seasonality for each water component and both basins. In the SAWC reconstruction, the seasonal anomaly is well reproduced for the discharge D, which was expected since this observation was used in the SAWC integration process. This translates into a pronounced anomaly in TWSC. This shows the benefits and the risks associated with the SAWC reconstruction: if the *in situ* D is reliable, then SAWC will reproduce it well and the impact on the other components can be important. If D measurements are erroneous, this can introduce considerable noise into the WC analysis.

This relates also to the question of the natural versus real/anthropised discharge D. Hydrological model will generally consider natural rivers. It is difficult to obtain all the necessary information to model true discharge (dams management, pumping for irrigation, etc.). An interesting way to constrain models to follow *in situ* measurements of the discharge would be to assimilate these measurements into the model (Wang et al., 2018).

Acknowledgements. The authors are grateful to Diego Fernandez for fruitful discussions on the water cycle monitoring during the ESA-WACMOS-MED project.

References

Adler, R. F., Huffman, G. J., Chang, A., Ferraro, R., Xie, P.-P., Janowiak, J., Rudolf, B., Schneider, U., Curtis, S., Bolvin, D., Gruber, A., Susskind, J., Arkin, P., and Nelkin, E.: The Version-2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979–Present), J. Hydrometeorol., 4, 1147–1167, https://doi.org/10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2, http://journals.ametsoc.org/doi/abs/10.1175/1525-7541%282003%29004%3C1147%3ATVGPCP%3E2.0.CO%3B2, 2003.

Aires, F.: Combining Datasets of Satellite-Retrieved Products. Part I: Methodology and Water Budget Closure, J. Hydrometeorol., 15, 1677–1691, https://doi.org/10.1175/JHM-D-13-0148.1, http://journals.ametsoc.org/doi/abs/10.1175/JHM-D-13-0148.1, 2014.

Albergel, C., Munier, S., Jennifer Leroux, D., Dewaele, H., Fairbairn, D., Lavinia Barbu, A., Gelati, E., Dorigo, W., Faroux, S., Meurey, C., Le Moigne, P., Decharme, B., Mahfouf, J. F., and Calvet, J. C.: Sequential assimilation of satellite-derived vegetation and soil moisture products using SURFEX-v8.0: LDAS-Monde assessment over the Euro-Mediterranean area, Geosci. Model Dev., 10, 3889–3912, https://doi.org/10.5194/gmd-10-3889-2017, 2017.

Asoka, A., Gleeson, T., Wada, Y., and Mishra, V.: Relative contribution of monsoon precipitation and pumping to changes in groundwater storage in India, Nat. Geosci., 10, 109–117, https://doi.org/10.1038/ngeo2869, http://www.nature.com/articles/ngeo2869, 2017.

Azarderakhsh, M., Rossow, W. B., Papa, F., Norouzi, M., and Khanbilvardi, R.: Diagnosing water variations with the Amazon basin using satellite data, J. Geophys. Res., 116, 2011.

Babel, M. S. and Wahid, S. M.: Freshawater under threat South Asia: Vulnerability Assessment of Freshwater Resources to Environmental Change, Tech. rep., United Nations Environment Program, Nairobi, 2008.

Beck, H. E., van Dijk, A. I. J. M., Levizzani, V., Schellekens, J., Miralles, D. G., Martens, B., and de Roo, A.: MSWEP: 3-hourly 0.25deg; global gridded precipitation by merging gauge, satellite, and reanalysis data, Hydrol. Earth Syst. Sci., 21, 589–615, https://doi.org/10.5194/hess-21-589-2017, https://www.hydrol-earth-syst-sci.net/21/589/2017/, 2017.

Chen, H., Zhang, W., Nie, N., and Guo, Y.: Long-term groundwater storage variations estimated in the Songhua River Basin by using GRACE products, land surface models, and in-situ observations, Sci. Total Environ., 649, 372–387, https://doi.org/10.1016/j.scitotenv.2018.08.352, 2019.

Decharme, B., Delire, C., Minvielle, M., and Colin, J.: Recent changes in the ISBA-CTRIP land surface system for using in the CNRM-CM6 climate model and in global off-line hydrological applications, J. Adv. Model. Earth Syst., 14, 1–92, 2019a.

Decharme, B., Delire, C., Minvielle, M., Colin, J., Vergnes, J., Alias, A., Saint-Martin, D., Séférian, R., Sénési, S., and Voldoire, A.: Recent Changes in the ISBA-CTRIP Land Surface System for Use in the CNRM-CM6 Climate Model and in Global Off-Line Hydrological Applications, J. Adv. Model. Earth Syst., 11, 1207–1252, https://doi.org/10.1029/2018MS001545, https://onlinelibrary.wiley.com/doi/abs/10.1029/2018MS001545, 2019b.

Döll, P., Douville, H., Güntner, A., Müller Schmied, H., and Wada, Y.: Modelling Freshwater Resources at the Global Scale: Challenges and Prospects, https://doi.org/10.1007/s10712-015-9343-1, 2016.

Emery, C. M., Paris, A., Biancamaria, S., Boone, A., Calmant, S., Garambois, P. A., and Da Silva, J. S.: Large-scale hydrological model river storage and discharge correction using a satellite altimetry-based discharge product, Hydrol. Earth Syst. Sci., 22, 2135–2162, https://doi.org/10.5194/hess-22-2135-2018, 2018.

Gouweleeuw, B., Gruber, C., Flechtner, F., Güntner, A., Gain, A., Kvas, A., and Mayer-Gürr, T.: Daily GRACE gravity field solutions track major flood events in the Ganges-Brahmaputra Delta, Hydrol. Earth Syst. Sci., pp. 2867–2880, 2018.

Haddeland, I., Heinke, J., Biemans, H., Eisner, S., Flörke, M., Hanasaki, N., Konzmann, M., Ludwig, F., Masaki, Y., Schewe, J., Stacke, T., Tessler, Z. D., Wada, Y., and Wisser, D.: Global water resources affected by human interventions and climate change, Proc. Natl. Acad. Sci. U. S. A., 111, 3251–3256, https://doi.org/10.1073/pnas.1222475110, 2014.

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

Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G., Hong, Y., Bowman, K. P., and Stocker, E. F.: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales, J. Hydrometeorol., 8, 38–55, https://doi.org/10.1175/JHM560.1, http://journals.ametsoc.org/doi/abs/10.1175/JHM560.1, 2007.

Humphrey, V., Gudmundsson, L., and Seneviratne, S. I.: A global reconstruction of climate-driven subdecadal water storage variability, Geophys. Res. Lett., 44, 2300–2309, https://doi.org/10.1002/2017GL072564, https://onlinelibrary.wiley.com/doi/abs/10.1002/2017GL072564, 2017.

IPCC: IPCC, 2014: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzer, Tech. rep., 2014.

Islam, A., Bala, S., and Haque, M.: Flood inundation map of Bangladesh using MODIS time-series images, J. Flood Risk Manag., 3, 210–222, https://doi.org/10.1111/j.1753-318X.2010.01074.x, http://doi.wiley.com/10.1111/j.1753-318X.2010.01074.x, 2010.

Jones, P. W.: First- and Second-Order Conservative Remapping Schemes for Grids in Spherical Coordinates, Mon. Weather Rev., 127, 2204–2210, https://doi.org/10.1175/1520-0493(1999)127<2204:FASOCR>2.0.CO;2, 1999.

Khaki, M., Forootan, E., Kuhn, M., Awange, J., Papa, F., and Shum, C.: A study of Bangladesh's sub-surface water storages using satellite products and data assimilation scheme, Sci. Total Environ., 625, 963–977, https://doi.org/10.1016/J.SCITOTENV.2017.12.289, https://www.sciencedirect.com/science/article/pii/S0048969717337221, 2018.

Landerer, F. W., Dickey, J. O., and Güntner, A.: Terrestrial water budget of the Eurasian pan-Arctic from GRACE satellite measurements during 2003-2009, J. Geophys. Res. Atmos., 115, 1–14, https://doi.org/10.1029/2010JD014584, 2010.

Lawford, R. G., Roads, J., Lettenmaier, D. P., and Arkin, P.: GEWEX Contributions to Large-Scale Hydrometeorology, J. Hydrometeorol., 8, https://doi.org/10.1175/JHM608.1, http://journals.ametsoc.org/doi/pdf/10.1175/JHM608.1, 2007.

Lehner, B., Verdin, K., and Jarvis, A.: HydroSHEDS Technical Documentation. World Wildlife Fund US, Tech. rep., http://hydrosheds.cr.usgs.gov, 2006.

Li, D., Long, D., Zhao, J., Lu, H., and Hong, Y.: Observed changes in flow regimes in the Mekong River basin, J. Hydrol., 551, 217–232, https://doi.org/10.1016/j.jhydrol.2017.05.061, 2017.

Lorenz, C., Kunstmann, H., Devaraju, B., Tourian, M. J., Sneeuw, N., and Riegger, J.: Large-scale runoff from landmasses: a global assessment of the closure of the hydrological and atmospheric water balances, J. Hydrometeorol., p. 140508110258003, https://doi.org/10.1175/JHM-D-13-0157.1, http://journals.ametsoc.org/doi/abs/10.1175/JHM-D-13-0157.1, 2014.

Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Férnandez-Prieto, D., Beck, H. E., Dorigo, W. A., and Verhoest, N. E. C.: GLEAM v3: satellite-based land evaporation and root-zone soil moisture, Geosci. Model Dev. Discuss., pp. 1–36, https://doi.org/10.5194/gmd-2016-162, http://www.geosci-model-dev-discuss.net/gmd-2016-162/, 2016.

Michel, D., Jiménez, C., Miralles, D. G., Jung, M., Hirschi, M., Ershadi, A., Martens, B., Mccabe, M. F., Fisher, J. B., Mu, Q., Seneviratne, S. I., Wood, E. F., and Fernández-Prieto, D.: The WACMOS-ET project - Part 1: Tower-scale evaluation of four remote-sensing-based evapotranspiration algorithms, Hydrol. Earth Syst. Sci., 20, 803–822, https://doi.org/10.5194/hess-20-803-2016, 2016.

Miralles, D. G., De Jeu, R. A. M., Gash, J. H., Holmes, T. R. H., and Dolman, A. J.: Magnitude and variability of land evaporation and its components at the global scale, Hydrol. Earth Syst. Sci, 15, 967–981, https://doi.org/10.5194/hess-15-967-2011, www.hydrol-earth-syst-sci.net/15/967/2011/, 2011.

Monteith, J. L.: Evaporation and environment., 1965.

Mu, Q., Zhao, M., and Running, S. W.: Improvements to a MODIS global terrestrial evapotranspiration algorithm, Remote Sens. Environ., 115, 1781–1800, https://doi.org/10.1016/j.rse.2011.02.019, http://linkinghub.elsevier.com/retrieve/pii/S0034425711000691, 2011.

Munier, S. and Aires, F.: A new global method of satellite dataset merging and quality characterization constrained by the terrestrial water cycle budget, Remote Sens. Environ., 205, 119–203, 2017.

Munier, S., Aires, F., Schlaffer, S., Prigent, C., Papa, F., Maisongrande, P., and Pan, M.: Combining data sets of satellite-retrieved products for basin-scale water balance study: 2. Evaluation on the Mississippi Basin and closure correction model, J. Geophys. Res., 119, 12,100–12,116, https://doi.org/10.1002/2014JD021953, 2014.

Pan, M. and Wood, E. F.: Data Assimilation for Estimating the Terrestrial Water Budget Using a Constrained Ensemble Kalman Filter, J. Hydrometeorol., 7, 534–547, https://doi.org/10.1175/JHM495.1, http://journals.ametsoc.org/doi/abs/10.1175/jhm495.1, 2006.

Pan, M., Sahoo, A. K., Troy, T. J., Vinukollu, R. K., Sheffield, J., Wood, and F, E.: Multisource estimation of long-term terrestrial water budget for major global river basins, J. Clim., 25, 3191–3206, https://doi.org/10.1175/JCLI-D-11-00300.1, 2012.

Papa, F., Güntner, A., Frappart, F., Prigent, C., and Rossow, W. B.: Variations of surface water extent and water storage in large river basins: A comparison of different global data sources, Geophys. Res. Lett., 35, L11 401, https://doi.org/10.1029/2008GL033857, http://doi.wiley.com/10.1029/2008GL033857, 2008.

Papa, F., Durand, F., Rossow, W. B., Rahman, A., and Bala, S. K.: Satellite altimeter-derived monthly discharge of the Ganga-Brahmaputra River and its seasonal to interannual variations from 1993 to 2008, J. Geophys. Res. Ocean., 115, 1–19, https://doi.org/10.1029/2009JC006075, 2010.

Papa, F., Bala, S. K., Pandey, R. K., Durand, F., Gopalakrishna, V. V., Rahman, A., and Rossow, W. B.: Ganga-Brahmaputra river discharge from Jason-2 radar altimetry: An update to the long-term satellite-derived estimates of continental freshwater forcing flux into the Bay of Bengal, J. Geophys. Res. Ocean., 117, 1–13, https://doi.org/10.1029/2012JC008158, 2012.

Papa, F., Frappart, F., Malbeteau, Y., Shamsudduha, M., Vuruputur, V., Sekhar, M., Ramillien, G., Prigent, C., Aires, F., Pandey, R. K., Bala, S., and Calmant, S.: Satellite-derived surface and sub-surface water storage in the Ganges-Brahmaputra River Basin, J. Hydrol. Reg. Stud., 4, 15–35, https://doi.org/10.1016/j.ejrh.2015.03.004, http://dx.doi.org/10.1016/j.ejrh.2015.03.004, 2015.

Pedinotti, V., Boone, A., Decharme, B., Crétaux, J. F., Mognard, N., Panthou, G., Papa, F., and Tanimoun, B. A.: Evaluation of the ISBA-TRIP continental hydrologic system over the Niger basin using in situ and satellite derived datasets, Hydrol. Earth Syst. Sci., 16, 1745–1773, https://doi.org/10.5194/hess-16-1745-2012, 2012.

Pellet, V., Aires, F., Mariotti, A., and Fernando, D.: Analyzing the Mediterranean water cycle via satellite data integration., Pure Appl. Geophys., https://doi.org/10.1007/s00024-018-1912-z, 2018.

Pellet, V., Aires, F., Munier, S., Fernández Prieto, D., Jordá, G., Arnoud Dorigo, W., Polcher, J., and Brocca, L.: Integrating multiple satellite observations into a coherent dataset to monitor the full water cycle - Application to the Mediterranean region, Hydrol. Earth Syst. Sci., 23, https://doi.org/10.5194/hess-23-465-2019, 2019.

PENMAN, H. L.: Natural evaporation from open water, hare soil and grass, Proc. R. Soc. Lond. A. Math. Phys. Sci., 193, 120–145, https://doi.org/10.1098/rspa.1948.0037, 1948.

Pokhrel, Y. N., Hanasaki, N., Wada, Y., and Kim, H.: 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, https://doi.org/10.1002/wat2.1150, http://doi.wiley.com/10.1002/wat2.1150, 2016.

Priestley, C. and Taylor, R.: On the Assessment of Surface Heat Flux and Evaporation Using Large-Scale Parameters, Mon. Weather Rev., 100, 81–92, https://doi.org/10.1175/1520-0493(1972)100<0081:OTAOSH>2.3.CO;2, http://journals.ametsoc.org/doi/abs/10.1175/1520-0493%281972%29100%3C0081%3AOTAOSH%3E2.3.CO%3B2, 1972.

Prigent, C., Papa, F., Aires, F., Rossow, W. B., and Matthews, E.: Global inundation dynamics inferred from multiple satellite observations, 1993-2000, J. Geophys. Res. Atmos., 112, https://doi.org/10.1029/2006JD007847, http://onlinelibrary.wiley.com.ezproxy.cul.columbia.edu/doi/10.1029/2006JD007847/full, 2007.

Rateb, A., Kuo, C.-Y., Imani, M., Tseng, K.-H., Lan, W.-H., Ching, K.-E., and Tseng, T.-P.: Terrestrial Water Storage in African Hydrological Regimes Derived from GRACE Mission Data: Intercomparison of Spherical Harmonics, Mass Concentration, and Scalar Slepian Methods, Sensors, 17, 566, https://doi.org/10.3390/s17030566, http://www.ncbi.nlm.nih.gov/pubmed/28287453http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5375852http://www.mdpi.com/1424-8220/17/3/566, 2017.

Rodell, B. Y. M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-j., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D., and Toll, D.: THE GLOBAL LAND DATA ASSIMILATION SYSTEM This powerful new land surface modeling system integrates data from advanced observing systems to support improved forecast model initialization and hydrometeorological investigations, Bull. Amer. Meteor. Soc., 85, 381–394, https://doi.org/10.1, https://journals.ametsoc.org/doi/pdf/10.1175/BAMS-85-3-381, 2004.

Rodell, M., Velicogna, I., and Famiglietti, J. S.: Satellite-based estimates of groundwater depletion in India, Nature, 460, 999–1002, https://doi.org/10.1038/nature08238, http://www.nature.com/doifinder/10.1038/nature08238, 2009.

Rodell, M., Beaudoing, H., L'Ecuyer, T., Olson, W., Famiglietti, J., Houser, P., Adler, R., Bosilovich, M., Clayson, C., Chambers, D., Clark, E., Fetzer, E., Gao, X., Gu, G., Hilburn, K., Huffman, G., Lettenmaier, D., Liu, W., Robertson, F., Schlosser, C., Sheffield, J., and Wood, E.: The Observed State of the Water Cycle in the Early 21st Century, J. Clim., pp. 8289–8318, https://doi.org/10.1175/JCLI-D-14-00555.1, http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-14-00555.1, 2015.

Rodell, M., Famiglietti, J. S., Wiese, D. N., Reager, J. T., Beaudoing, H. K., Landerer, F. W., and Lo, M. H.: Emerging trends in global freshwater availability, Nature, 557, 651–659, https://doi.org/10.1038/s41586-018-0123-1, 2018.

Sahoo, A. K., Pan, M., Troy, T. J., Vinukollu, R. K., Sheffield, J., and Wood, E. F.: Reconciling the global terrestrial water budget using satellite remote sensing, Remote Sens. Environ., 115, 1850–1865, https://doi.org/10.1016/j.rse.2011.03.009, 2011.

Salameh, E., Frappart, F., Papa, F., Güntner, A., Venugopal, V., Getirana, A., Prigent, C., Aires, F., Labat, D., and Laignel, B.: Fifteen years (1993-2007) of surface freshwater storage variability in the ganges-brahmaputra river basin using multi-satellite observations, Water (Switzerland), 9, https://doi.org/10.3390/w9040245, 2017.

Save, H., Bettadpur, S., and Tapley, B. D.: High-resolution CSR GRACE RL05 mascons, J. Geophys. Res. Solid Earth, 121, 7547–7569, https://doi.org/10.1002/2016JB013007, http://doi.wiley.com/10.1002/2016JB013007, 2016.

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

Schneider, U., Rudolf, B., Becker, A., Ziese, M., Finger, P., Meyer-Christoffer, A., and Schneider, U.: GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle, Theor. Appl. Climatol., 115, 15–40, https://doi.org/10.1007/s00704-013-0860-x, 2011.

Schneider, U., Becker, A., Ziese, M., and Rudolf, B.: Global Precipitation Analysis Products of the GPCC, Internet Publ., pp. 1–13, ftp://ftp-anon.dwd.de/pub/data/gpcc/PDF/GPCC_intro_products_2008.pdf, 2014.

Shamsudduha, M. and Panda, D. K.: Spatio-temporal changes in terrestrial water storage in the Himalayan river basins and risks to water security in the region: A review, Int. J. Disaster Risk Reduct., 35, 101 068, https://doi.org/10.1016/j.ijdrr.2019.101068, https://doi.org/10.1016/j.ijdrr.2019.101068, 2019.

Sheffield, J., Ferguson, C. R., Troy, T. J., Wood, E. F., and McCabe, M. F.: Closing the terrestrial water budget from satellite remote sensing, Geophys. Res. Lett., 36, 1–5, https://doi.org/10.1029/2009GL037338, 2009.

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., and Hsu, K.: A Review of Global Precipitation Data Sets: Data Sources, Estimation, and Intercomparisons, Rev. Geophys., 56, 79–107, https://doi.org/10.1002/2017RG000574, https://onlinelibrary.wiley.com/doi/abs/10.1002/2017RG000574, 2018.

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

Szczypta, C., Calvet, J.-C., Maignan, F., Dorigo, W., Baret, F., and Ciais, P.: Suitability of modelled and remotely sensed essential climate variables for monitoring Euro-Mediterranean droughts, Geosci. Model Dev., 7, 931–946, https://doi.org/10.5194/gmd-7-931-2014, https://www.geosci-model-dev.net/7/931/2014/, 2014.

Tang, Y., Hooshyar, M., Zhu, T., Ringler, C., Sun, A. Y., Long, D., and Wang, D.: Reconstructing annual groundwater storage changes in a large-scale irrigation region using GRACE data and Budyko model, J. Hydrol., 551, 397–406, https://doi.org/10.1016/j.jhydrol.2017.06.021, 2017.

Tapley, B. D., Bettadpur, S., Watkins, M., and Reigber, C.: The gravity recovery and climate experiment: Mission overview and early results, Geophys. Res. Lett., 31, https://doi.org/10.1029/2004GL019920, 2004.

Tiwari, V. M., Wahr, J., and Swenson, S.: Dwindling groundwater resources in northern India, from satellite gravity observations, Geophys. Res. Lett., 36, L18 401, https://doi.org/10.1029/2009GL039401, http://doi.wiley.com/10.1029/2009GL039401, 2009.

Tootchi, A., Jost, A., and Ducharne, A.: Multi-source global wetland mapping: combining surface water imagery and groundwater constraints, Hydrol. Earth Syst. Sci. Discuss., pp. 1–34, https://doi.org/10.5194/hess-2018-48, https://www.hydrol-earth-syst-sci-discuss.net/hess-2018-48/, 2018.

Vergnes, J.-P. and Decharme, B.: A simple groundwater scheme in the TRIP river routing model: global off-line evaluation against GRACE terrestrial water storage estimates and observed river discharges, 2012.

Vergnes, J.-P., Decharme, B., Alkama, R., Martin, E., Habets, F., and Douville, H.: A Simple Groundwater Scheme for Hydrological and Climate Applications: Description and Offline Evaluation over France, J. Hydrometeorol., 13, 1149–1171, https://doi.org/10.1175/JHM-D-11-0149.1, http://journals.ametsoc.org/doi/10.1175/JHM-D-11-0149.1, 2012.

Wahr, J., Swenson, S., and Velicogna, I.: Accuracy of GRACE mass estimates, Geophys. Res. Lett., 33, L06 401, https://doi.org/10.1029/2005GL025305, http://doi.wiley.com/10.1029/2005GL025305, 2006.

Wang, F., Polcher, J., Peylin, P., and Bastrikov, V.: Assimilation of river discharge in a land surface model to improve estimates of the continental water cycles, Hydrol. Earth Syst. Sci., 22, 3863–3882, https://doi.org/10.5194/hess-22-3863-2018, 2018.

Watkins, M. M. and Yuan, D.-N.: GRACE Gravity Recovery and Climate Experiment JPL Level-2 Processing Standards Document For Level-2 Product Release 05.1, 2014.

Watkins, M. M., Wiese, D. N., Yuan, D.-N., Boening, C., and Landerer, F. W.: Improved methods for observing Earth's time variable mass distribution with GRACE using spherical cap mascons, J. Geophys. Res. Solid Earth, 120, 2648–2671, https://doi.org/10.1002/2014JB011547, http://doi.wiley.com/10.1002/2014JB011547, 2015a.

Watkins, M. M., Wiese, D. N., Yuan, D.-N., Boening, C., and Landerer, F. W.: Improved methods for observing Earth's time variable mass distribution with GRACE using spherical cap mascons, J. Geophys. Res. Solid Earth, 120, 2648–2671, https://doi.org/10.1002/2014JB011547, http://doi.wiley.com/10.1002/2014JB011547, 2015b.

Webster, P. J., Jian, J., Hopson, T. M., Hoyos, C. D., Agudelo, P. A., Chang, H.-R., Curry, J. A., Grossman, R. L., Palmer, T. N., Subbiah, A. R., Webster, P. J., Jian, J., Hopson, T. M., Hoyos, C. D., Agudelo, P. A., Chang, H.-R., Curry, J. A., Grossman, R. L., Palmer, T. N., and Subbiah, A. R.: Extended-Range Probabilistic Forecasts of Ganges and Brahmaputra Floods in Bangladesh, Bull. Am. Meteorol. Soc., 91, 1493–1514, https://doi.org/10.1175/2010BAMS2911.1, http://journals.ametsoc.org/doi/10.1175/2010BAMS2911.1, 2010.

Wijngaard, R. R., Biemans, H., Lutz, A. F., Shrestha, A. B., Wester, P., and Immerzeel, W. W.: Climate change vs. socio-economic development: understanding the future South Asian water gap, Hydrol. Earth Syst. Sci., 22, 6297–6321, https://doi.org/10.5194/hess-22-6297-2018, https://www.hydrol-earth-syst-sci.net/22/6297/2018/, 2018.

Yang, P., Xia, J., Zhan, C., and Wang, T.: Reconstruction of terrestrial water storage anomalies in Northwest China during 1948–2002 using GRACE and GLDAS products, Hydrol. Res., 49, 1594–1607, https://doi.org/10.2166/nh.2018.074, 2018.

Yilmaz, M. T., DelSole, T., and Houser, P. R.: Improving Land Data Assimilation Performance with a Water Budget Constraint, J. Hydrometeorol., 12, 1040–1055, https://doi.org/10.1175/2011JHM1346.1, 2011.

Zhang, Y., Pan, M., Sheffield, J., Siemann, A., Fisher, C., Liang, M., Beck, H., Wanders, N., MacCracken, R., Houser, P. R., Zhou, T., Lettenmaier, D. P., Ma, Y., Pinker, R. T., Bytheway, J., Kummerow, C. D., and Wood, E. F.: A Climate Data Record (CDR) for the global terrestrial water budget: 1984-2010, Hydrol. Earth Syst. Sci. Discuss., pp. 1–40, https://doi.org/10.5194/hess-2017-192, https://www.hydrol-earth-syst-sci-discuss.net/hess-2017-192/, 2017.

Zhang Yongqiang et al.: Monthly global observation-driven Penman-Monteith-Leuning (PML) evapotranspiration and components. v2. CSIRO. Data Collection., https://doi.org/10.4225/08/5719A5C48DB85, 2016.

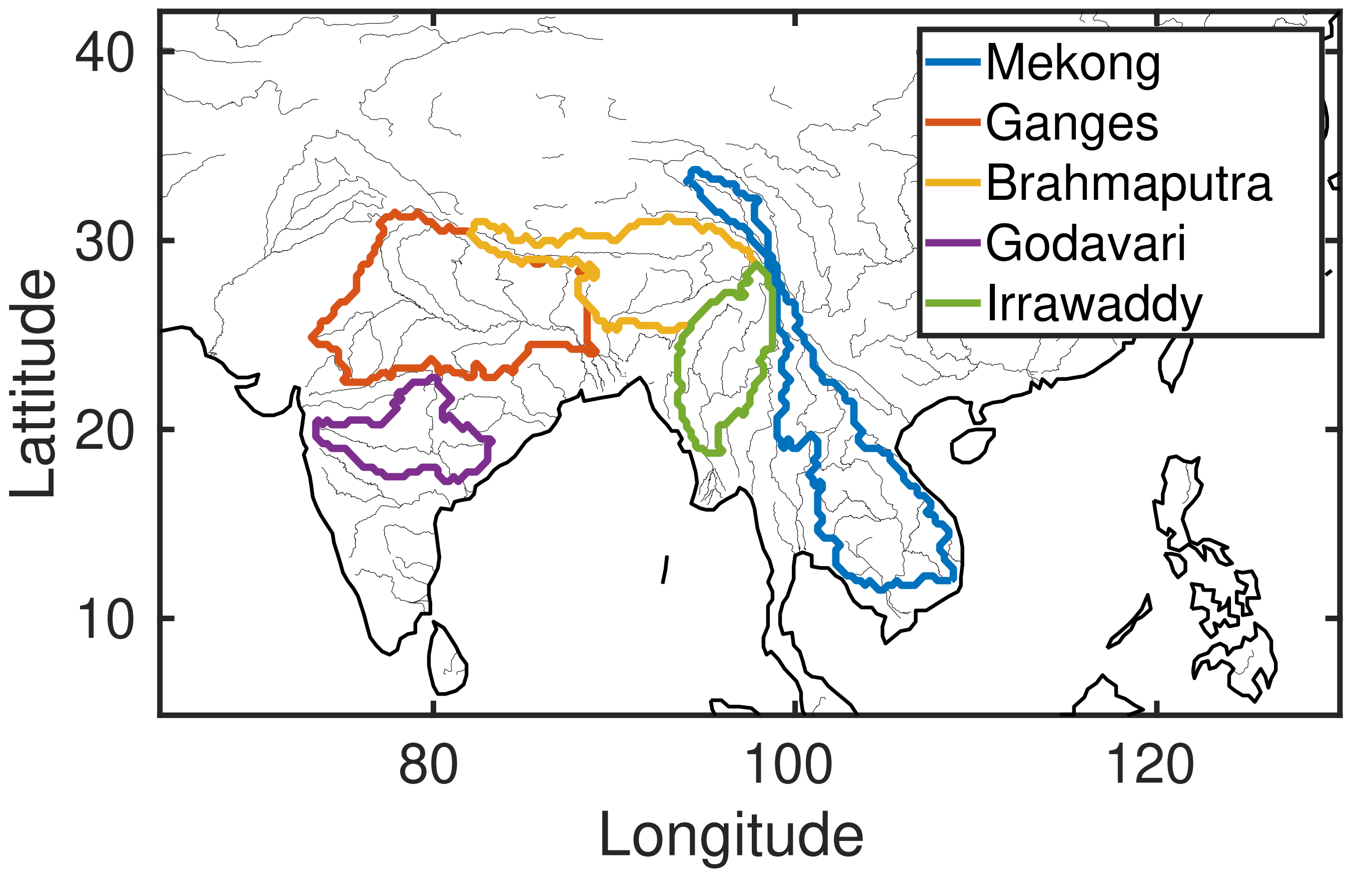

Figure 1. Five Himalayan basins considered in this study.

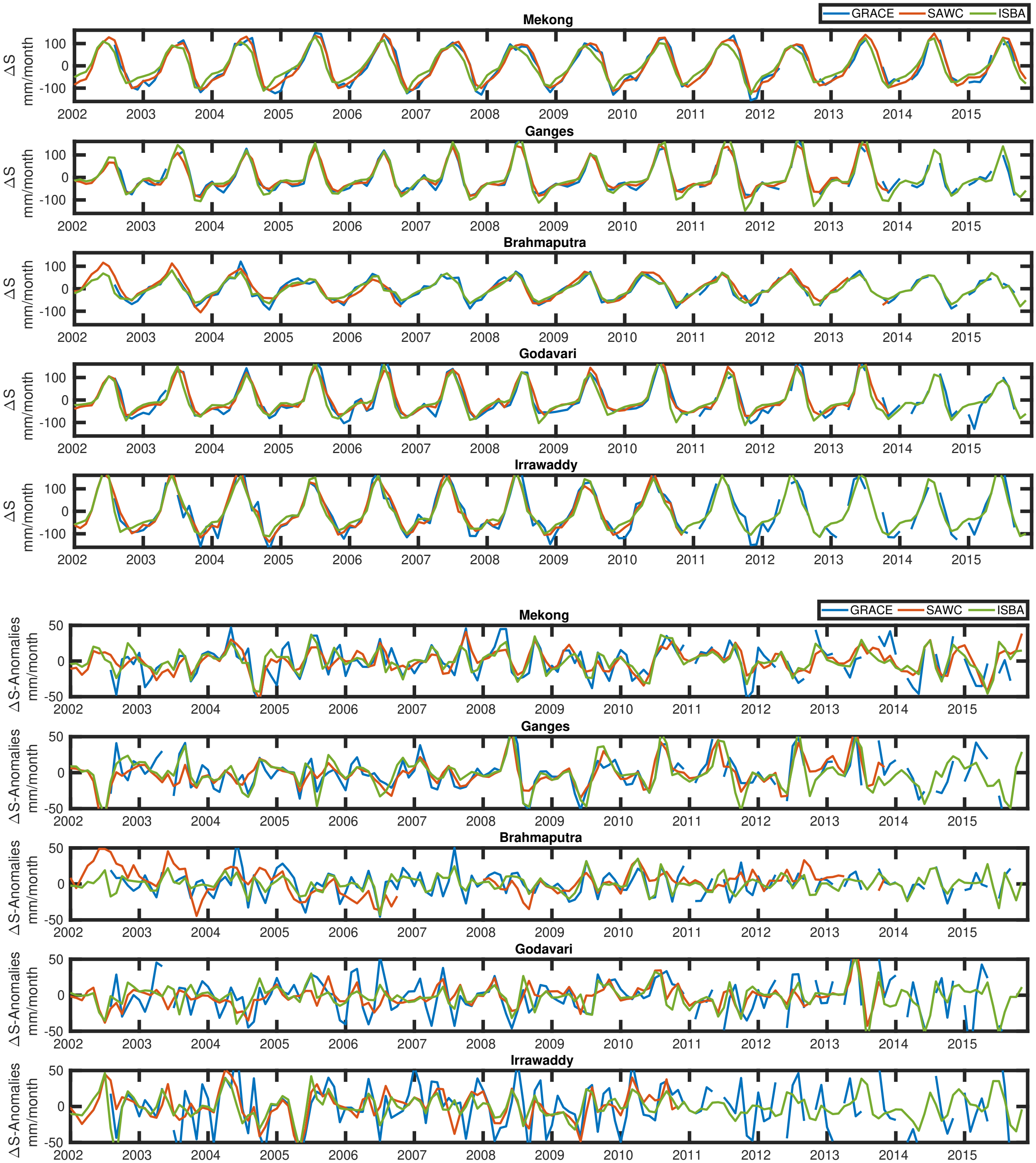

Figure 2. TWSC (top) and TWSC seasonal anomaly (bottom), for the three estimates (GRACE, SAWC and ISBA), for the five Himalayan basins, over the GRACE time period.

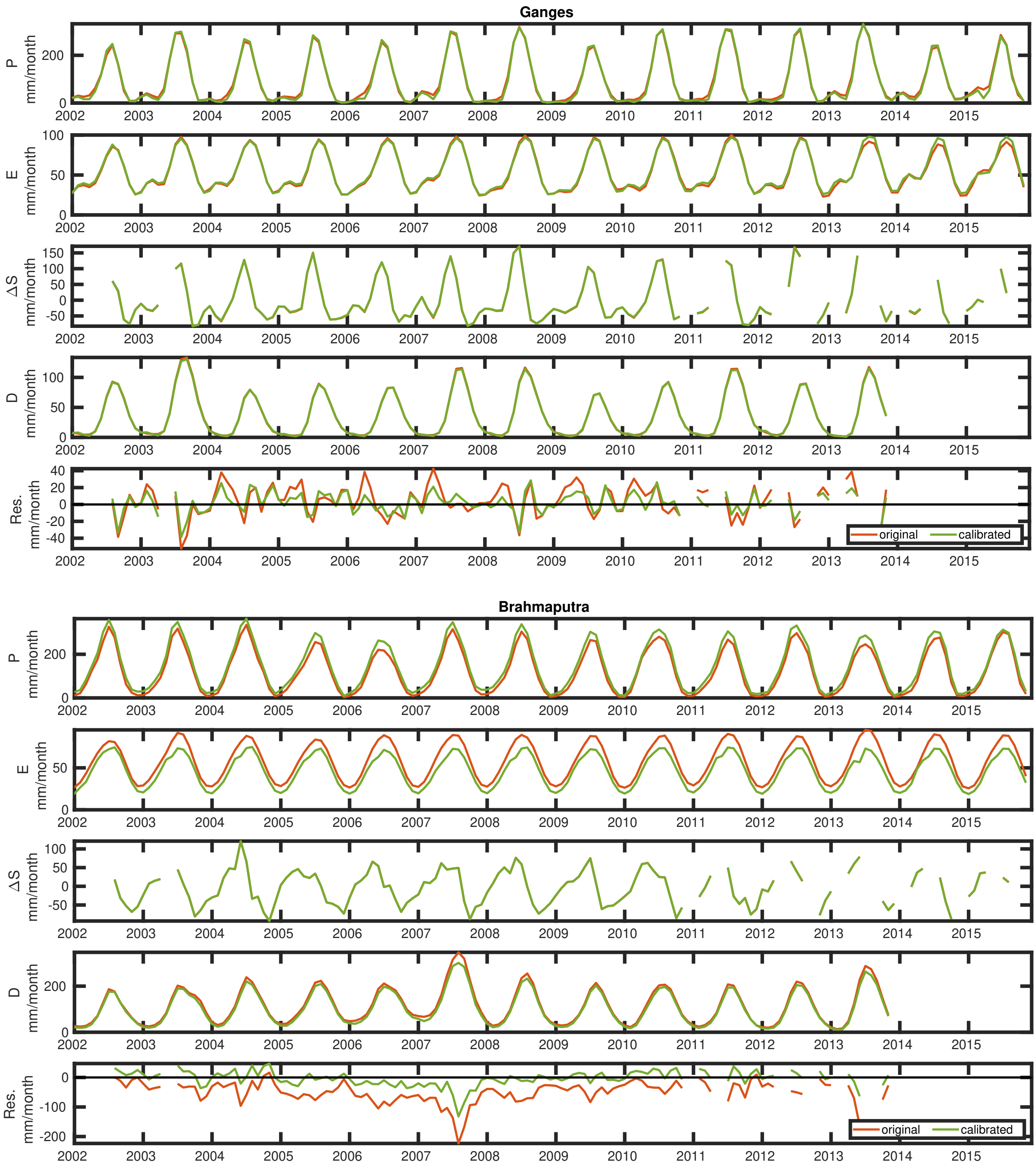

Figure 3. Comparison of the four water components estimates and the WC budget residuals (in row), for the original X_{SW} (red) and calibrated X_{CAL} (green) estimates. The estimates are for the Ganges (top) and the Brahmaputra (bottom) basins.

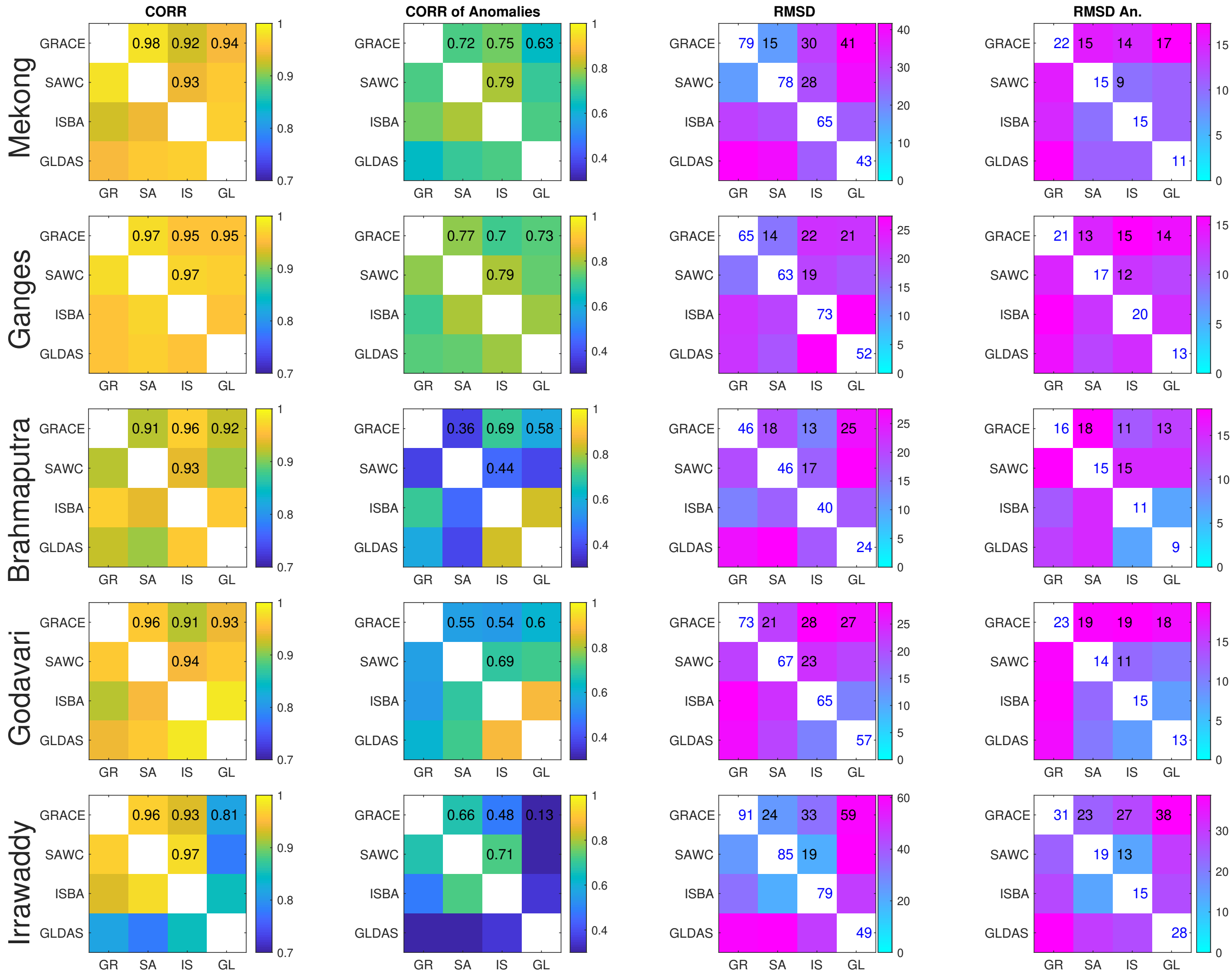

Figure 4. The correlation (left), the correlation of the anomalies (middle), and the Root Mean Square of the Difference (right) between the four estimates (GRACE, SAWC, ISBA, GLDAS), for the five Himalayan river basins (in row), over the GRACE (2002-2015). Some of the commented statistics are also indicated numerically.

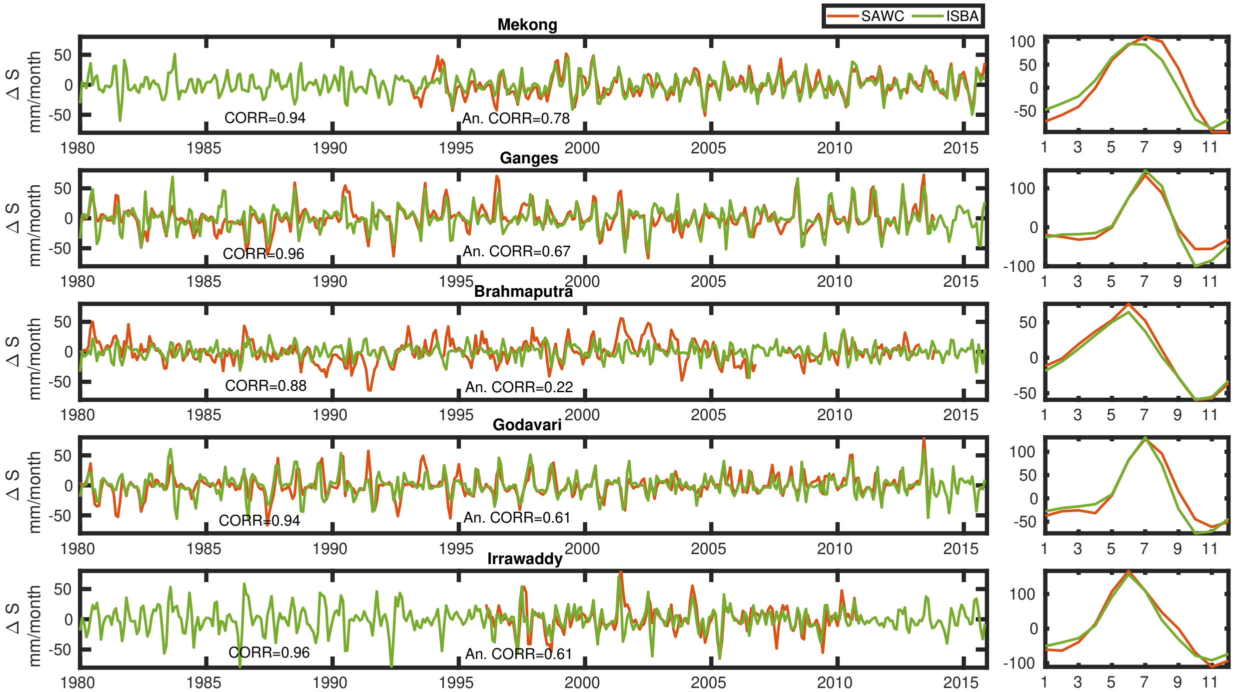

Figure 5. Times series of the TWSC (left) and seasonal cycle (right) from 1980 to 2015, for the SAWC and ISBA model estimates, over the five considered basins. Correlations of raw and anomaly values are also provided.

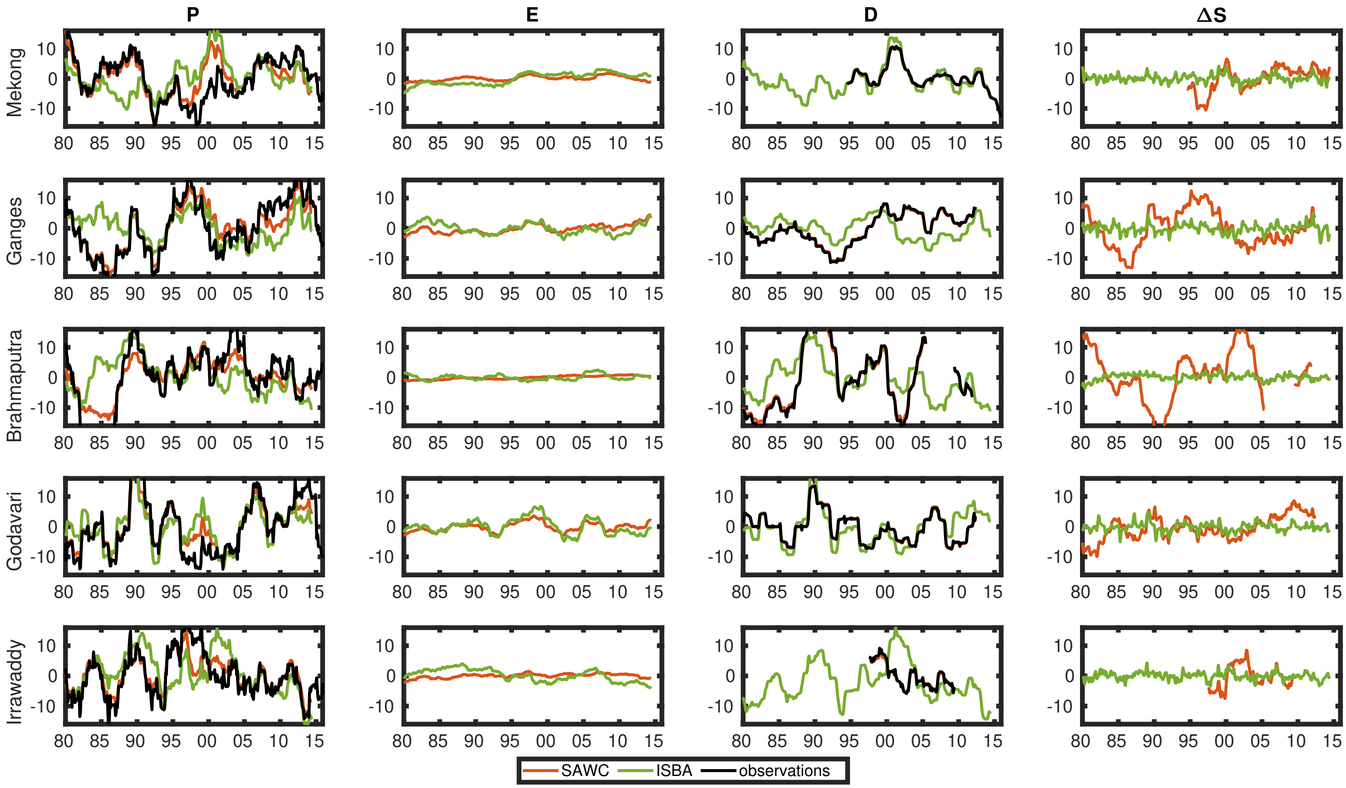

Figure 6. Times series of the WC components (mm/month) for 1980-2015, in terms of anomalies (with respect to the climatological season) smoothed using a 3-year moving window; on the five considered basins; for SAWC (red) and ISBA (green) estimates. Observations are also represented in black (MWEWP for P and *in situ* measurements for D).

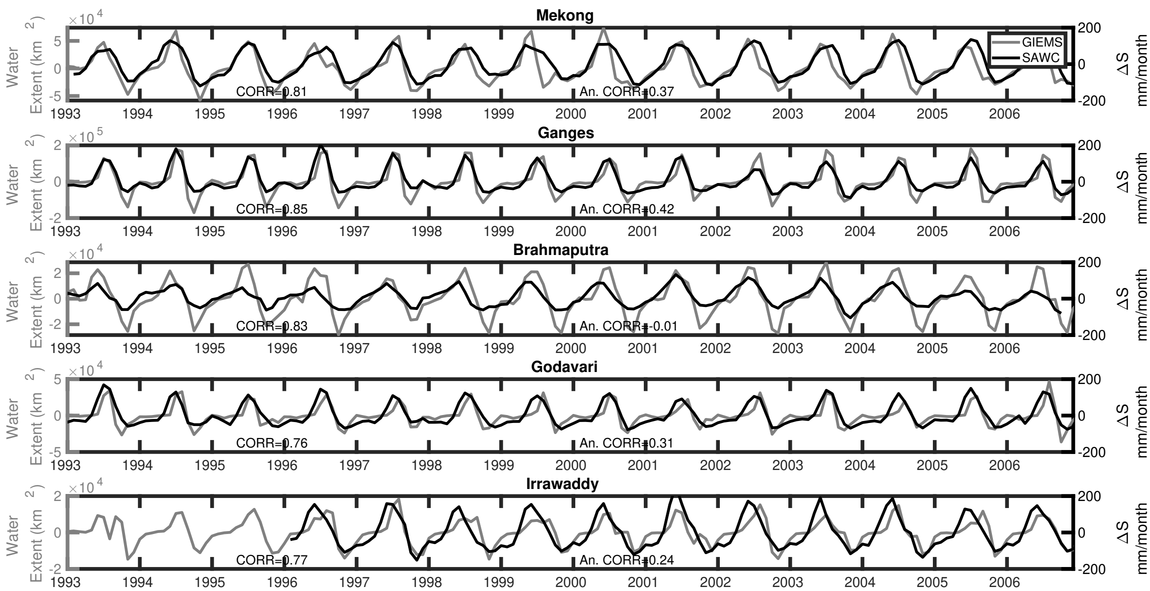

Figure 7. Evaluation of TWSC from the SAWC reconstruction using the GIEMS inundated area, from 1993 to 2007, over the figure considered basins. The correlation between them is also provided, together with the correlation of the anomalies.

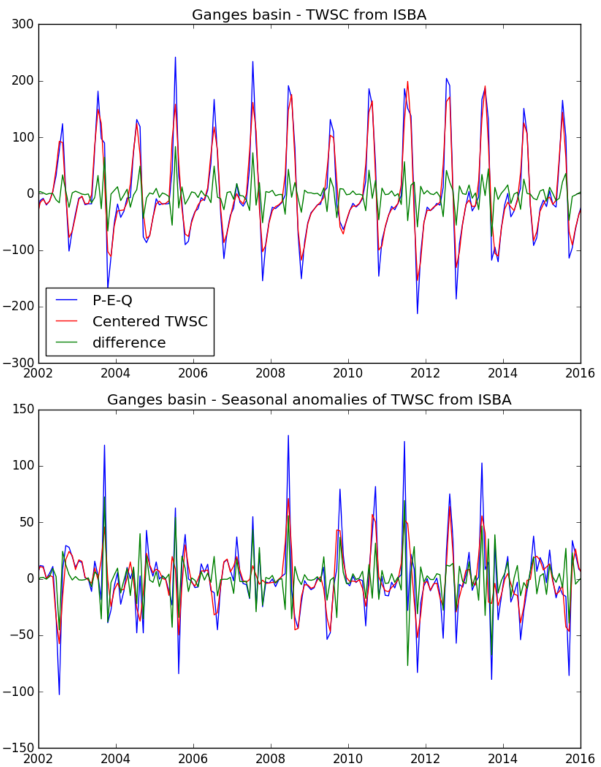

Figure A1. Comparison between the two estimates of the TWSC over the Ganges basin: the closure at any time step in the ISBA model (P-E-Q, in blue) and the centered difference of the observed TWS anomalies with GRACE (in red). The difference of the two estimate is also shown (in green). The figure shows the original time series (top) and the seasonal anomalies (bottom).

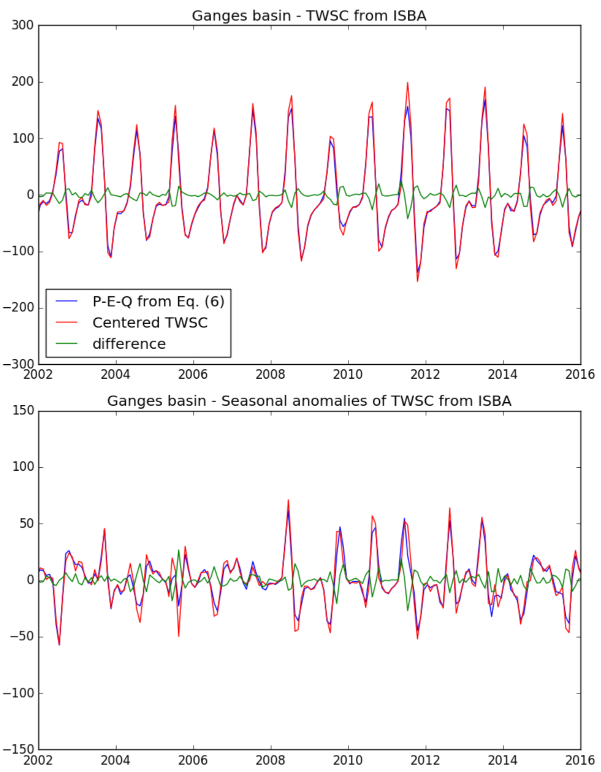

Figure A2. Same as Fig.A1 but the closure is now ensured with Eq.(8), following Landerer et al. (2010). This approximation lead to a better match of the two estimates.

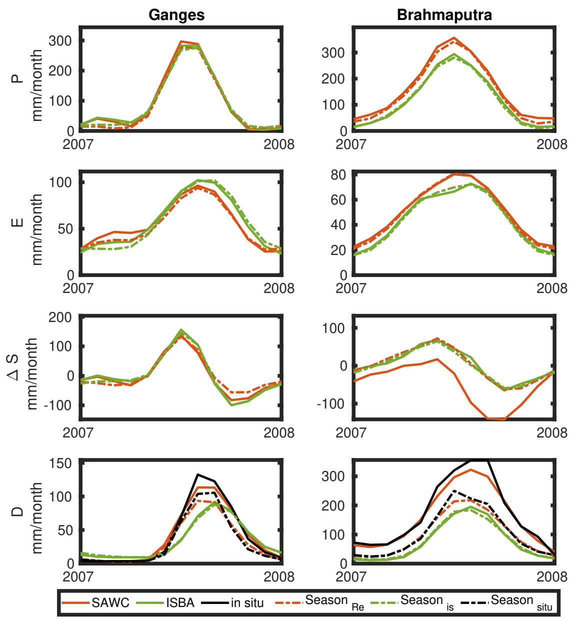

Figure A3. Comparison between SAWC reconstruction (red) and the ISBA model output (Green) estimates; for the year 2007 with a large inondation in the Brahmaputra basin; for the four water components estimates (in row); and the Ganges (left) and the Brahmaputra (right) basins. The climatological season are also represented in dashed lines.

Name	Area (10^5 km^2)	Outlet station	Location	Time record	Source
Mekong	7.7	Phonm Pehn	11.5°N; 104.9°E	1993-2016	pers. comm. (1)
Ganges	9.6	Hardinge	24.1°N; 89.0°E	1980-2013	pers. comm. (2)
Brahmaputra	5.2	Bahadurabad	25.1°N; 89.7°E	1980-2013	pers. comm. (2)
Godavari	3.2	Polavaram	17.2°N; 81.7°E	1965-2015	Water Resources Information System of India
Irrawaddy	3.6	Pyay	18.8°N; 95.2°E	1996-2010	GRDC

Table 1. Characteristics of the five considered basins and associate *in situ* measurement stations. (1) Personnal communication from Biancamaria et al., 2017, EGU. Data derived from radar altimetry water level estimations and calibrated against *in situ* measurements following a similar technique as in (Papa et al., 2010) (2) personnal communication and obtained from BDWB (Bangladesh Water Development Board (http://www.bwdb.gov.bd/) as in Papa et al. (2012)

.

Basin	Original X_{SW} (mm/month)	Calibrated X_{CAL} (mm/month)	Improvement (%)
Mekong	23	16	30
Ganges	20	15	25
Brahmaputra	61	28	54
Godavari	29	20	31
Irrawaddy	44	24	45

Table 2. RMS of the WC budget residuals in mm/month for the original (X_{SW}) and the calibrated (X_{CAL}) estimates, over the five considered basins.

	Dataset	Basins				
		Mekong	Gang.	Brahm.	Goda.	Irrawa.
P	GPCP	14	15	35	12	10
	TMPA	14	13	31	13	11
	MSEWP	15	17	32	16	19
E	GLEAM	6	5	6	6	6
	MOD16	4	8	6	9	8
	CSIRO	8	5	7	5	8
D	*In situ*	3.4	2.7	4.7	2.8	7.7

Table 3. Uncertainty estimates (in mm/month) in terms of STD error compared to X_{PF}, for P, E and D, over the five considered basins, for the original datasets.

Long-term Total Water Storage Change from a SAtellite Water Cycle (SAWC) reconstruction over large south Asian basins

Reviewer 1

GENERAL COMMENTS

• **In summary, this article tries to reconstruct Total Water Storage Change (TWSC) using satellite-based integrated water cycle components of P and E, and observation-based D in five larger basins in South Asia. Then, TWSC obtained from GRACE, ISBA, and GLDAS were used to evaluate the performance of the reconstructed TWSC here. The topic is interesting, but many attempts have already been made by previous studies (Tang et al., 2017; Humphrey et al., 2017). Major revision is needed before publication, here, a few suggestions that authors should consider while revising are listed below:**
- Thank you for your comments, we hope that the new version of the manuscript is now in a better shape.

The introduction describes better the state of the art in the reconstruction of TWS anomalies (TWSA) and change (TWSC) in the literature and states the novelty of this article : *If classical approaches to retrieve TWS rely on land surface model (Decharme et al., 2019; Tootchi et al., 2018), studies have recently attempted to using statistical model and various climate drivers. For instance, Humphrey et al. (2017) reconstruct the TWSA using a linear regression from precipitation and temperature, while Chen et al. (2019) use an artificial neural network to reconstruct TWSA based on precipitation, temperature and surface variables (e.g. soil moisture and NDVI). Yang et al. (2018) review and compare several statistical methods (linear, random forest, artificial neural network and support vector machine) to reconstruct TWSA from soil moisture, canopy water and snow water equivalent. These studies focus on the TWSA without monitoring the whole WC.*

If statistical methods offer the opportunity to estimate TWS anomalies at global scale in a simpler way than the LSM, they do not consider the water balance and the related TWS estimate may not be coherent with the other water components. The water balance at basin scale has been used to estimate TWSC using satellite observations. For instance, Tang et al. (2017) use the Budyko model to estimate annual TWSC based on P and E. A more sophisticated method has been developed where satellite observation were assimilated in the Variable Capicity Model (VIC) at basin scale (Pan et al., 2012) and at the 0.5° LSM pixel (Zhang et al., 2017). In order to obtain a TWSC estimate independent of the LSM, another framework has been developed (Aires, 2014; Munier et al., 2014; Pellet et al., 2019). It is based on an integration of satellite observations and in situ river discharge measurement, using the conservation of the water as a constraint, to optimize all sources of information. It has been shown that using the constraint on the observations gives as good results as with the assimilation framework (Munier et al., 2014). We follow here on this framework to reconstruct TWSC with good accuracy.

• **First, as shown in Fig. 2 and Fig. 4, SAtellite Water Cycle (SAWC) estimates generally has higher correlation with that from GRACE, ISBA and GLDAS except for the Irrawaddy Basin. As for the correlation of anomalies, relative higher correlation between SAWC estimates and the other three were found in Mekong and Ganges basins, while much lower correlation was found in the left basins especially in Brahmaputra basin. This highlighted the spatial and temporal variation of the performance of SAWC estimates, therefore, more discussions on such uncertainties are needed.**

- Thank you for this remark. Indeed, the performance of all TWSC estimates varies spatially and numerous factors explain these differences mainly the accuracy of GRACE estimate in terms of seasonal anomalies over the small sub-basins and the large uncertainty of the evapotranspiration estimates used in SAWC. This is now clearer in the text: *SAWC estimate has generally high correlation values with GRACE, ISBA and GLDAS estimates (except for the Irrawaddy Basin). In terms of correlation of anomalies, SAWC estimate is always closer to ISBA than to GRACE even if SAWC has high correlation of anomalies with GRACE (between 0.69 and 0.79) except over the Brahmaputra basin (0.36). Comparatively, GLDAS estimate is less correlated to GRACE over the four basins (except Brahmaputra basin). The RMSD and RMSD of anomalies show similar pattern than the correlation* *values over all the basins. GRACE presents relatively low spatial resolution (300 km^2 at the equator) that can decrease the accuracy of TWSC anomaly estimate for small basins (e.g. Godavari, Irrawaddy). The smaller the basin is, the larger the gap between SAWC-GRACE and SAWC-ISBA correlation becomes. SAWC estimate is based on precipitation and evapotranspiration obtained at finner spatial resolution than GRACE (0.25°). Therefore, SAWC, as ISBA (at the 0.25° spatial resolution) better represents the anomaly over small basins as far as the precipitation and evapotranspiration are accurate. Over the Brahmaputra basin, the large uncertainty of satellite evapotranspiration products over the mountainous area (see the impact of the calibration for the evapotranspiration estimate over this basin in Figure 3) might impact the SAWC TWSC accuracy and explains why GLDAS and ISBA are better over this basin. This assumption is later confirmed in Figure 6 in which precipitation in ISBA and SAWC are close but the anomalies of E differ. Finally, the discrepancy between simulated TWSC from ISBA and GLDAS can be explained by the different representation of aquifers in these two models. While a two-dimensional diffusive groundwater scheme in ISBA represents unconfined aquifer processes (Vergnes and Decharme, 2012; Vergnes et al., 2012), the Noah land model used in the GLDAS simulations did not include surface and groundwater storage. Therefore, the simulated mean seasonal cycle and the inter-annual variability of the TWSC is improved in ISBA (Decharme et al., 2019). On the contrary, deviations from GRACE TWSC can thus be expected with GLDAS (Syed et al., 2008).*

Based on the results presented in Figure 4, we decided to compare our SAWC solution over the long time period only to ISBA. Nevertheless, none of these models included anthropogenic effects and this is now also discussed (see next comment).

• Second, ISBA was used to evaluate the SAWC estimates due to the long series historical data. However, ISBA model does not represent anthropogenic factors such as groundwater extraction, river regulation or irrigation, which may significantly impact D and TWSC. This is much different from SAWC estimates which might already considered the anthropogenic disturbances. This difference can lead to some big discrepancy as shown in Fig.5 and Fig. 6 for the terms of D and delta S. Therefore, the authors are encouraged to clarify which anthropogenic disturbances have been considered in SAWC estimates and how they affect the discrepancy among different basins in corresponding years.

- Thank you for this remark. With the use of actual river discharge observations, SAWC estimate considers all anthropogenic effects that impact the river along its path (mainly water withdrawal for irrigation and flow regulation by dams). Several points have been added to the comments on the results in order to discuss the impact of dams construction on the Mekong river discharge in the observation and in the model:

- *In Figure 6, Mekong river discharge anomalies show lower min-max range in the observations than in ISBA. Li et al. (2017) highlight the impact of the construction of the Xiaowan and the Nuozhadu dams starting in 1991. The dam reduces the streamflow in particularly wet seasons and increases the streamflow in particularly dry seasons which lowers the anomaly variations.*

- *D is more correlated to precipitation in ISBA (0.94) than in SAWC solutions (0.63). This shows that D in a model is more straightforwardly dependent of the precipitation than in observe state.*

- *On the contrary, TWSC anomaly is less linked to precipitation in the ISBA model than in SAWC solutions where natural recharge is better represent. These difference is also discussed in the Appendix A.*

• Also, if possible, adding the results from other hydrological models that considers the human activity is highly encouraged.
- Significant efforts that have been made during the last two decades to incorporate anthropogenic impacts in LSM (Hanasaki et al., 2006; Haddeland et al., 2014). These new schemes in LSM are developed offline and mainly at regional scale but significant challenges still remain in their standardization into global LSM as in the availability of the observations (e.g. irrigation, pumping rate) (Pokhrel et al., 2016). Global LSM do not include the global representation of flow regulation and irrigation water needs. Therefore, analyzing the impact of anthropogenic effect into a LSM is beyond the scope of the study. The previous citations and comment have been added to the manuscript in Section 2.2.2.: *"These two global and well known models have been chosen for comparison even if none of them included anthropogenic effects on the river discharge and groundwater storage. Significant efforts have been made during the last two decades to incorporate anthropogenic impacts in LSM (Hanasaki et al., 2006; Haddeland et al., 2014) but crucial challenges still remain. Most of these new schemes in LSMs have been developed and used offline for regional scale studies and without common and standardized framework (Pokhrel et al., 2016; Döll et al., 2016). At global*

scale, a state of art does not include the global representation of flow regulation and irrigation water needs."

• **Third, compared with previous studies of TWSC derive, an integrated utilization of satellite products to retrieve TWSC seems an advantage of this study, but similar idea has been reported in (Pan et al., 2012; Zhang et al., 2017). , therefore, clear illustration of the novelty of this study is needed.**

- As now stated in the introduction, the Princeton (Pan et al., 2012; Sahoo et al., 2011; Zhang et al., 2017) and the WATCHFULL / WACMOS-MED initiatives (Aires, 2014; Munier et al., 2014; Pellet et al., 2019) are both based on the combination of numerous satellite information and the physical law of water conservation to optimize the latter. However, the first is based on the assimilation of the satellite information into the VIC model while our approach attempted to be as observational as possible. A study has already compared TWSC reconstruction between Princeton and WATCHFULL initiative over the Mississippi (Munier et al., 2014). This is now indicated in the introduction.

• **Fourth, as shown in B of Eq.4, a priori specification of the uncertainties seems important in obtaining optimized solution through Post-Filtering", so more explanations of the advantage for current specification scheme is needed.**

- Characterizing the uncertainties of satellite-retrieved products is a difficult task. These specifications are now clearer in the text: *"Such characterizations are generally product and site specific. Some studies (Pan et al., 2012; Sahoo et al., 2011; Zhang et al., 2017) estimate the a priori uncertainty of particular water components based on the spread among the various estimates (taking the spread of estimates as an estimate of the uncertainties can sometimes be dangerous). In our case, this approach would not take into account the fact that the precipitation estimates are not independent. The value used here are derived from (Munier et al., 2014) in which the authors reviewed carefully the literature on this topic. The partitioning of uncertainty between P and E has however been modified to allow larger uncertainty in P since datasets are dependent in our case. As the objective of the current study is to reconstruct GRACE TWSC, the approach assumes lower errors in GRACE estimate that becomes our reference."*

SPECIFIC COMMENTS

• **Page 5, Line 60-65, this study used one gravity solution based on MASCON-JPL. Other solutions from Center for Space Research (CSR) at the University of Texas at Austin, and GeoForschungsZentrum (GFZ) are available. The comparison of different solutions among different basins are needed to be clarified to support the choice of the solution or using the resembled solution.**

- I may misunderstand the comment. To our knowledge, GFZ does not provide a GRACE MASCON solution but only Spherical decomposition one. The MASCON solutions from CSR and JPL differ in their processing and we choose here the JPL solution because it is more independent of the spherical solutions. This information has been added to the manuscript at Section 2.2.1: *"Another MASCON solution exists : the CSR-MASCON solution. The MASCON solutions from CSR and JPL differ in their processing: while JPL solution is based on the explicit estimation of mass anomalies at specific mass concentration block location using the analytical partial derivatives of the inter-satellite range-rate measurements (Watkins et al., 2015), the CSR developed MASCON solution is first based on a Spherical decomposition of the inter-satellite range-rate measurements that is truncated spatially at the location of mass concentration (Save et al., 2016). The two solutions have been compared to the spherical solutions in terms of uncertainty in both min-max range and trend in (Scanlon et al., 2016; Save et al., 2016). We choose here the JPL solution because it is more independent of the spherical solution."*. If it is admitted that for a Spherical solution (JPL,CSR,GFZ), the use of the ensemble mean (simple arithmetic mean of JPL, CSR, GFZ fields) is the most effective in reducing the noise in the gravity field solutions (Sakumura et al., 2014) this might not be the case for MASCON solution. The community uses independently the JPL-MSC or CSR-MSC solutions (Scanlon et al., 2016). In our preliminary test, the JPL-MSC overperforms the spherical solution over particular latitudinal oriented river basin (e.g. Irrawaddy).

• **Page5, Line 70, "with respect to averaged season", the time period should be specified.**

- The time period used to computed the average season (2002-2015) is now specified.

• **Caption of fig. 3, for the original XSW (blue), it should be green as shown in the figure.**

- Thank you. Caption has been modified.

References

Aires, F. (2014). Combining Datasets of Satellite-Retrieved Products. Part I: Methodology and Water Budget Closure. *J. Hydrometeorol.*, 15(4):1677–1691.

Chen, H., Zhang, W., Nie, N., and Guo, Y. (2019). Long-term groundwater storage variations estimated in the Songhua River Basin by using GRACE products, land surface models, and in-situ observations. *Sci. Total Environ.*, 649:372–387.

Decharme, B., Delire, C., Minvielle, M., Colin, J., Vergnes, J., Alias, A., SaintMartin, D., Séférian, R., Sénési, S., and Voldoire, A. (2019). Recent Changes in the ISBACTRIP Land Surface System for Use in the CNRMCM6 Climate Model and in Global OffLine Hydrological Applications. *J. Adv. Model. Earth Syst.*, 11(5):1207–1252.

Döll, P., Douville, H., Güntner, A., Müller Schmied, H., and Wada, Y. (2016). Modelling Freshwater Resources at the Global Scale: Challenges and Prospects.

Haddeland, I., Heinke, J., Biemans, H., Eisner, S., Flörke, M., Hanasaki, N., Konzmann, M., Ludwig, F., Masaki, Y., Schewe, J., Stacke, T., Tessler, Z. D., Wada, Y., and Wisser, D. (2014). Global water resources affected by human interventions and climate change. *Proc. Natl. Acad. Sci. U. S. A.*, 111(9):3251–3256.

Hanasaki, N., Kanae, S., and Oki, T. (2006). A reservoir operation scheme for global river routing models. *J. Hydrol.*, 327(1-2):22–41.

Humphrey, V., Gudmundsson, L., and Seneviratne, S. I. (2017). A global reconstruction of climatedriven subdecadal water storage variability. *Geophys. Res. Lett.*, 44(5):2300–2309.

Li, D., Long, D., Zhao, J., Lu, H., and Hong, Y. (2017). Observed changes in flow regimes in the Mekong River basin. *J. Hydrol.*, 551:217–232.

Munier, S., Aires, F., Schlaffer, S., Prigent, C., Papa, F., Maisongrande, P., and Pan, M. (2014). Combining data sets of satellite-retrieved products for basin-scale water balance study: 2. Evaluation on the Mississippi Basin and closure correction model. *J. Geophys. Res.*, 119(21):12,100–12,116.

Pan, M., Sahoo, A. K., Troy, T. J., Vinukollu, R. K., Sheffield, J., Wood, and F, E. (2012). Multisource estimation of long-term terrestrial water budget for major global river basins. *J. Clim.*, 25(9):3191–3206.

Pellet, V., Aires, F., Munier, S., Fernández Prieto, D., Jordá, G., Arnoud Dorigo, W., Polcher, J., and Brocca, L. (2019). Integrating multiple satellite observations into a coherent dataset to monitor the full water cycle - Application to the Mediterranean region. *Hydrol. Earth Syst. Sci.*, 23(1).

Pokhrel, Y. N., Hanasaki, N., Wada, Y., and 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(4):548–574.

Sahoo, A. K., Pan, M., Troy, T. J., Vinukollu, R. K., Sheffield, J., and Wood, E. F. (2011). Reconciling the global terrestrial water budget using satellite remote sensing. *Remote Sens. Environ.*, 115(8):1850–1865.

Sakumura, C., Bettadpur, S., and Bruinsma, S. (2014). Ensemble prediction and intercomparison analysis of GRACE time-variable gravity field models. *Geophys. Res. Lett.*, 41(5):1389–1397.

Save, H., Bettadpur, S., and Tapley, B. D. (2016). High-resolution CSR GRACE RL05 mascons. *J. Geophys. Res. Solid Earth*, 121(10):7547–7569.

Scanlon, B. R., Zhang, Z., Save, H., Wiese, D. N., Landerer, F. W., Long, D., Longuevergne, L., and Chen, J. (2016). Global evaluation of new GRACE mascon products for hydrologic applications. *Water Resour. Res.*, 52(12):9412–9429.

Syed, T. H., Famiglietti, J. S., Rodell, M., Chen, J., and Wilson, C. R. (2008). Analysis of terrestrial water storage changes from GRACE and GLDAS. *Water Resour. Res.*, 44(2).

Tang, Y., Hooshyar, M., Zhu, T., Ringler, C., Sun, A. Y., Long, D., and Wang, D. (2017). Reconstructing annual groundwater storage changes in a large-scale irrigation region using GRACE data and Budyko model. *J. Hydrol.*, 551:397–406.

Tootchi, A., Jost, A., and Ducharne, A. (2018). Multi-source global wetland mapping: combining surface water imagery and groundwater constraints. *Hydrol. Earth Syst. Sci. Discuss.*, pages 1–34.

Vergnes, J.-P. and Decharme, B. (2012). A simple groundwater scheme in the TRIP river routing model: global off-line evaluation against GRACE terrestrial water storage estimates and observed river discharges.

Vergnes, J.-P., Decharme, B., Alkama, R., Martin, E., Habets, F., and Douville, H. (2012). A Simple Groundwater Scheme for Hydrological and Climate Applications: Description and Offline Evaluation over France. *J. Hydrometeorol.*, 13(4):1149–1171.

Watkins, M. M., Wiese, D. N., Yuan, D.-N., Boening, C., and 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(4):2648–2671.

Yang, P., Xia, J., Zhan, C., and Wang, T. (2018). Reconstruction of terrestrial water storage anomalies in Northwest China during 19482002 using GRACE and GLDAS products. *Hydrol. Res.*, 49(5):1594–1607.

Zhang, Y., Pan, M., Sheffield, J., Siemann, A., Fisher, C., Liang, M., Beck, H., Wanders, N., MacCracken, R., Houser, P. R., Zhou, T., Lettenmaier, D. P., Ma, Y., Pinker, R. T., Bytheway, J., Kummerow, C. D., and Wood, E. F. (2017). A Climate Data Record (CDR) for the global terrestrial water budget: 1984-2010. *Hydrol. Earth Syst. Sci. Discuss.*, (April):1–40.

Long-term Total Water Storage Change from a SAtellite Water Cycle (SAWC) reconstruction over large south Asian basins

Reviewer 2

COMMENTS

- **This paper explains how to estimate the total water storage change of a large basin using GRACE estimates by satellite. The water conservation equation is used to have an independent constraint, and uses satellite estimates of precipitation and evaporation together with a direct measure of river discharge near the mouth of the river. These complementary measures have to be at the monthly scale, as this is the temporal resolution of the GRACE estimates. The methodology is applied to four large basins in India and Indochina and the methodology is able to produce estimates that compare well with GRACE observations. I find the paper and the methodology interesting and the results of application, since they allow to monitor the water status of large basins with very little in-situ observations (essentially only a discharge measurement is needed). The paper is clearly written and well organized.**
- Thank you for your appreciation on this work.

- **My questions, being a meteorologist, are about the determination of the precipitation and evaporation by satellite. In the integration part, three sources are used for precipitation. More than providing the references, nothing is said about the characteristics of these data sets, how are they produced, what are the differences between them, which is the uncertainty for each of them, and how is the total uncertainty obtained. Similarly, more information about the ET databases should be provided**.

- Thank you for this remark. These global precipitation and ET database are widely used in remote sensing community (Rodell et al., 2015; Pan et al., 2012; Sahoo et al., 2011; Zhang et al., 2017; Munier et al., 2014; Pellet et al., 2018, 2019), but more information are needed in this section. The manuscript specifies some characteristics of these datasets for P, E and TWSC, and cite some inter-comparison studies in which uncertainty assessments can be found:

1. *Precipitation, P - All these datasets are global datasets widely used in the community. GPCP and TMPA use the same algorithm Threshold Matched Precipitation Index (TMPI) to estimate instantaneous precipitation from multiple satellites by combining high-quality passive micro-wave observations and infrared data and differ only in the use of gauge analyses (GPCC) to obtain calibrated estimates. While TMPA is based on inverse random-error variance weighting, GPCP assumes that the precipitation distribution estimated from combined satellite estimates is optimal and uses the gauge observations only for debiasing. The MSWEP dataset merges the highest quality precipitation data sources available as a function of timescale and location. It uses a combination of rain gauge measurements, the two previous satellite datasets, and reanalysis. These datasets have been compared in terms of uncertainties and performance in (Sun et al., 2018). It should be noted that these datasets are not independent of each other but represent the best up-to-date precipitation estimates for hydrological studies.*

2. *Evapotranspiration, E - The Global Land Evaporation Amsterdam Model (GLEAM-V3B, Martens et al., 2016; Miralles et al., 2011), uses Priestley and Taylor 1972 (Priestley and Taylor, 1972) empirical energy-based equation to calculate the reference evapotranspiration and separately estimate the different components of land evaporation: transpiration, bare-soil evaporation, interception loss, open-water evaporation and sublimation. GLEAM uses reanalysis (vA) or satellite (vB) precipitation inputs. The global observation-driven Penman-Monteith-Leuning (PML, Zhang Yongqiang et al., 2016) evapotranspiration introduced by the Commonwealth Scientific and Industrial Research Organisation (CSIRO) and the MODIS Global Evapotranspiration Project (MOD16, Mu et al., 2011) are both evapotranspiration estimates based on Penman-Monteith equations (PENMAN, 1948; Monteith, 1965). We choose these three datasets due to their different equations of parametrization for the evapotranspiration. Inter-comparison of global evapotranspiration algorithms and datasets can be found in (Michel et al., 2016).*

3. *Total Water Storage Change (TWSC), ΔS - Another MASCON solution exists : the CSR-MASCON solution. The MASCON solutions from CSR and JPL differ in their processing: while JPL solution is based on the explicit estimation of mass anomalies at specific mass concentration block location using the analytical partial derivatives of the inter-satellite range-rate measurements (Watkins et al., 2015), the CSR developed MASCON solution is first based on a Spherical decomposition of the inter-satellite range-rate measurements that is truncated spatially at the location of mass concentration (Save et al., 2016). The two solutions have been compared to the spherical solutions in terms of uncertainty in both min-max range and trend in (Scanlon et al., 2016; Save et al., 2016). We choose here the JPL solution because it is more independent of the spherical solution.*

Characterizing the uncertainties of satellite-retrieved products is a difficult task (see the answer to the next comment). In this study, the precipitation and evapotranspiration uncertainties are derived from the literature (Munier et al., 2014). All datasets describing a water components have the same uncertainty and the resulting uncertainty of ensemble mean is derive assuming the independence of the sources. This simplification is usually done: (Pan et al., 2012; Munier et al., 2014; Pellet et al., 2018, 2019). This is now clearer in the manuscript where the Simple Weighting (i.e. arithmetic average) estimate and its uncertainty are introduced in Section 2.2.1.

$$P_{SW} = \frac{1}{p-1}\sum_{i=1}^{p}\frac{\sum_{k\neq i}(\sigma_k)^2}{\sum_k(\sigma_k)^2}P_i. \tag{1}$$

This equation is valid when there is no bias error in the P_is (thanks to the preliminary bias correction) and is optimal when the errors ϵ_i are statistically independent from each other. This expression is valid for the other water components. The variance of the P_{SW} uncertainties is then given by:

$$\sigma_{P_{SW}} = \frac{1}{(p-1)^2}\sum_{i=1}^{p}\left(\frac{\sum_{k\neq i}(\sigma_k)^2}{\sum_k(\sigma_k)^2}\right)^2\sigma_i^2. \tag{2}$$

• I believe that the paper would benefit of related precipitation and evaporation maps and a discussion in depth of the uncertainties of the terms of water closure budget (P, ET, D). The last paragraph of subsection 2.2.1, or Table 3, only give the values imposed for the uncertainties, not how they are obtained. Also subsection 2.3.4 is vague on the subject.

- Thank you for this comment. Uncertainty analysis at grid scale is beyond the scope of this study which focus on the basin scale, however particular analysis on precipitation (resp. evapotranspiration) uncertainties can be found in (Sun et al., 2018) (resp. (Michel et al., 2016)). These references have been added in section 2.2.1. The following specification are now clearer in the text : *"Characterizing the uncertainties of satellite-retrieved products is a difficult task. Such characterizations are generally product, and site, specific. Some studies (Pan et al., 2012; Sahoo et al., 2011; Zhang et al., 2017) estimate the a priori uncertainty of particular water component based on the spread among the various estimates. In our case, this approach would not take into account the fact that our precipitation estimate are not independent. Finally, the values considered here are derived from (Munier et al., 2014) in which the authors reviewed the literature on this topic. Compared to this study, the partitioning of uncertainty between P and E has been modified to allow larger uncertainty in P since all P datasets are dependent with each others. As the objective of the current study is to reconstruct GRACE TWSC, the approach assumes low error in the GRACE estimate."*

Noted that Table 3 does not give *a priori* uncertainty value but provided uncertainty estimates computed *a posteriori* as the distance between the original datasets and the reference (our new estimate). This is why the various original products have different *a posteriori* uncertainty even if a same *a priori* uncertainty was specified. This is now clearer in the text.

- **On the other hand, ISBA-CTRIP and GLDAS are used as evaluation tools. Being these utilities models themselves, it is unclear if the results are good enough for validation in this area of the world. More details should be provided about the quality of these models in this region so that it appears legitimate to use it as a validation tool, discussing at least their uncertainties.**

- Thank you for this remark. We prefer to state "evaluation" with ISBA instead of "validation" in section 3.2 since validating TWSC over the long time period (1980-2015) is a difficult task. Often, this type of evaluation is performed by comparing to other independent estimates. For instance, Figure 4 shows that ISBA can simulate more accurately the TWSC than GLDAS. The following statement has been aded to the manuscript : *"Finally, the discrepancy between simulated TWSC from ISBA and GLDAS can be explained by the representation of aquifers in these two models. While a two-dimensional diffusive groundwater scheme in ISBA represents unconfined aquifers process (Vergnes and Decharme, 2012; Vergnes et al., 2012),* *the Noah land model used in the GLDAS simulations did not include surface and groundwater storage. Therefore, the simulated mean seasonal cycle and the inter-annual variability of the TWSC is improved in ISBA (Decharme et al., 2019). On the contrary, deviations from GRACE TWSC can thus be expected with GLDAS simulated TWSC (Syed et al., 2008). Based on the results presented in Figure 4, we decided to compare SAWC estimate only to ISBA over long time period".* Nevertheless, none of these models included anthropogenic effect and this is also now discussed.

• **Furthermore, having a better description of the rationale in Section 3, more specifically in subsection 3.2, may be of help for the reader. In subsection 3.1 all the available sources (GRACE, SAWC, ISBA and GLDAS) are compared and it is stated that SAWC fits best with GRACE, admitting that it is by construction. Then, in subsection 3.2, it retains ISBA for the further comparison considering that it performs better than GLDAS. In this part a discussion on the uncertainties of all methods is missing.**
- See previous answer for the rational in Section 3. The manuscript now justifies the ISBA comparison.

• **For a non-specialist, the paper is interesting and the methodology seems powerful**
- Thank you, we hope that the revised manuscript will answer your concerns.

References

Decharme, B., Delire, C., Minvielle, M., Colin, J., Vergnes, J., Alias, A., SaintMartin, D., Séférian, R., Sénési, S., and Voldoire, A. (2019). Recent Changes in the ISBACTRIP Land Surface System for Use in the CNRMCM6 Climate Model and in Global OffLine Hydrological Applications. *J. Adv. Model. Earth Syst.*, 11(5):1207–1252.

Michel, D., Jiménez, C., Miralles, D. G., Jung, M., Hirschi, M., Ershadi, A., Martens, B., Mccabe, M. F., Fisher, J. B., Mu, Q., Seneviratne, S. I., Wood, E. F., and Fernández-Prieto, D. (2016). The WACMOS-ET project - Part 1: Tower-scale evaluation of four remote-sensing-based evapotranspiration algorithms. *Hydrol. Earth Syst. Sci.*, 20(2):803–822.

Monteith, J. L. (1965). Evaporation and environment.

Munier, S., Aires, F., Schlaffer, S., Prigent, C., Papa, F., Maisongrande, P., and Pan, M. (2014). Combining data sets of satellite-retrieved products for basin-scale water balance study: 2. Evaluation on the Mississippi Basin and closure correction model. *J. Geophys. Res.*, 119(21):12,100–12,116.

Pan, M., Sahoo, A. K., Troy, T. J., Vinukollu, R. K., Sheffield, J., Wood, and F, E. (2012). Multisource estimation of long-term terrestrial water budget for major global river basins. *J. Clim.*, 25(9):3191–3206.

Pellet, V., Aires, F., Mariotti, A., and Fernández-Prieto, D. (2018). Analyzing the Mediterranean Water Cycle Via Satellite Data Integration. *Pure Appl. Geophys.*, 175(11).

Pellet, V., Aires, F., Munier, S., Fernández Prieto, D., Jordá, G., Arnoud Dorigo, W., Polcher, J., and Brocca, L. (2019). Integrating multiple satellite observations into a coherent dataset to monitor the full water cycle - Application to the Mediterranean region. *Hydrol. Earth Syst. Sci.*, 23(1).

PENMAN, H. L. (1948). Natural evaporation from open water, hare soil and grass. *Proc. R. Soc. Lond. A. Math. Phys. Sci.*, 193(1032):120–145.

Priestley, C. and Taylor, R. (1972). On the Assessment of Surface Heat Flux and Evaporation Using Large-Scale Parameters. *Mon. Weather Rev.*, 100(2):81–92.

Rodell, M., Beaudoing, H., L'Ecuyer, T., Olson, W., Famiglietti, J., Houser, P., Adler, R., Bosilovich, M., Clayson, C., Chambers, D., Clark, E., Fetzer, E., Gao, X., Gu, G., Hilburn, K., Huffman, G., Lettenmaier, D., Liu, W., Robertson, F., Schlosser, C., Sheffield, J., and Wood, E. (2015). The Observed State of the Water Cycle in the Early 21st Century. *J. Clim.*, pages 8289–8318.

Sahoo, A. K., Pan, M., Troy, T. J., Vinukollu, R. K., Sheffield, J., and Wood, E. F. (2011). Reconciling the global terrestrial water budget using satellite remote sensing. *Remote Sens. Environ.*, 115(8):1850–1865.

Save, H., Bettadpur, S., and Tapley, B. D. (2016). High-resolution CSR GRACE RL05 mascons. *J. Geophys. Res. Solid Earth*, 121(10):7547–7569.

Scanlon, B. R., Zhang, Z., Save, H., Wiese, D. N., Landerer, F. W., Long, D., Longuevergne, L., and Chen, J. (2016). Global evaluation of new

GRACE mascon products for hydrologic applications. *Water Resour. Res.*, 52(12):9412–9429.

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., and Hsu, K. (2018). A Review of Global Precipitation Data Sets: Data Sources, Estimation, and Intercomparisons. *Rev. Geophys.*, 56(1):79–107.

Syed, T. H., Famiglietti, J. S., Rodell, M., Chen, J., and Wilson, C. R. (2008). Analysis of terrestrial water storage changes from GRACE and GLDAS. *Water Resour. Res.*, 44(2).

Vergnes, J.-P. and Decharme, B. (2012). A simple groundwater scheme in the TRIP river routing model: global off-line evaluation against GRACE terrestrial water storage estimates and observed river discharges.

Vergnes, J.-P., Decharme, B., Alkama, R., Martin, E., Habets, F., and Douville, H. (2012). A Simple Groundwater Scheme for Hydrological and Climate Applications: Description and Offline Evaluation over France. *J. Hydrometeorol.*, 13(4):1149–1171.

Watkins, M. M., Wiese, D. N., Yuan, D.-N., Boening, C., and 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(4):2648–2671.

Zhang, Y., Pan, M., Sheffield, J., Siemann, A., Fisher, C., Liang, M., Beck, H., Wanders, N., MacCracken, R., Houser, P. R., Zhou, T., Lettenmaier, D. P., Ma, Y., Pinker, R. T., Bytheway, J., Kummerow, C. D., and Wood, E. F. (2017). A Climate Data Record (CDR) for the global terrestrial water budget: 1984-2010. *Hydrol. Earth Syst. Sci. Discuss.*, (April):1–40.