DOI 10.1007/978-3-319-32449-4_4
Reprinted from *Surveys in Geophysics* Journal, DOI 10.1007/s10712-015-9354-y

On Creating Global Gridded Terrestrial Water Budget Estimates from Satellite Remote Sensing

Yu Zhang[1] · **Ming Pan**[1] · **Eric F. Wood**[1]

Received: 16 March 2015 / Accepted: 12 December 2015 / Published online: 4 January 2016
© Springer Science+Business Media Dordrecht 2015

Abstract The increasing availability and reliability of satellite remote sensing products [e.g., precipitation (*P*), evapotranspiration (*ET*), and the total water storage change (*TWSC*)] make it feasible to estimate the global terrestrial water budget at fine spatial resolution. In this study, we start from a reference water budget dataset that combines all available data sources, including satellite remote sensing, land surface model (LSM) and reanalysis, and investigate the roles of different non-satellite remote sensing products in closing the terrestrial water budget through a sensitivity analysis by removing/replacing one or more categories of products during the budget estimation. We also study the differences made by various satellite products for the same budget variable. We find that the gradual removal of non-satellite data sources will generally worsen the closure errors in the budget estimates, and remote sensing retrievals of *P*, *ET*, and *TWSC* together with runoff (*R*) from LSM give the worst closure errors. The gauge-corrected satellite precipitation helps to improve the budget closure (4.2–9 % non-closure errors of annual mean precipitation) against using the non-gauge-corrected precipitation (7.6–10.4 % non-closure errors). At last, a data assimilation technique, the constrained Kalman filter, is applied to enforce the water balance, and it is found that the satellite remote sensing products, though with worst closure, yield comparable budget estimates in the constrained system to the reference data. Overall, this study provides a first comparison between the water budget closure using the satellite remote sensing products and a full combination of remote sensing, LSM, and reanalysis products on a quasi-global basis. This study showcases the capability and potential of the satellite remote sensing in closing the terrestrial water budget at fine spatial resolution if properly constrained.

Keywords Global terrestrial water budget · Satellite remote sensing · CKF · Constraint

✉ Yu Zhang
yz5@Princeton.EDU

[1] Department of Civil and Environmental Engineering, Princeton University, E-208 E-Quad, Princeton, NJ 08544, USA

1 Introduction

The evolution and shift of the terrestrial water cycle pose a significant impact on the climate system, the availability of water resources, the occurrence of hydrological extremes, etc. The global terrestrial water cycle is also key in understanding the complex interactive feedbacks and mechanisms among the land surface, ocean and the atmosphere. Better understanding of the global water cycle can be enabled by accurate and reliable estimation of the global terrestrial water budget that is continuous in time and space from various data sources such as traditional in situ observations, advanced satellite remote sensing, land surface model (LSM), and reanalysis. Though in situ observations always serve as the "truth", their limited spatial fetch and high cost make them less economical. By contrast, satellite remote sensing, with its mission in observing Earth at a fine spatial resolution with temporal continuity, makes it possible to estimate the water budget in less developed regions where the in situ gauge stations are sparse or non-existent. In addition, satellite remote sensing products are also always used as the forcing (e.g., precipitation) or the basic setups (e.g., land cover, topography) for land surface, weather, and climate models, which can also provide water budget estimates such as evapotranspiration and runoff at large scales that supplement the point scale in situ measurement. However, the accuracy and reliability of the budget estimates from these models highly depend on factors like the parameterization, initial condition, forcing, and calibration/validation.

Terrestrial water budget consists of four major components: the precipitation (P), evapotranspiration (ET), runoff (R), and total water storage change ($TWSC$), and the mass balance of water requires that:

$$TWSC = P - ET - R \tag{1}$$

With the development and improvement in satellite remote sensing techniques, all the components of the terrestrial water budget can be estimated from the space-borne remote sensing, though the accuracy and resolution vary across the different water budget components due to varying sensor characteristics. For example, precipitation can be estimated by merging microwave and infrared information, such as with the Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA: Huffman et al. 2007, 2010), the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks-Cloud Classification System (PERSIANN-CCS: Hong et al. 2007), and the Climate Prediction Center morphing method (CMORPH: Joyce et al. 2004). Global estimations of evapotranspiration can be derived from satellite surface radiation budget, surface meteorology, and vegetation cover (e.g., Fisher et al. 2008; Mu et al. 2007; Vinukollu et al. 2011). Potentially runoff can be retrieved from satellite altimetry. The Surface Water Ocean Topography (SWOT: Durand et al. 2010) mission, which is expected to be launched in 2020, will play a leading role in surface hydrological observations by providing information for major rivers and water bodies at near global coverage with a repeating period of 21 days. SWOT will use the Ka-band radar and provide sea surface height and terrestrial water heights at 120-km-wide swath. The radar measurements will also be processed for measuring rivers with widths larger than 100 meters width and lakes with areas larger than $(250\ \text{m})^2$. SWOT will provide river elevation (with an accuracy of 10 cm), slope (with an accuracy of 1 cm/1 km) and width which can be used in estimating river discharge (Paiva et al. 2015; Pavelsky et al. 2014). The surface and subsurface total water storage (TWS) can be measured by the NASA Gravity Recovery And Climate Experiment (GRACE) twin satellites (Landerer and Swenson 2012; Tapley et al. 2004;

Wahr et al. 2004), which were launched on March 17, 2002, at the coarse spatial resolution of ~220 km and monthly timescale. Then the measurements of micro-gravity at their original resolution are processed onto 1° spatial resolution and monthly timescale by three centers, Geoforschungs Zentrum Potsdam (GFZ), Center for Space Research at University of Texas, Austin (CSR), and Jet Propulsion Laboratory (JPL), with the gravity anomalies attributed to changes in total water storage. GRACE has been widely used in water budget estimation (e.g., Gao et al. 2010; Pan et al. 2012; Sahoo et al. 2011; Sheffield et al. 2009; Wang et al. 2014) as well as drought analysis (e.g., Famiglietti 2014; Thomas et al. 2014). As a successor to the original GRACE mission, a GRACE Follow-on (GRACE-FO, http://grace.jpl.nasa.gov/mission/grace-fo/) is planned for launch in 2017 to continue measuring the Total Water Storage (*TWS*).

At regional scales where in situ observations are available, Pan et al. (2012) estimated the errors in each water budget component against the in situ observations and then merged those products based on their error information. In addition, earlier studies (Gao et al. 2010; Sahoo et al. 2011; Sheffield et al. 2009; Troy et al. 2011; Vinukollu et al. 2011) attempted to close the water budget at the basin scale by using satellite remote sensing. However, the closure cannot be achieved without enforcing the water balance through approaches such as data assimilation. Sahoo et al. (2011) applied a constrained Kalman filter (CKF) to close the water budget using satellite remote sensing and provided balance-constrained best estimates of the water budget for ten major basins.

At the global scale, previous studies (e.g., Dirmeyer et al. 2006; Haddeland et al. 2011; Oki et al. 1995; Trenberth et al. 2007; Weedon et al. 2011) estimated the terrestrial water budget from either single or multiple land surface and/or hydrologic models for limited periods mostly during the 1990s. A recent study (Rodell et al. 2015) blended multiple sourced datasets into a "best guess" by utilizing the standard deviation/spread of the uncertainties for each component and estimated the global terrestrial water budget for 2000–2010. But none of these studies provides a multi-decadal global terrestrial water budget record covering most of the satellite era. Currently the authors are carrying out an additional analysis for estimating and closing the global terrestrial water budget that combine multiple data sources (Table 1) that include in situ observations, satellite remote sensing, and LSM outputs and reanalysis, at 0.5° spatial resolution and monthly timescale for the period of 1984–2010. A subset of these data records (2004–2007) is used in this paper as the reference data to evaluate the water budget estimation using different combinations of remote sensing data sources.

Though current satellite remote sensing offers the potential to estimate gridded terrestrial water budget over the globe, which is especially of significant importance for ungauged and sparsely gauged regions, challenges exist in quantifying the errors in each satellite remote sensing products when reliable in situ observations are lacking—and even when they are available. This study aims at creating a global terrestrial gridded (0.5°) monthly water budget from satellite remote sensing that has budget closure by applying a CKF algorithm. The period of 2004–2007 is selected based on the common availability of all the datasets listed in Table 1. This paper is organized as follows: Sect. 2 introduces the datasets, the methodology in estimating and closing water budget, and the experiments used to assess the impact of the data sources on the water budget; Sect. 3 presents the results from using data products that range from all available products to remotely sensed products along on estimating and closing water budget; and Sect. 4 presents the conclusions and findings from the study.

Table 1 Data summary

Dataset	Period	Spatial resolution	Temporal resolution	References
Precipitation				
CSU	1998–2010	0.25°	3 h	(Bytheway and Kummerow 2013)
PGF	1948–2010	0.25°	3 h	(Sheffield et al. 2006)
CHIRPS	1981–present	0.5°	Monthly	(Funk et al. 2014)
GPCC(v6)	1901–2010	0.5°	Monthly	(Schneider et al. 2014)
TMPA-RT	2001–present	0.25°	3 h	(Huffman et al. 2007, 2010)
Evapotranspiration				
SRB–PGF–PM	1984–2007	0.5°	3 h	(Vinukollu et al. 2011)
VIC	1948–2010	0.25°	3 h	(Sheffield and Wood 2007)
ERA-interim	1979–present	T255	–	(Simmons et al. 2006)
MERRA	1979–present	2/3 × 1/2 H grids	–	(Rienecker et al. 2011)
GLEAM	1984–2007	0.5°	Daily	(Miralles et al. 2011)
SRB-CFSR-SEBS	1984–2007	0.5°	Daily	(Vinukollu et al. 2011)
SRB-CFSR-PM				
SRB-CFSR-PT				
Runoff				
VIC	1948–2010	0.25°	3 h	(Sheffield and Wood 2007)
Total water storage (TWS)				
VIC	1948–2010	0.25°	3 h	(Sheffield and Wood 2007)
GRACE	2002–present	1°	Monthly	(Landerer and Swenson 2012)

The Princeton Global Forcing (PGF) dataset used in this study, which is an updated version of the PGF described in (Sheffield et al. 2006), provides near-surface meteorological data for driving land surface models and other terrestrial modeling systems. All other acronyms are defined Sect. 2.1

Forced by the near-surface meteorological variables from PGF listed above, VIC model simulates evapotranspiration and runoff at 0.25°, 3 h over the land from 1948 to 2010 as an updated version of (Sheffield and Wood 2007)

2 Data and Methodology

2.1 Utilized Data

Three satellite precipitation products used in this study. Two are the Colorado State University (CSU) and the Real Time product of TRMM Multi-satellite Precipitation Analysis (TMPA-RT). The CSU product (Bytheway and Kummerow 2013) is the TMPA (Huffman et al. 2007, 2010) accumulated rainfall with uncertainty estimates and is a 3-h, 0.25° spatial resolution gridded product available between 50°N and 50°S. Different from the standard TMPA products, the TMPA-RT product provides the precipitation estimation at near real time over 50°N– 50°S, but without rain gauge adjustment. The difference between TMPA-RT and CSU reveals the role of correcting the satellite retrievals using ground gauges. The third satellite product is from the Climate Hazard group InfraRed Precipitation with Stations (CHIRPS, Funk et al. 2014) product. In addition, the in situ based Global Precipitation Climate Center (GPCC) product (Schneider et al. 2014) and the multi-source merged product of Princeton Global Forcing (PGF, Sheffield et al. 2006)

dataset that is based on in situ, satellite and reanalysis model precipitation are also used in our study.

Five satellite remote sensing products are used for global evapotranspiration (*ET*) estimation. The products are various combinations of data sources and *ET* algorithms (radiation–surface meteorology–*ET* algorithm). The algorithms are run at a daily time step and then aggregated to monthly totals. The products are: (1) SRB–PGF–PM: Surface Radiation Budget–Princeton Global Forcing–Penman–Monteith; (2) SRB–CFSR–PM: Surface Radiation Budget–Climate Forecast System Reanalysis–Penman–Monteith; (3) SRB–CFSR–PT: Surface Radiation Budget–Climate Forecast System Reanalysis–Priestly-Taylor; (4) SRB–CFSR–SEBS: (the Surface Radiation Budget–Climate Forecast System Reanalysis–Surface Energy Balance System (SEBS); and (5) GLEAM: the Global Land-surface Evaporation: the Amsterdam Methodology, (Miralles et al. 2011). The algorithms for models (1)–(4) are described in Vinukollu et al. (2011). Three *ET* models (2)–(4) use CFSR meteorology and CFSR surface radiation that has been adjusted to match the monthly SRB surface radiation. Additionally, two reanalysis *ET* products from ERA-interim (Simmons et al. 2006) and NASA's Modern-Era Retrospective Analysis for Research and Application (MERRA, Rienecker et al. 2011), as well as one *ET* from the Variable Infiltration Capacity (VIC) land surface model (LSM), forced by an updated version of PGF (Sheffield and Wood 2007), are also used.

The runoff is generated at 0.25° spatial resolution and 3 h, from the same VIC LSM (Sheffield and Wood 2007). For the runoff fields used in here, VIC was calibrated over 43 globally well-distributed river basins.

The terrestrial Total Water Storage Change (*TWSC*) is estimated from both the VIC LSM and GRACE (Landerer and Swenson 2012) ReLease05 (RL05) (http://grace.jpl.nasa.gov/data/get-data/monthly-mass-grids-land/). Since most of the soil water dynamics occur in the upper portion of the column, it was decided to use both a LSM and GRACE retrievals as *TWSC* estimates even though it is recognized that GRACE should be a more inclusive product. The GRACE gravity anomaly retrievals are post-processed onto 1° spatial resolution and monthly timescale by GFZ, CSR, and JPL, as discussed earlier. The provided scaling grid was then multiplied to the 1° GRACE land data in order to reduce the attenuation of the surface mass variations at small spatial scales due to the sampling and post-processing (e.g., the de-stripping filter and the 300-km Gaussian Filter). It is noted that the current 1° GRACE land data cannot accurately observe ice mass changes over Greenland and Antarctica, or glacier and ice caps (Jacob et al. 2012). The ensemble mean (equal weighting of JPL, CSR, GFZ products) is calculated as the *TWSC* from GRACE. Sakumura et al. (2014) points out that this is the most appropriate method in reducing the noise in the gravity field solutions within the available scatter of the solutions.

All the products in Table 1 are either aggregated or disaggregated onto 0.5° spatial resolution and the monthly timescale. The reader is referred to the references in Table 1 for detailed information for each product. In addition, monthly streamflow observations from Global Runoff Data Center (GRDC) are also used in this study, as well as for validation.

2.2 Product Merging and Water Budget Closure

There are insufficient in situ observations, especially for *ET*, *R*, and *TWSC*, to estimate the errors and/or biases for each water budget component at the grid scale over the globe. Therefore, in this study, the deviations from the ensemble mean of all the data sources for the same budget variable is used as a proxy of the uncertainty/error in the individual products. The merging procedure for each budget component is a weighted averaging,

where the optimal merging weight w_i is given by the following equation (Luo et al. 2007; Sahoo et al. 2011):

$$w_i = \frac{1}{\sigma_i^2} \Big/ \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \tag{2}$$

in which w_i is the merging weight for product i, σ_i^2 is the error variance of product i calculated against the ensemble mean, and n is the total number of the products. Note that $\sum w_i$ equals to 1. The larger the error variance of product i, the lower is its merging weight. The merging process is conducted in the unconstrained system.

In the constrained system, the CKF algorithm (Sahoo et al. 2011) is applied that assures budget closure at each grid cell over the globe. In short, CKF redistributes the non-closure errors back onto the different water budget components according to their error levels and correlations. The water balance residual is defined as $r = P - ET - R - TWSC$. The budget components can be written as the column vector x, $x = [P, ET, R, TWSC]^T$, and then the residual of the water balance can be expressed as a linear function of the vector, $r = G\,x$, where $G = [1, -1, -1, -1]$. The error covariance matrix of x is calculated as $\varepsilon_{xx} = \overline{(\hat{x} - x)(\hat{x} - x)^T}$, where $\hat{x}$ is an estimate of x, its "true value" and the overbar represents expectation over the time series. In this study, $(\hat{x} - x)$ is replaced with the spread of the ensemble in each water budget component. This uncertainty estimation method was first proposed by Adler et al. (2001) and then applied in Tian and Peters-Lidard (2010) to generate a global precipitation uncertainty map for a variety of satellite remote sensing products. ε_{xx} has the dimension 4×4 since x consists of four budget variables. Then the balance-constrained estimate is calculated via $\hat{x}' = \hat{x} - \varepsilon_{xx} G^T (G \varepsilon_{xx} G^T)^{-1} \hat{r}$. The residual term $\hat{r}$ is redistributed back into different water budget components through the above equation. Mathematically, the CKF algorithm mimics assimilating a zero-error observation of r (i.e., $r = 0$).

2.3 Design of the Budget Closure Experiments

Five experiments listed in Table 2 are carried out in terms of a sensitivity analysis to understand how datasets from different sources affect the estimation and closure of the terrestrial water budget. Starting with the complete suite of measurements and products [experiment (a)], referred to here as the reference dataset, and removing from the reference data, one at a time, in situ observations [experiment (b)], reanalysis products [experiment (c)], in situ and reanalysis products together [experiment (d)], it is possible to determine the impact of including various data sources in the global water budget estimates and its closure. The five experiments are conducted in both unconstrained (that is without applying the CKF algorithm) and the constrained (with CKF) systems.

3 Results and Discussion

3.1 Roles of Non-satellite Sources in Closing the Water Budget

To assess the contribution from non-satellite data sources to the non-closure/imbalance, an error sensitivity analysis, in terms of experiment (a–e), is conducted by removing/replacing datasets with ground observations, or reanalysis, one category at a time, and then all

Table 2 Data sources used in each sensitivity experiment

Experiment	(a)	(b)	(c)	(d)	(e)
P					
PGF	✔		✔		
CSU	✔		✔		
GPCC	✔		✔		
CHIRPS	✔		✔		
TMPA-RT		✔		✔	✔
ET					
PGF–PM	✔	✔	✔	✔	✔
VIC	✔	✔	✔	✔	
ERA	✔	✔			
MERRA	✔	✔			
GLEAM	✔	✔	✔	✔	✔
SRB–CFSR–SEBS	✔	✔	✔	✔	✔
SRB–CFSR–PM	✔	✔	✔	✔	✔
SRB–CFSR–PT	✔	✔	✔	✔	✔
TWSC					
VIC	✔	✔	✔	✔	
GRACE–CSR	✔	✔	✔	✔	✔
GRACE–GFZ	✔	✔	✔	✔	✔
GRACE–JPL	✔	✔	✔	✔	✔
R					
VIC	✔	✔	✔	✔	✔

The data in use in experiment (a) reference; (b) remove gauge; (c) remove reanalysis; (d) remove gauge and reanalysis; (e) remote sensing only, are checked in the table above; the experiments listed above are corresponding to the experiment (a–e) shown in Fig. 1, Fig. 2 in the unconstrained system and Fig. 7 in the constrained system

together, in the unconstrained system. Figure 1a shows the average monthly non-closure estimated from the reference data [experiment (a)]. Figure 1b–e shows the monthly average non-closure estimated by replacing the gauge-adjusted precipitation products with satellite-only product [experiment (b)], removing two reanalysis *ET* products [experiment (c)], removing both gauge-adjusted precipitation and reanalysis *ET* products [experiment (d)], and removing all gauge corrections, reanalysis, and LSM outputs [experiment (e)]. Figure 2 further shows the impact of removing/replacing different category/categories of the datasets from the reference dataset in terms of the mean absolute error (MAE, mm/yr) of the imbalance as a function of mean annual precipitation. It is noticed that the reference data give the best water budget closure across different climate regimes within different mean annual precipitation bins (dark line in Fig. 2). Only removing the reanalysis *ET* products [purple line in Fig. 2, experiment (c)] gives a better budget closure than the other three cases [red, blue and orange lines in Fig. 2, experiments (b), (d) and (e)] when the mean annual precipitation is higher than 30 mm; it gives a poorer closure than the cases with gauge removal [red line in Fig. 2, experiment (b)] and both gauge and reanalysis removal [blue line in Fig. 2, experiment (d)] when the mean annual precipitation is below 30 mm. As the area fraction of mean annual precipitation that is under 30 mm maintains a small portion (<5 %) in which only removing reanalysis gives worse closure, the spatial map of averaged imbalance also shows generally better closure estimates when only the reanalysis *ET* is removed [Fig. 1c, experiment (c)] than the other three cases in which the gauge-adjusted precipitation products are replaced with only using remote sensing products [Fig. 1b, d, e, experiments (b), (d) and (e)]. This indicates the critical role of gauge

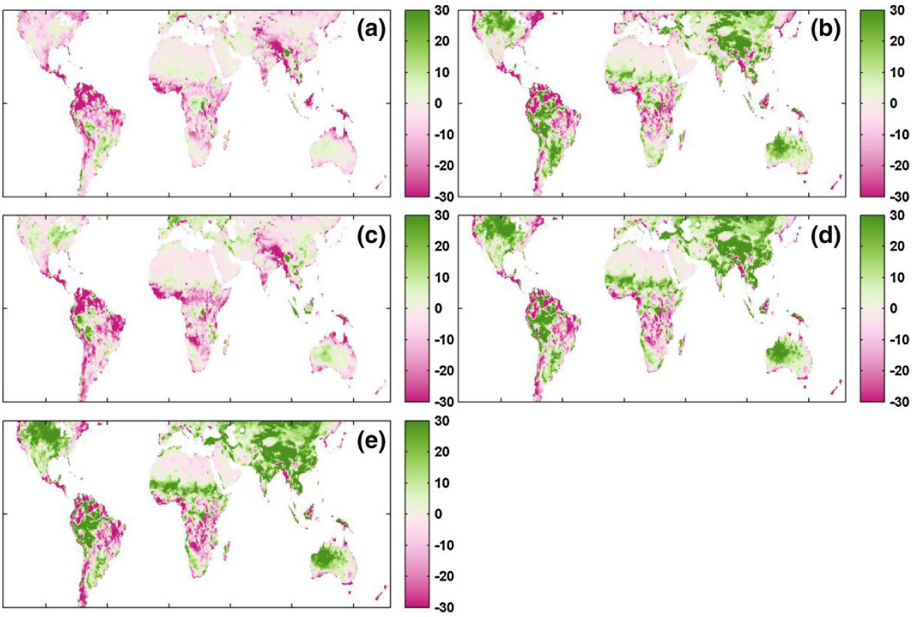

Fig. 1 Sensitivity analysis according to experiments **a**–**e** in terms of the average monthly water budget closure (*P–ET–R–TWSC*, mm/month) during 2004–2007 in the unconstrained system for c **a** reference dataset; **b** in situ precipitation gauge products removed from the reference; **c** reanalysis removed from the reference; **d** both in situ precipitation gauge products and reanalysis removed from the reference; **e** replace the reference data with satellite-only remote sensing products and runoff from VIC

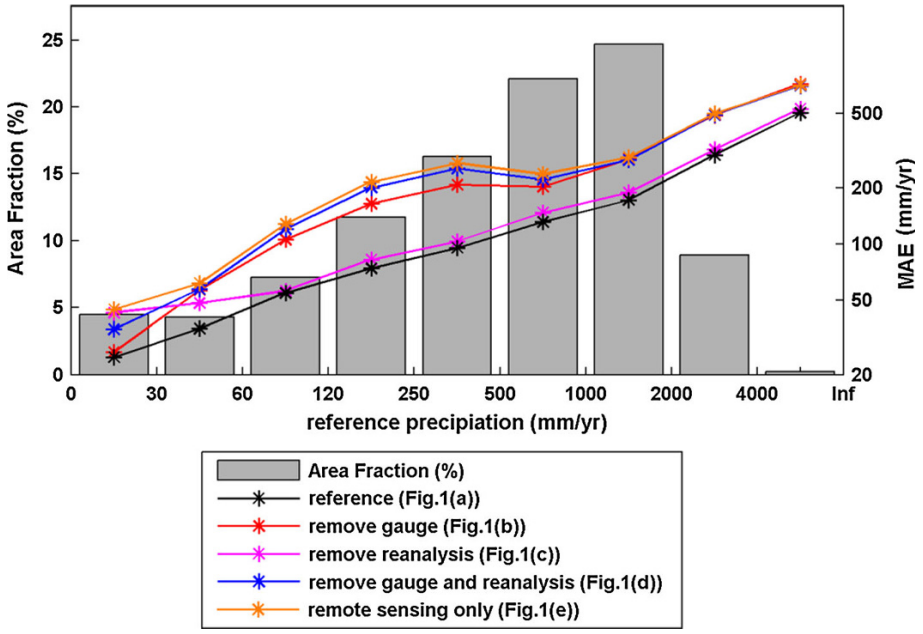

Fig. 2 Mean absolute error (MAE, mm/yr) of the non-closure as a function of the amount of the mean annual precipitation from the reference for experiments **a**–**e** in the unconstrained system. *MAE is in log scale

correction in closing the water budget. The slight difference between Fig. 1b and d reveals the negative impact of further removing reanalysis *ET* products when the gauge-involved precipitation products are replaced with satellite remote sensing only product when the mean annual precipitation is smaller than 1000 mm [red line vs. blue line in Fig. 2, experiment (b) vs. (d)]; meanwhile, the difference between Fig. 1d and e suggests that the further removal of the *TWSC* and *ET* from LSM VIC has a slightly negative impact in the budget closure when the mean annual precipitation is lower than 1000 mm [blue line vs. orange line in Fig. 2, experiment (d) vs. (e)]. However, the three cases in which the gauge-adjusted precipitation products are replaced with remote sensing only product (red, blue, and orange lines in Fig. 2) show very similar performance in closing the budget when the precipitation is larger than 1000 mm/yr.

3.2 Roles of In Situ Precipitation Observations in Water Budget Closure

The role of in situ precipitation observations are investigated by comparing the water budget estimates between gauge-corrected satellite precipitation CSU and the TMPA-RT precipitation estimated solely from satellite observations, together with various remote sensing *ET* products, runoff from VIC, and *TWSC* from GRACE, at multiple temporal scales. In general, at the annual scale, the water budget is more balanced when the gauge-corrected CSU is combined with different remote sensing *ET* products than TMPA-RT as the markers are more aligned with the diagonal lines in Fig. 3**a1**–**e1**. The exceptions are South America, which shows an obvious balance deterioration when combined with *ET* products from SRB–PGF–PM, GLEAM and SRB–CFSR–PM, and Europe which shows a balanced deterioration as well when combined with *ET*s from SRB–CSFR–SEB and SRB–CFSR–PM. Similar to what is revealed by Fig. 3, Fig. 4 also shows improvement in budget estimation, represented by a general better agreement between $P - ET$ (y axis) and VIC-simulated runoff (x axis), from gauge-corrected CSU rather than using the non-gauge-corrected TMPA-RT at basin scales. Both Figs. 3 and 4 show a downshift of $P - ET$

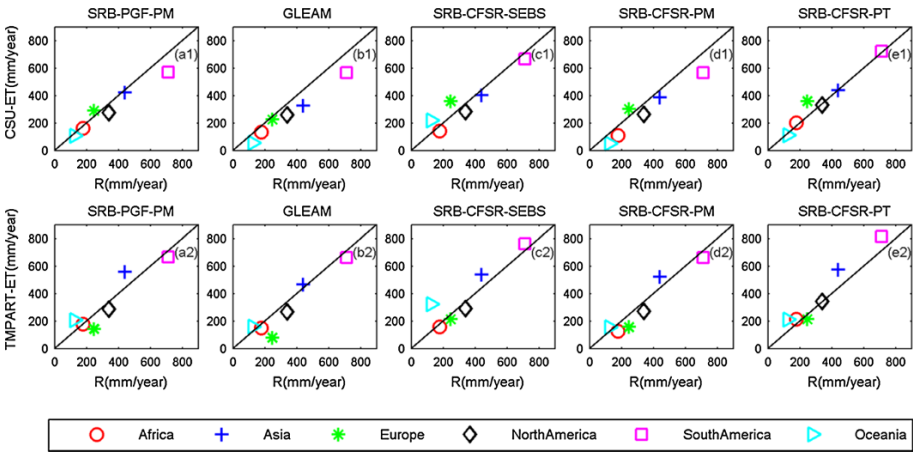

Fig. 3 Comparison of the average annual runoff (mm/year) between the satellite inferred runoff (y axis) from different *ET* products subtracted from ground rain gauge-corrected (CSU, **a1**–**e1**) and non-gauge-corrected (TMPA-RT, **a2**–**e2**) satellite precipitation, and land surface model simulated runoff (R, x axis) during 2004–2007 at continental scales

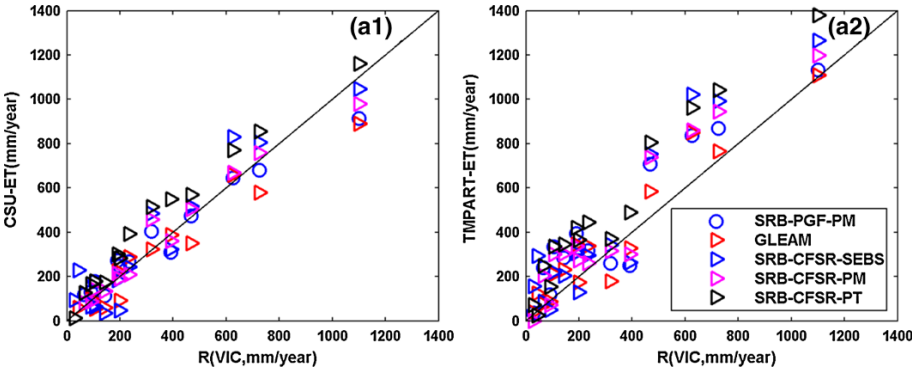

Fig. 4 Comparison of the average annual runoff (mm/year) between the satellite inferred runoff (*y* axis) from different *ET* products subtracted from ground rain gauge-corrected (CSU, **a1**) and non-gauge-corrected (TMPA-RT, **a2**) satellite precipitation, and land surface model simulated runoff (R, *x* axis) during 2004–2007 at basin scales

(*y* axis) calculated from CSU to TMPA-RT at both continental and basin scales, which suggests an overall overestimation in TMPA-RT with respect to gauge observations.

The seasonal cycle of non-closure error estimated by CSU (Figs. 5**a1–f1**, 6**a1–f1**) and TMPA-RT (Figs. 5**a2–f2**, 6**a2–f2**) combined with different remote sensing *ET* products at both continental and basin scales are illustrated in Figs. 5 and 6, respectively. The spread of the non-closure error is relatively higher in South America in both cases when CSU (Fig. 5**e1**) and TMPA-RT (Fig. 5**e2**) are used for water budget estimation, and this indicates high uncertainties among different remote sensing *ET*s exist in South America. In general, the rain gauge-adjusted CSU precipitation products improve the water budget closure at both continental and basin scales relative to the remote sensing only TMPA-RT product, except in South America where the improvements are minimal compared to other continents (Fig. 5). The same happens for the Amazon and Murray Darling basins (Fig. 6).

The non-closure error ratio (%) is calculated as the absolute value of the non-closure divided by the precipitation and then aggregated over space and time at both annual and monthly scales as shown in Table 2. In general, CSU shows lower non-closure ratio than TMPA-RT when combining with most of the *ET* products. However, it shows higher non-closure when CSU is combined with GLEAM *ET* at annual scale, and SRB–CFSR–PM *ET* at both annual and monthly scales.

3.3 Effects of Different Remote Sensing *ET* Products in the Water Budget Closure

Five remote sensing *ET* products are evaluated in terms of non-closure ratio over land at multiple temporal scales when combined with CSU and TMPA-RT in estimating the budgets as listed in Table 3. Globally, SRB–CSFR–PT reaches the minimal non-closure ratio when combined with CSU at both annual (4.2 %) and monthly (8 %) scales. TMPA-RT shows the best estimation of the water budget when they are combined with SRB–CSFR–PM at the annual scale (7.6 %) and with the SRB–CSFR–PT at the monthly scale (9.3 %). Note that such an "optimal" combination of *P* and *ET* (in terms of budget closure) will depend on time and location.

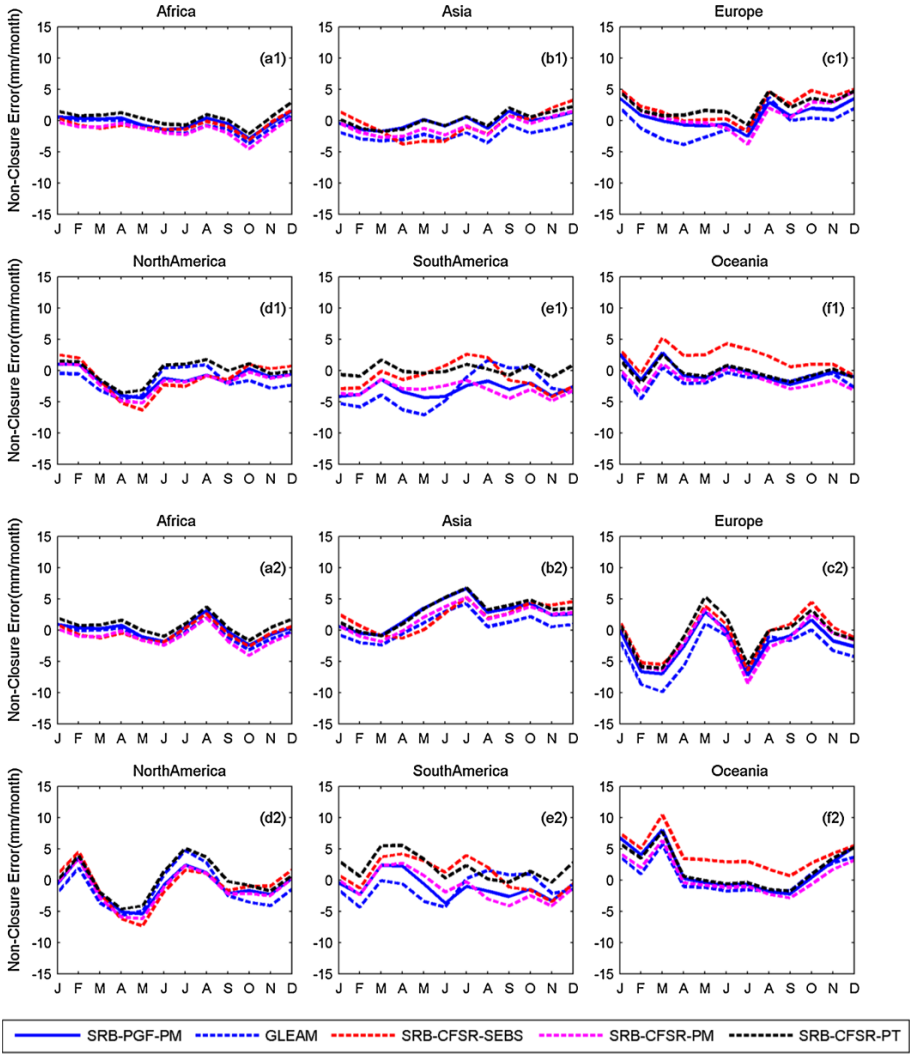

Fig. 5 Comparison of the seasonal cycle of non-closure error estimated by ground rain gauge-corrected (CSU, **a1**–**e1**) and non-gauge-corrected (TMPA-RT, **a2**–**e2**) satellite precipitation, together with different *ET* products, *TWSC* from GRACE, and runoff simulated from VIC during 2004–2007 at continental scales

3.4 Roles of CKF in Constraining the Water Balance

Though the gauge-adjusted CSU product gives better water budget closure than the non-gauge-adjusted product TMPA-RT, neither exactly closes the water budget in the unconstrained system. In order to enforce the water balance, the CKF algorithm is applied, which assures closure of the budget at each global grid cell for experiments (a–e) in the constrained system. The precipitation and *ET* from the reference data [experiment (a)] in the constrained system are considered as one benchmark here to evaluate the performance of experiments (b–e). The relative root-mean-square error [RMSE (%)] of the constrained

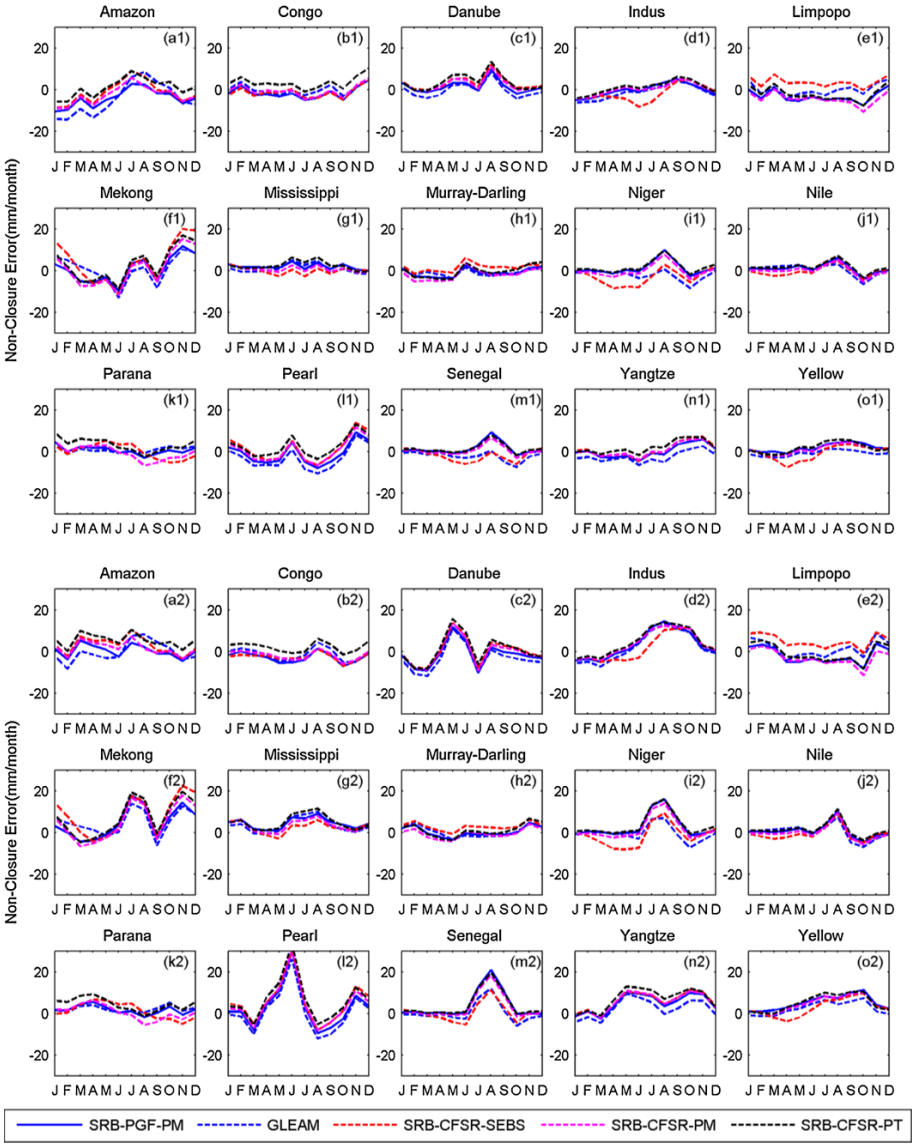

Fig. 6 Comparison of the seasonal cycle of non-closure error estimated by ground rain gauge-corrected (CSU, **a1**–**o1**) and non-gauge-corrected (TMPA-RT, **a2**–**o2**) satellite precipitation, together with different *ET* products, *TWSC* from GRACE, and runoff simulated from VIC during 2004–2007 at basin scales

precipitation and *ET* estimates from experiments (b–e) is calculated against the reference dataset [experiment (a)]. In general, it shows that the closest *P* and *ET* estimates to the budget estimates, from the reference in the constrained system, when only the reanalysis is removed [purple line in Fig. 7, experiment (c)], while the poorest estimates are when all possible non-satellite sources are removed [orange line in Fig. 7, experiment (e)]. It is also found that removing the reanalysis causes less impact to the precipitation estimate (top

Table 3 Non-closure ratio (%, non-closure/precipitation) at multiple temporal scales over land

		Annual	Monthly
CSU	SRB–PGF–PM	4.9	9.4
	GLEAM	8.6	9.6
	SRB–CFSR–SEBS	8.7	10.7
	SRB–CFSR–PM	9.0	11.6
	SRB–CFSR–PT	**4.2**	**8.0**
TMPA-RT	SRB–PGF–PM	8.7	10.4
	GLEAM	8.3	10.0
	SRB–CFSR–SEBS	10.4	11.2
	SRB–CFSR–PM	**7.6**	10.3
	SRB–CFSR–PT	7.7	**9.3**

The numbers in bold indicate the best water budget estimates (with the least non-closure) that are optimally combined of *P* and *ET* from different products

panel in Fig. 7) than to the *ET* estimate (bottom panel in Fig. 7), especially when the reference annual precipitation and *ET* are relatively low.

This study aims at exploring the potential of the satellite remote sensing in replacing the reference data for budget estimation, which corresponds to experiment (e). Figure 8 shows an example of the water budget components (*P*, *ET* and *R* in Fig. 8a–c, and *TWSC* in Fig. 8d–f); water budget closure in the unconstrained system (Fig. 8g); the average attribution of each component with the remote sensing products [Fig. 8h corresponding to experiment (e)]; and the reference data [Fig. 8i corresponding to experiment (a)] are used to estimate the water budget throughout 2004–2007 for the unconstrained (first column) and constrained (second and third columns) systems for the Amazon River basin. Over the Amazon basin where the rainfall is heavy and the gauges are relatively sparse, the uncertainties in rainfall are higher, making precipitation the major recipient of the budget error attribution (Fig. 8h–i). The constrained water budget fluxes (*P*, *ET* and *R*) and *TWSC* estimated from remote sensing are similar to that estimated from the reference data. This suggests that though the reference data seem to close better the water budget (Fig. 1a) compared to the remote sensing data (Fig. 1e) in the unconstrained system, the budget estimates in the constrained system via CKF data assimilation approach are quite comparable between the remote sensing (Fig. 8b, e) and the reference data (Fig. 8c, f).

Further comparison of the monthly average terrestrial water budget estimates in both the unconstrained and constrained systems during 2004–2007 is shown in Fig. 9. The differences between the water budget components estimated from remote sensing and the reference data in both unconstrained system and constrained system are shown in Fig. 9**f1–l1** and **f2–l2**, respectively. There is no difference in runoff between reference and remote sensing in the unconstrained system (Fig. 9k1) as VIC runoff is the only source for both experiments. In the constrained system, such differences show up as the CKF redistributes budget errors. For precipitation, the difference between reference and remote sensing drops from the unconstrained system to the constrained system (Fig. 9**i1** vs. **i2** with the bias (%) dropping from 232 to −2 %, RMSE (%) dropping from 548 to 45 % and the correlation coefficient (CC) increasing from 0.43 to 0.94). Decreases in the differences between *ET*/*TWSC* estimated from the reference and remote sensing are also found when the constraint is enforced (Fig. 9**j1** vs. **j2** and **l1** vs. **l2**). The estimates from the reference data and the remote sensing in the constrained system are quite comparable, and this suggests good

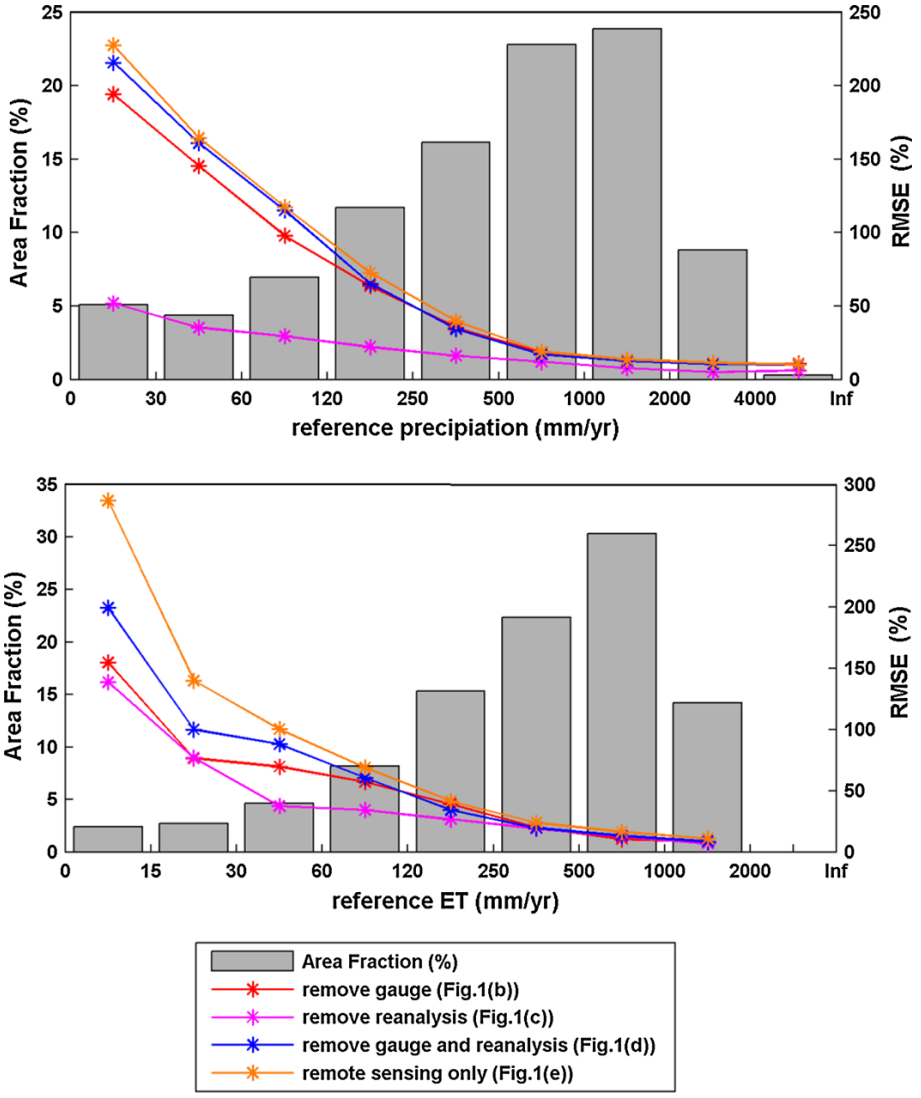

Fig. 7 RMSE (%) of the non-closure as a function of mean annual precipitation and ET from the reference budget estimates for experiments (*b–e*) in the constrained system

potential of satellite remote sensing to provide a very good estimate of the global water budget if properly constrained.

3.5 Runoff Validation Against GRDC Data at the Basin Scale

In situ observational runoff data for 32 large basins as used in Pan et al. (2012) and 331 medium-sized basins (with sizes no larger than 10^4 km^2) were collected from Global Runoff Data Centre (GRDC). The data availabilities vary from basin to basin. By filtering

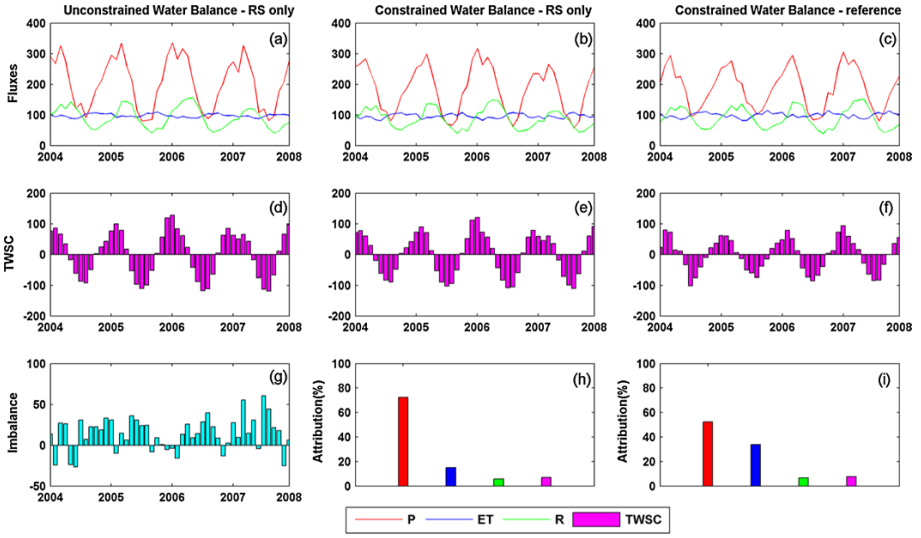

Fig. 8 Unconstrained water budget (*P*, *ET* and *R* in **a**, *TWSC* in **d**) estimated from satellite remote sensing, constrained water budget estimated from satellite remote sensing (*P*, *ET* and *R* in **b**, *TWSC* in **e**) and the reference data (*P*, *ET* and *R* in **c**, *TWSC* in **f**) over the Amazon River basin. **g** is the imbalance before the water budget constraint. The imbalance after the water budget constraint equals to zero and the attribution from each water budget components are shown in **h**, **i** for unconstrained and constrained system, respectively

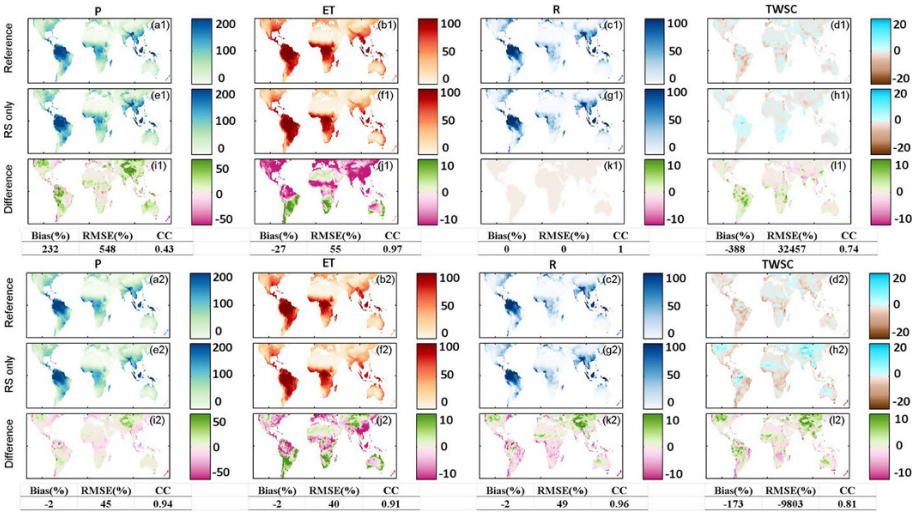

Fig. 9 Comparison of the average monthly water budgets (mm/month) estimated from the reference data (**a1**–**d1**) and (**a2**–**d2**), the satellite remote sensing data (**e1**–**h1**) and (**e2**–**h2**), and the differences between the reference data and the remote sensing products in estimating the water budgets in the unconstrained (**i1**–**l1**) and constrained systems (**i2**–**l2**). *The corresponding statistic indices of water budgets estimated from remote sensing relative to the reference data are listed under each budget terms for both unconstrained and constrained system

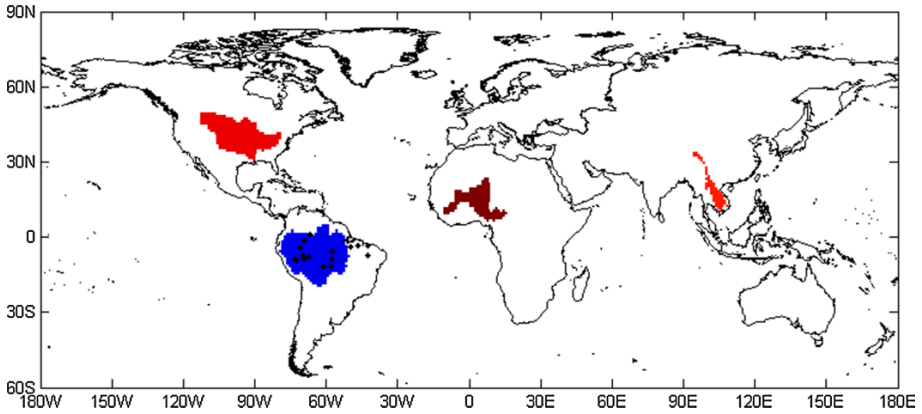

Fig. 10 Locations of the four large river basins and the outlets of the 16 medium river basins for validation which are shown as + symbols

out those gauge observations with data availability that cover our research period 2004–2007, only four large basins and 16 medium-sized basins are left for validation (Fig. 10). This illustrates the challenge of identifying in situ runoff data for validating Earth observations. The basin sizes of the 16 medium basins are ranging from 10,400 to 363,000 km^2. The four large basins are evenly distributed in different continents, but all of the 16 medium-sized basins are located in South America. Figure 11 illustrated the monthly runoff time series estimated from a combination of different data sources, which correspond to experiments (a–e) in the constrained system, compared against the GRDC observation. The runoffs estimated and constrained from different data sources from experiments (a–e) align well with each other for those four large river basins except for Mississippi river basin, which has the densest in situ observations that may lead to large variations among different data sources for the same water budget component (particularly precipitation during the summer season). The larger variation in any of the budget component will impact the budgets estimated and constrained from the constrained system. The poor estimates of runoff in Niger River basin might be caused by the precipitation error over that region due to limited in situ observation of precipitation in Africa that would impact the data quality of precipitation products from different sources, or features like wetlands that are not well represented in VIC.

Figure 12 further provides the comparison of monthly mean runoff estimated from the five experiments with different data sources against GRDC observations for the sixteen medium-sized river basins. Similar to what is shown in Fig. 11, the runoff estimated and constrained from different data sources from experiments (a–e) are in general agreement, which demonstrate that using satellite remote sensing only data sources [experiment (e)] together with the CKF in the constrained system can achieve similar budget estimates (e.g., for runoff) to other experiments (a–d) with different data sources. However, some of the runoff estimated from the experiments does not align well with GRDC observations. This is possibly due to the coarse spatial resolution (0.5°) of the estimated budgets from the experiments. Uncertainties exist in the basin mask files for extracting the monthly runoff

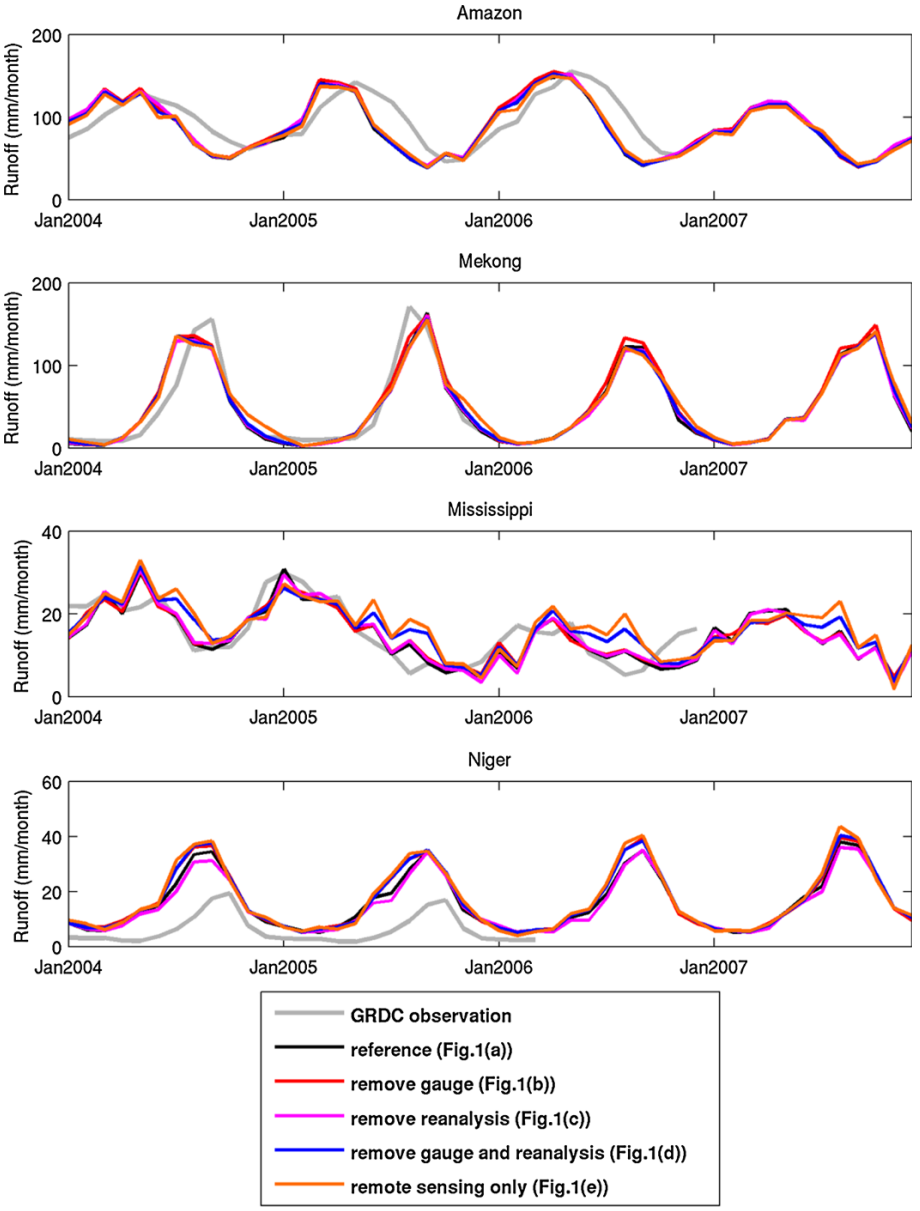

Fig. 11 Comparison of monthly time series of runoff estimated from experiments (*a–e*) from the constrained system over the four large river basins against GRDC observation

from the global datasets, and these uncertainties would be particular large for those basins which only cover a limited number of pixels, though we have assigned a fraction to those boundary pixels in our basin mask extraction algorithm.

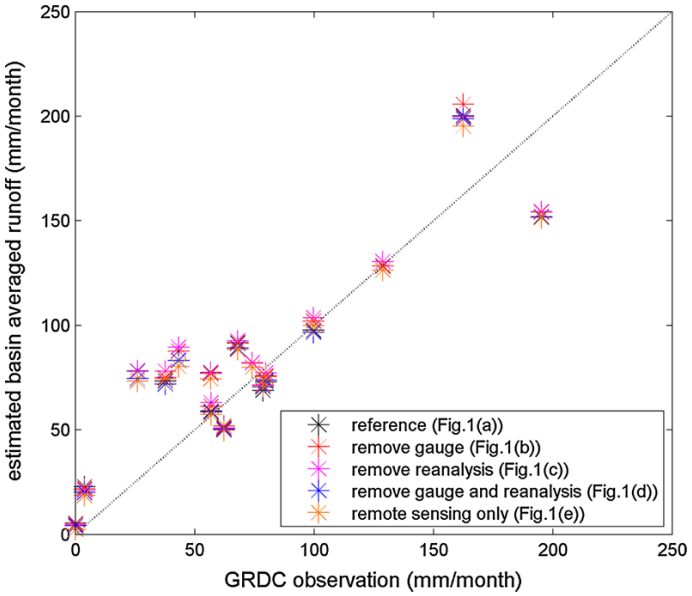

Fig. 12 Comparison of monthly mean runoff estimated from experiments (*a*–*e*) from the constrained system over the 16 medium-sized river basins against GRDC observation

4 Conclusions

This study creates closed terrestrial, global water budget estimates using satellite remote sensing at a 0.5° spatial resolution and monthly timescale for the period of 2004–2007. We also investigate the roles of data from different sources in estimating and closing the budget, and particularly assess the capability of the satellite remote sensing in closing the terrestrial water budget at the 0.5° spatial resolution through constrained data assimilation.

The water budget closure is sensitive to the gauge-adjusted precipitation but is less sensitive to reanalysis *ET* products in the unconstrained system. Comparison among the water budget closure datasets that were estimated by removing/replacing one or more categories of data from the reference set shows that non-satellite data sources can help to reduce the non-closure errors significantly. Particularly, replacing the gauge-corrected precipitation datasets with satellite-only TMPA-RT for the budget estimation leads to larger non-closure errors, which reflects a high bias in TMPA-RT.

The rain gauge-corrected CSU precipitation product, when combined with different remote sensing *ET* products, *TWSC* from GRACE and runoff from VIC, outperforms the non-gauge-corrected TMPA-RT precipitation at multiple spatial and temporal scales in the water budget closure. The discrepancy between the non-closure errors based on using CSU or TMPA-RT precipitation demonstrates the need of a rain gauge adjustment for global gridded satellite-based water budget estimates.

As neither the reference data [experiment (a)] nor the remote sensing data [experiment (e)] can exactly close the water budget in the unconstrained system, the CKF data assimilation approach is applied to constrain the budget closure. Once constrained, the differences between budget estimates from the reference data and from satellite-only remote sensing become fairly small, suggesting a good potential for being able to use remote sensing alone to reasonably reproduce the water budgets.

Runoff estimated from experiments (a–e) with constraints is validated against GRDC observations. A limited number of basins are applied for validation due to the limited data availability for the study period. The validation results show that the runoff estimated from remote sensing only in experiment (e) can achieve very similar runoff estimates from other experiments (a–d) with more categories of data sources, and meanwhile is lined up fairly well with GRDC runoff expect for those smaller basins with mask uncertainties for the coarse spatial resolution at 0.5°.

As we demonstrate the potential for remote sensing in water budget studies, further efforts are still needed to understand the discrepancies among different data sources such as in situ observations, satellite remote sensing, land surface models, and reanalysis data products in closing the water budget. With the current satellite missions such as the Global Precipitation Mission (GPM) and Soil Moisture Active Passive (SMAP) mission, and future missions such as Surface Water Ocean Topography (SWOT) mission for global surface elevations and river stage and discharge, it is very promising that the community can retrieve accurate depiction of the global water budget at fine resolutions and over a longer time period.

Acknowledgments This paper is part of the special issue of the ISSI workshop on remote sensing and water resources. This work is supported by NASA research grants NNX08AN40A "Developing Consistent Earth System Data Records for the Global Terrestrial Water Cycle" and NNX09AK35G "Development and diagnostic analysis of a multi-decadal global evaporation product for NEWS"

References

Adler RF, Kidd C, Petty G, Morissey M, Goodman HM (2001) Intercomparison of global precipitation products: the third Precipitation Intercomparison Project (PIP-3). Bull Am Meteorol Soc 82:1377–1396

Bytheway JL, Kummerow CD (2013) Inferring the uncertainty of satellite precipitation estimates in data-sparse regions over land. J Geophys Res Atmos 118:9524–9533. doi:10.1002/jgrd.50607

Dirmeyer PA, Gao X, Zhao M, Guo Z, Oki T, Hanasaki N (2006) GSWP-2: multimodel analysis and implications for our perception of the land surface. Bull Am Meteorol Soc 87:1381–1397. doi:10.1175/BAMS-87-10-1381

Durand M, Fu L-L, Lettenmaier DP, Alsdorf DE, Rodriguez E, Esteban-Fernandez D (2010) The surface water and ocean topography mission: observing terrestrial surface water and oceanic submesoscale eddies. Proc IEEE 98:766–779

Famiglietti JS (2014) The global groundwater crisis. Nat Clim Change 4:945–948. doi:10.1038/nclimate2425

Fisher JB, Tu KP, Baldocchi DD (2008) Global estimates of the land–atmosphere water flux based on monthly AVHRR and ISLSCP-II data, validated at 16 FLUXNET sites. Remote Sens Environ 112:901–919

Funk CC et al (2014) A quasi-global precipitation time series for drought monitoring. US Geol Surv Data Ser. doi:10.3133/ds832

Gao H, Tang Q, Ferguson CR, Wood EF, Lettenmaier DP (2010) Estimating the water budget of major US river basins via remote sensing. Int J Remote Sens 31:3955–3978

Haddeland I et al (2011) Multimodel estimate of the global terrestrial water balance: setup and first results. J Hydrometeorol 12:869–884

Hong Y, Gochis D, Cheng J-T, Hsu K-L, Sorooshian S (2007) Evaluation of PERSIANN-CCS rainfall measurement using the NAME event rain gauge network. J Hydrometeorol 8:469–482

Huffman GJ et al (2007) The TRMM multisatellite precipitation analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. J Hydrometeorol 8:38–55

Huffman GJ, Adler RF, Bolvin DT, Nelkin EJ (2010) The TRMM multi-satellite precipitation analysis (TMPA). Springer, Satellite rainfall applications for surface hydrology, pp 3–22

Jacob T, Wahr J, Pfeffer WT, Swenson S (2012) Recent contributions of glaciers and ice caps to sea level rise. Nature 482:514–518

Joyce RJ, Janowiak JE, Arkin PA, Xie P (2004) CMORPH: a method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. J Hydrometeorol 5:487–503
Landerer FW, Swenson SC (2012) Accuracy of scaled GRACE terrestrial water storage estimates. Water Resour Res 48(4)
Luo L, Wood EF, Pan M (2007) Bayesian merging of multiple climate model forecasts for seasonal hydrological predictions. J Geophys Res Atmos 1984–2012:112
Miralles DG, Holmes TRH, De Jeu RAM, Gash JH, Meesters AGCA, Dolman AJ (2011) Global land-surface evaporation estimated from satellite-based observations. Hydrol Earth Syst Sci 15:453–469. doi:10.5194/hess-15-453-2011
Mu Q, Heinsch FA, Zhao M, Running SW (2007) Development of a global evapotranspiration algorithm based on MODIS and global meteorology data. Remote Sens Environ 111:519–536
Oki T, Musiake K, Matsuyama H, Masuda K (1995) Global atmospheric water balance and runoff from large river basins. Hydrol Process 9:655–678
Paiva RCD, Durand MT, Hossain F (2015) Spatiotemporal interpolation of discharge across a river network by using synthetic SWOT satellite data. Water Resour Res 51:430–449
Pan M, Sahoo AK, Troy TJ, Vinukollu RK, Sheffield J, Wood EF (2012) Multisource estimation of long-term terrestrial water budget for major global river basins. J Clim 25:3191–3206
Pavelsky TM, Durand MT, Andreadis KM, Beighley RE, Paiva RCD, Allen GH, Miller ZF (2014) Assessing the potential global extent of SWOT river discharge observations. J Hydrol 519:1516–1525
Rienecker MM et al (2011) MERRA: NASA's modern-era retrospective analysis for research and applications. J Clim 24:3624–3648. doi:10.1175/JCLI-D-11-00015.1
Rodell M et al (2015) The observed state of the water cycle in the early 21st century. J Clim. doi:10.1175/JCLI-D-14-00555.1
Sahoo AK, Pan M, Troy TJ, Vinukollu RK, Sheffield J, Wood EF (2011) Reconciling the global terrestrial water budget using satellite remote sensing. Remote Sens Environ 115:1850–1865
Sakumura C, Bettadpur S, Bruinsma S (2014) Ensemble prediction and intercomparison analysis of GRACE time-variable gravity field models. Geophys Res Lett 41:1389–1397
Schneider U, Becker A, Finger P, Meyer-Christoffer A, Ziese M, Rudolf B (2014) 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. doi:10.1007/s00704-013-0860-x
Sheffield J, Wood EF (2007) Characteristics of global and regional drought, 1950–2000: analysis of soil moisture data from off-line simulation of the terrestrial hydrologic cycle. J Geophys Res Atmos 112:D17115. doi:10.1029/2006JD008288
Sheffield J, Goteti G, Wood EF (2006) Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. J Clim 19:3088–3111
Sheffield J, Ferguson CR, Troy TJ, Wood EF, McCabe MF (2009) Closing the terrestrial water budget from satellite remote sensing. Geophys Res Lett 36(7)
Simmons A, Uppala S, Dee D, Kobayashi S (2006) ERA-interim: new ECMWF reanalysis products from 1989 onwards. ECMWF Newsl 110:26–35
Tapley BD, Bettadpur S, Ries JC, Thompson PF, Watkins MM (2004) GRACE measurements of mass variability in the Earth system. Science 305:503–505
Thomas AC, Reager JT, Famiglietti JS, Rodell M (2014) A GRACE-based water storage deficit approach for hydrological drought characterization. Geophys Res Lett 41:1537–1545
Tian Y, Peters-Lidard CD (2010) A global map of uncertainties in satellite-based precipitation measurements. Geophys Res Lett 37(24)
Trenberth KE, Smith L, Qian T, Dai A, Fasullo J (2007) Estimates of the global water budget and its annual cycle using observational and model data. J Hydrometeorol 8:758–769
Troy TJ, Sheffield J, Wood EF (2011) Estimation of the terrestrial water budget over northern eurasia through the use of multiple data sources. J Clim 24:3272–3293. doi:10.1175/2011JCLI3936.1
Vinukollu RK, Meynadier R, Sheffield J, Wood EF (2011) Multi-model, multi-sensor estimates of global evapotranspiration: climatology, uncertainties and trends. Hydrol Process 25:3993–4010. doi:10.1002/hyp.8393
Wahr J, Swenson S, Zlotnicki V, Velicogna I (2004) Time-variable gravity from GRACE: first results. Geophys Res Lett 31(11)
Wang S, Huang J, Li J, Rivera A, McKenney DW, Sheffield J (2014) Assessment of water budget for sixteen large drainage basins in Canada. J Hydrol 512:1–15
Weedon GP et al (2011) Creation of the WATCH forcing data and its use to assess global and regional reference crop evaporation over land during the twentieth century. J Hydrometeorol 12:823–848