



Bridging the gap between GRACE and GRACE-FO using a hydrological model

Xu Zhang ^{a,*}, Jinbao Li ^{a,b}, Qianjin Dong ^c, Zifeng Wang ^a, Han Zhang ^a, Xiaofeng Liu ^a

^a Department of Geography, University of Hong Kong, Hong Kong SAR, China

^b HKU Shenzhen Institute of Research and Innovation, Shenzhen 518057, China

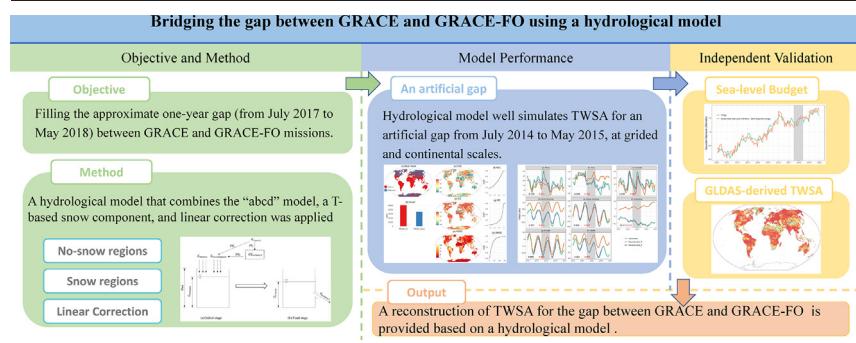
^c State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, China



HIGHLIGHTS

- We bridged the gap between GRACE and GRACE-FO using a hydrological model.
- Our reconstruction for an artificial gap achieves high correlation with observations.
- The reconstruction for the real gap fits well against sea-level budget and GLDAS.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:

Received 5 October 2021

Received in revised form 26 December 2021

Accepted 30 January 2022

Available online 3 February 2022

Editor: Martin Drews

Keywords:

Terrestrial water storage

Gap filling

GRACE

GRACE-FO

ABSTRACT

Gravity Recovery and Climate Experiment (GRACE) and GRACE Follow-On (GRACE-FO), two successive satellite-based missions starting in 2002, have provided an unprecedented way of measuring global terrestrial water storage anomalies (TWSA). However, a temporal gap exists between GRACE and GRACE-FO products from July 2017 to May 2018, which introduces bias and uncertainties in TWSA calculations and modeling. Previous studies have incorporated hydroclimatic factors as predictors for filling the gap, but most of them utilized artificial intelligence or pure statistical models that generally de-trended TWSA and had no physical foundation. Thus, a physically-based reconstruction is required for increasing robustness. In this study, we bridge the temporal gap by developing an empirical hydrological model. The "abcd" model, a T-based snow component, and linear correction are utilized to represent runoff generation, snow dynamics, and long-term trends. The testing results indicate that our hydrological model can successfully reconstruct TWSA in tropical, temperate, and continental climates, although further improvement is needed for arid climates. Our reconstruction for the gap achieves high accuracy and robustness as shown by the evaluations against sea-level budget and GLDAS-derived TWSA. Compared to previous studies using artificial intelligence or statistical techniques, our hydrological model performs similarly in the gap filling but does not involve de-trended or de-seasonalized transformations, which will facilitate the combination of GRACE and GRACE-FO products and improve the physical understanding of global TWSA.

1. Introduction

The Gravity Recovery and Climate Experiment (GRACE) mission and its successor GRACE Follow-On (GRACE-FO) provide a global measure of terrestrial water storage anomalies (TWSA) with unprecedented accuracy since

2002 (Kornfeld et al., 2019; Landerer et al., 2020; Tapley et al., 2004; Tapley et al., 2019). The distance between two twin satellites is used to derive the changes in global gravity fields, which enables the quantification of large-scale water redistribution and movement in different forms (Fatolazadeh and Goita, 2021; Mehrnegan et al., 2021; Tapley et al., 2019; Wang et al., 2020a; Wang et al., 2020b). The GRACE products have been extensively applied in the fields of environmental studies, earth science, and so forth. For instance, the ice melt, ocean mass change, and groundwater depletion, previously

* Corresponding author.

E-mail address: zhangxu_hku@connect.hku.hk (X. Zhang).

measured using altimetry method or inferred from model simulation, can be directly quantified from GRACE products (Chen et al., 2019; Luthcke et al., 2006; Reager et al., 2016; Schlegel et al., 2016; Velicogna and Wahr, 2005; Wang et al., 2018a). The occurrence of floods and droughts can also be monitored or predicted using GRACE products (Houborg et al., 2012; Reager et al., 2014; Satish Kumar et al., 2021; Sun et al., 2018; Wang et al., 2020a). Moreover, in combination with the hydrological models, GRACE data have been widely adopted to determine human impacts on TWSA (Deng and Chen, 2017; Felfelani et al., 2017; Forootan et al., 2019; Li et al., 2017), estimate water budget in typical catchments (Chen et al., 2020; Swann and Koven, 2017), and improve hydrological model performance using data assimilation (Houborg et al., 2012; Li et al., 2019a; Werth et al., 2009).

Despite the extensive applications, a noteworthy drawback of GRACE products is that there exists an approximate one-year gap (from July 2017 to May 2018) between GRACE and GRACE-FO missions, which inhibits full utilization of the available TWSA measurements. Filling this gap is thus crucial for more reliable and wider application of GRACE and GRACE-FO products. Products from Swarm satellites and satellite laser ranging (SLR) have been attempted to reconstruct TWSA and reproduce TWSA variations at large scales (Löcher and Kusche, 2020; Richter et al., 2021). However, the spatial resolution is not comparable to GRACE and GRACE-FO datasets, resulting in large uncertainty for some small catchments and high-latitudes (Li et al., 2021). In light of the close relationship between TWSA and hydroclimatic variables, some studies used hydroclimatic factors as predictors and reconstructed GRACE-derived TWSA at the gridded resolution. For example, drawing on a statistical model, Humphrey et al. (2017) and Humphrey and Gudmundsson (2019) used precipitation (P) and temperature (T) as explanatory variables and reconstructed the de-seasonalized and de-trended TWSA. Sun et al. (2020) employed a deep neural network with P , T , and soil moisture in the Global Land Data Assimilation System (GLDAS) model as explanatory variables for the prediction of TWSA. Li et al. (2020) and Li et al. (2021) reconstructed global terrestrial water storage change (TWSC) and TWSA for the gap period through various statistical methods. Moreover, regional TWSA reconstructions have been accomplished for India (Sun et al., 2019), Africa (Ahmed et al., 2019), and China (Jing et al., 2020).

Although many studies have been made to reconstruct TWSA using hydroclimatic factors, they are based upon either AI models or pure statistical models. The rationale behind AI models is yet not well understood, which may introduce uncertainty when applying them into practice, especially when the values are out of the range of calibration (Chollet, 2018). Some fundamental hydrological processes (e.g., antecedent water storage) and the principle of water balance are not fully accounted in the statistical models (Humphrey and Gudmundsson, 2019). Moreover, some studies attributed the long-term trend in TWSA to human activities, and thus only reconstructed the de-trended TWSA based on hydrological variables (Humphrey and Gudmundsson, 2019; Li et al., 2021). However, previous studies have shown that the long-term TWSA trend in Australia, South America, and Africa is likely caused by climate change and/or natural environmental variability, which should be expressed from hydroclimatic variables (Boening et al., 2012; Gaughan and Waylen, 2012; Rodell et al., 2018). Consequently, the de-trended transformation may further undermine the physical basis of the reconstruction and decrease its validity and suitability for potential applications of the GRACE products.

Therefore, the main objective of this study is to fill the circa one-year gap of GRACE and GRACE-FO products using an empirical hydrological model. Hydrological processes are incorporated and raw series without de-trended transformation is reconstructed. P , E , and T will be used as three input variables to drive the water storage change, with a T -based snow component added for the high-latitudes. The long-term trend that may be caused by human activities will be simulated from linear correction. The performance of the TWSA reconstruction will be tested against an artificial gap from July 2014 to May 2015, and the reconstruction for the actual

gap will be examined against sea-level budget and GLDAS-derived TWSA. This study is expected to advance the simulation of TWSA with hydrological models and promote the full usage of GRACE and GRACE-FO products in various fields.

2. Data and methods

2.1. Data

2.1.1. GRACE and GRACE-FO products

The Release 06 Version 02 GRACE and GRACE-FO datasets (RL06_v02) based on mascon resolution by the Jet Propulsion Laboratory (JPL) were used in this study, which span from May 2002 to February 2021 (Kornfeld et al., 2019; Landerer et al., 2020; Watkins et al., 2015; Wiese et al., 2016; Wiese et al., 2019). The period from May 2002 to June 2017 was measured from GRACE, while the period from June 2018 to February 2021 was measured from GRACE-FO. The remaining 11 months from July 2017 to May 2018 constitute the gap period, and will be reconstructed in this study. All forms of water, including surface water, soil moisture, groundwater, snow and ice, are provided at a 0.5-degree resolution with an appropriately 1-month time interval by JPL. The TWSA data in GRACE and GRACE-FO are the anomalies and differ a constant from real terrestrial water storage (TWS). The raw measurements in GRACE and GRACE-FO products are unevenly distributed over time, with certain measurements not located in the middle of a month. Therefore, cubic spline interpolation was used to derive the TWSA in the middle of each month to represent monthly mean value. Cubic Hermite and linear interpolation methods were also examined, which produced very similar results (Fig. S1).

2.1.2. Climatic data

Precipitation (P), temperature at 2 m above land surface (T), and evaporation (E) are used as three hydroclimatic driving factors of TWS, with data collected from the European Centre for Medium-Range Weather Forecasts (ERA5) dataset from January 1998 to February 2021 at $0.25^\circ \times 0.25^\circ$ resolution (Hersbach et al., 2020). The P in ERA5 shows higher correlation with the observed monthly precipitation than ERA-Interim, MERRA-2, and JRA-55 reanalysis (Hersbach et al., 2020). Therefore, the ERA5 has been used to replace the ERA-Interim for GRACE-based analysis (Chen et al., 2020; Eicker et al., 2020; Sun et al., 2020). The meteorological data in ERA5 were regridded to $0.5^\circ \times 0.5^\circ$ resolution based on the first-order conservative remapping technique (Jones, 1999).

2.2. Hydrological model for reconstruction

2.2.1. Model for no-snow area (Model_no)

Our model is first introduced to simulate the hydrological processes for no-snow areas (model_no), and then a T -based snow component is added in Section 2.2.2 for snow areas (model_snow). The model_no is based on the “abcd” model (Thomas, 1981) and Budyko hypothesis (Budyko, 1974; Fu, 1981; Zhang et al., 2020). Two stages are assumed and schematically shown in Fig. 1. In the initial stage of month t , some terrestrial water has already been saved and is denoted by $G_{initial,t}$. Then, the difference between P_t and E_t in month t is equal to the meteorological input water to the ground (Fig. 1a). To consider the large uncertainty in E estimation, here we applied a linear transformation to correct E_t , and $G_{input,t}$ is expressed as

$$G_{input,t} = P_t - (\alpha_0 + \alpha_1 E_t) \quad (1)$$

The $G_{input,t}$ and $G_{initial,t}$ constitute the available water for the following runoff generation. Therefore, the summation of $G_{input,t}$ and $G_{initial,t}$ is termed as water availability ($G_{available,t}$).

$$G_{available,t} = G_{input,t} + G_{initial,t} \quad (2)$$

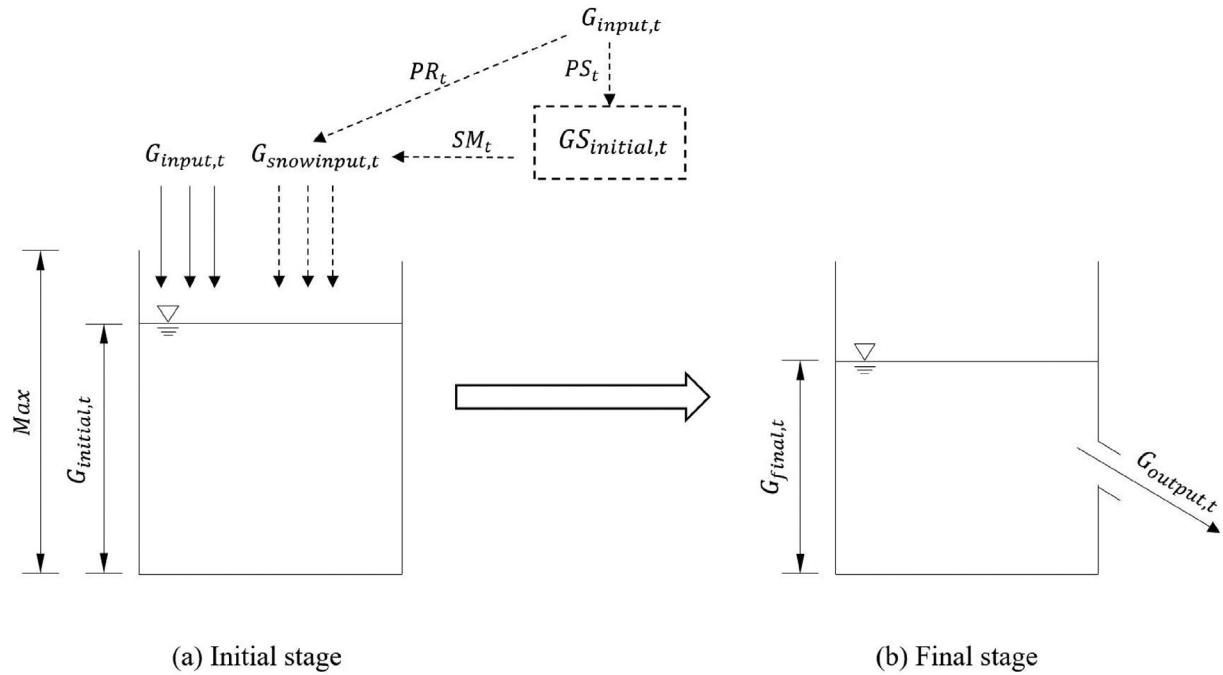


Fig. 1. Structure of the conceptual hydrological model at the initial (a) and final (b) stage. Dashed lines denote the component of snow formation.

In month t , some water leaves the land as R , and $G_{available,t}$ will decrease to $G_{final,t}$ at the end of month t (Fig. 1b). A supply and storage limit of $G_{final,t}$ was assumed to derive the equation of $G_{final,t}$ (Thomas, 1981; Zhang et al., 2008). Specifically, the boundary conditions with large storage capacity ($\frac{G_{available}}{Max} \rightarrow 0$) and sufficient water supply ($\frac{G_{available}}{Max} \rightarrow \infty$) are

$$G_{final} \rightarrow G_{available}, \text{ as } \frac{G_{available}}{Max} \rightarrow 0 \quad (3)$$

$$G_{final} \rightarrow \text{Max}, \text{ as } \frac{G_{available}}{Max} \rightarrow \infty \quad (4)$$

where Max is the maximum value for water storage. Yang equation, a Budyko-like equation, is used here to simulate the relationship between $G_{final,t}$ and $G_{available,t}$ with a parameter w (Yang et al., 2008).

$$G_{final,t} = (G_{available,t}^{-w} + Max^{-w})^{-\frac{1}{w}}, w > 0 \quad (5)$$

The $G_{final,t}$ will be used as the $G_{initial,t+1}$ in the next month to represent the accumulative influence of antecedent hydroclimatic drivers. The mean monthly TWS is defined as the average of $G_{initial,t}$ and $G_{final,t}$

$$G_{mean,t} = \frac{1}{2} (G_{initial,t} + G_{final,t}) \quad (6)$$

The objective function is to find the minimum squared error between the observed and simulated $G_{mean,t}$. However, GRACE can only measure the anomalies of TWS that differ a constant from our prediction. Thus, the mean difference between the predicted and observed values is removed at first. Then, the squared error of TWSA is used as the objective function.

$$Diff = \frac{\sum_{t=1}^n G_{mean,t}}{n} - \frac{\sum_{t=1}^n G_{GRACE,t}}{n} \quad (7)$$

$$\text{Squared Error} = \sum_{t=1}^n (G_{mean,t} - Diff - G_{GRACE,t})^2 \quad (8)$$

where $G_{GRACE,t}$ is the observed TWSA derived from GRACE in month t .

2.2.2. Model for snow area (Model_snow)

For snow-dominated areas, a T -based snow accumulation and melt component is added on model_no to consider the impacts of snow on TWSA (Hay and McCabe, 2002; Martinez and Gupta, 2010). The quantity $G_{input,t}$ is first partitioned into water as rainfall (PR_t) and snow (PS_t)

$$PR_t = \begin{cases} 0 & , T_t < T_{snow} \\ G_{input,t} \frac{T_t - T_{snow}}{T_{rain} - T_{snow}}, & T_{snow} \leq T_t \leq T_{rain} \\ G_{input,t} & , T_{rain} < T_t \end{cases} \quad (9)$$

$$PS_t = G_{input,t} - PR_t \quad (10)$$

where T_t is the average temperature in month t , T_{snow} and T_{rain} are the thresholds of snow and rainfall formation. The P above T_{rain} will be rainfall and the P below T_{snow} will be snow. Part of P will be snow when T_t falls into the range $[T_{snow}, T_{rain}]$. The difference between T_{rain} and T_{snow} is defined as a new parameter for estimation

$$\Delta T = T_{rain} - T_{snow}, \Delta T > 0 \quad (11)$$

Initial snow equivalent storage ($GS_{initial,t}$) in month t plus PS_t constitute the accumulated snow depth that releases snowmelt (SM_t)

$$SM_t = \begin{cases} 0 & , T_t < T_{snow} \\ (GS_{initial,t} + PS_t)m \frac{T_t - T_{snow}}{T_{rain} - T_{snow}} & , T_{snow} \leq T_t \leq T_{rain} \\ (GS_{initial,t} + PS_t)m & , T_{rain} < T_t \end{cases} \quad (12)$$

where parameter m is the snowmelt coefficient, representing the maximum fraction of snow storage that could melt when $T_t > T_{rain}$ (Hay and McCabe, 2002). Based on the water balance principle, the final snow water storage ($GS_{final,t}$) is

$$GS_{final,t} = GS_{initial,t} + PS_t - SM_t \quad (13)$$

The summation of PR_t and SM_t is used as the new input liquid water to the catchment in snow-dominant areas ($G_{snowinput,t}$).

$$G_{snowinput,t} = PR_t + SM_t \quad (14)$$

The remaining part of model_snow is the same as model_no except that $G_{input,t}$ is substituted by $G_{snowinput,t}$. The TWS in snow regions include both liquid water and snow storage, and $G_{mean,t}$ is

$$G_{mean,t} = \frac{1}{2} (G_{initial,t} + G_{final,t} + GS_{initial,t} + GS_{final,t}) \quad (15)$$

The objective function is to find the least squared error between GRACE-derived TWSA and $G_{mean,t}$ after the subtraction of $Diff$ as the model_no.

2.2.3. Linear correction

One difficulty in reconstructing GRACE-derived TWSA is the trend that is caused by human activities and ice melt (Li et al., 2021; Sun et al., 2020). Domestic water consumption, dam construction, agricultural irrigation, and ice melt have largely influenced TWSA in regions like the Sahara Desert, China, Mediterranean coast, India, and Greenland (Li et al., 2019b; Rodell et al., 2018; Tapley et al., 2019). However, accurate and physical quantification of their impacts at a global scale is challenging due to the limitation of related data. Fortunately, bridging the gap and data availability at both ends allow to infer the trend from temporal patterns. Thus, a time-dependent linear correction was applied to consider the trend near the gap. The observed TWSA ($TWSA_{obs}$) from July 2016 to May 2019 (excluding the gap from July 2017 to May 2018) were linearly regressed to the reconstructed TWSA ($TWSA_{rec}$) and year as follows

$$TWSA_{obs} = \alpha_0 + \alpha_1 TWSA_{rec} + \alpha_2 year \quad (16)$$

where observations in two years around the gap were used for calibrating the parameters. The component expressed as a function of year can be extended to the gap to quantify the trends relating to the factors that are not incorporated in the hydrological model, such as domestic water consumption, dam construction, and agricultural irrigation. As a result, the reconstructed TWSA in the gap period using hydrological models were linearly corrected using Eq. (16).

2.2.4. Parameters estimation

The model in this study was applied at the 0.5-degree resolution, which is the same as the resolution in GRACE dataset provided by JPL (Wiese et al., 2019). An "L-BFGS-B" method was used to estimate parameters (Byrd et al., 1995), with feasible bounds $\alpha_0 \in [-50, 50]$, $\alpha_1 \in [0, 2]$, $w \in [0, 5]$, $Max \in [0, 7000]$, $T_{rain} \in [-10, 10]$, $\Delta T \in [0, 50]$, and $m \in [0, 1]$ from previous hydrological studies (Gnann et al., 2019; Long et al., 2020; Martinez and Gupta, 2010; Zhang et al., 2020). The antecedent 52 months (January 1998 to April 2002) were used as a burn-in period for training $G_{initial,1}$ and $GS_{initial,1}$.

2.3. Evaluation metrics

Three metrics were used to access the reconstruction performance in calibration and validation periods, including the Nash-Sutcliffe efficiency (NSE; Nash and Sutcliffe (1970)), Pearson correlation coefficient (CC), and root mean square error (RMSE), representing the ratio of explained variance, linear correlation, and mean error between predicted and observed values, respectively. The closer the NSE and CC are to 1, the better the performance. The closer the RMSE to 0, the better the performance.

3. Reconstruction performance for an artificial gap

3.1. Performance evaluation

An artificial gap from July 2014 to May 2015 is assumed as validation period for evaluating the reconstruction performance of the hydrological model. The GRACE-derived TWSA from May 2002 to June 2017, excluding the artificial gap, are used for calibration. The structure of the artificial gap is similar to the real gap between GRACE and GRACE-FO, such that the reconstruction performance in the artificial gap can reflect the model accuracy for filling the real gap. Model_snow is only applied to the regions where monthly T is below 1 °C for at least 1 month in the study period. The areas where model_snow is not applied may not be largely affected by snow cover from remote sensing (Hall et al., 2002), and model_snow is the same as model_no.

The evaluation performance of the hydrological model in the calibration (May 2002 to June 2017 excluding the artificial gap) and validation (artificial gap from July 2014 to May 2015) periods are shown in Figs. 2 and 3, respectively. The performance in the calibration period is adequate in terms of all three metrics. NSE is larger than 0.5, and the CC is larger than 0.7 in most areas (Figs. 2a-c and 3a-c). In the validation period, model_no shows the best performance mainly in northern South America, eastern Northern America, central Africa, India, Europe, and southern China (Fig. 2d-i). Model_snow performed the best mainly in northern Asia, Europe, Canada, and central Antarctica (Fig. 3d-i). Snow accumulation is an important component in these regions and the incorporation of T indeed improves the performance compared with model_no. The difference between NSE and CC indicates that the hydrological model can predict the short-term variations of TWSA but may not be accurate for the long-term trend, which shows the necessity of correction for trends.

The linear correction, which is calibrated from July 2013 to May 2016 excluding the artificial gap, largely improves the NSE and RMSE (Figs. 2g-i, 3g-i, and S2), and reveals that considering the trend is necessary in order to improve the performance. Specifically, linear correction increases NSE by 0.3 at 59.9% and 48% of grids, and decreases RMSE by 3 mm at 67.0% and 49.1% of grids for model_no and model_snow, respectively. Meanwhile, the RMSE nearly remains the same for the grids with decreasing NSE, suggesting the robustness of linear correction (Fig. S2). The improvement in performance mainly occur in polar regions, North America, north Africa, eastern Asia, and southern Australia, where hydrological model does not perform well. Linear correction may supplement the hydrological processes, such as ice melt, human water consumption, and groundwater that were ignored in current hydrological model. A longer period was also tested for linear correction and the evaluation metrics are nearly the same (Figs. 2j-l and 3j-l). Thus, the trend has been sufficiently trained from two years around the gap.

The parameters for model_no and model_snow are shown in Figs. S3 and S4. From the values of α_0 and α_1 , ERA 5 reanalysis may underestimate E in high-latitudes and overestimate that in low- and mid-latitudes in TWSA analysis. Parameter w , representing runoff generation, ranges from 0 to 2, which are smaller than regional studies (Yang et al., 2008; Zhang et al., 2020). The reason may be that E has been removed before runoff generation, and thus the ratio of final water storage to storage capacity decreased. The Max is close to 2000 mm (i.e., initial values for parameters estimation), and may not play an important role in current model because anomalies are used and absolute values in TWS do not impact our results. Other initial values, including 1000 mm and 4000 mm, were also tested for Max , and the evaluation metrics are very similar. T_{rain} and ΔT are close to 0 and 6 °C, which are consistent with snow dynamics from regional studies (Hay and McCabe, 2002; Martinez and Gupta, 2010). Values of m demonstrate that snow is more likely to melt when $T_t > T_{rain}$ in high-latitudes.

The TWSA in arid climates is not well predicted in all models, such as central Australia, central contiguous United States, and Mongolia. The main reason may be that the GRACE product includes large uncertainty and shows a small signal-to-noise ratio in arid areas (Long et al., 2015; Sun et al., 2020). Other areas also demonstrate low levels of performance, such as eastern Asia, northern India, and southern South America. Uncertainty in E may

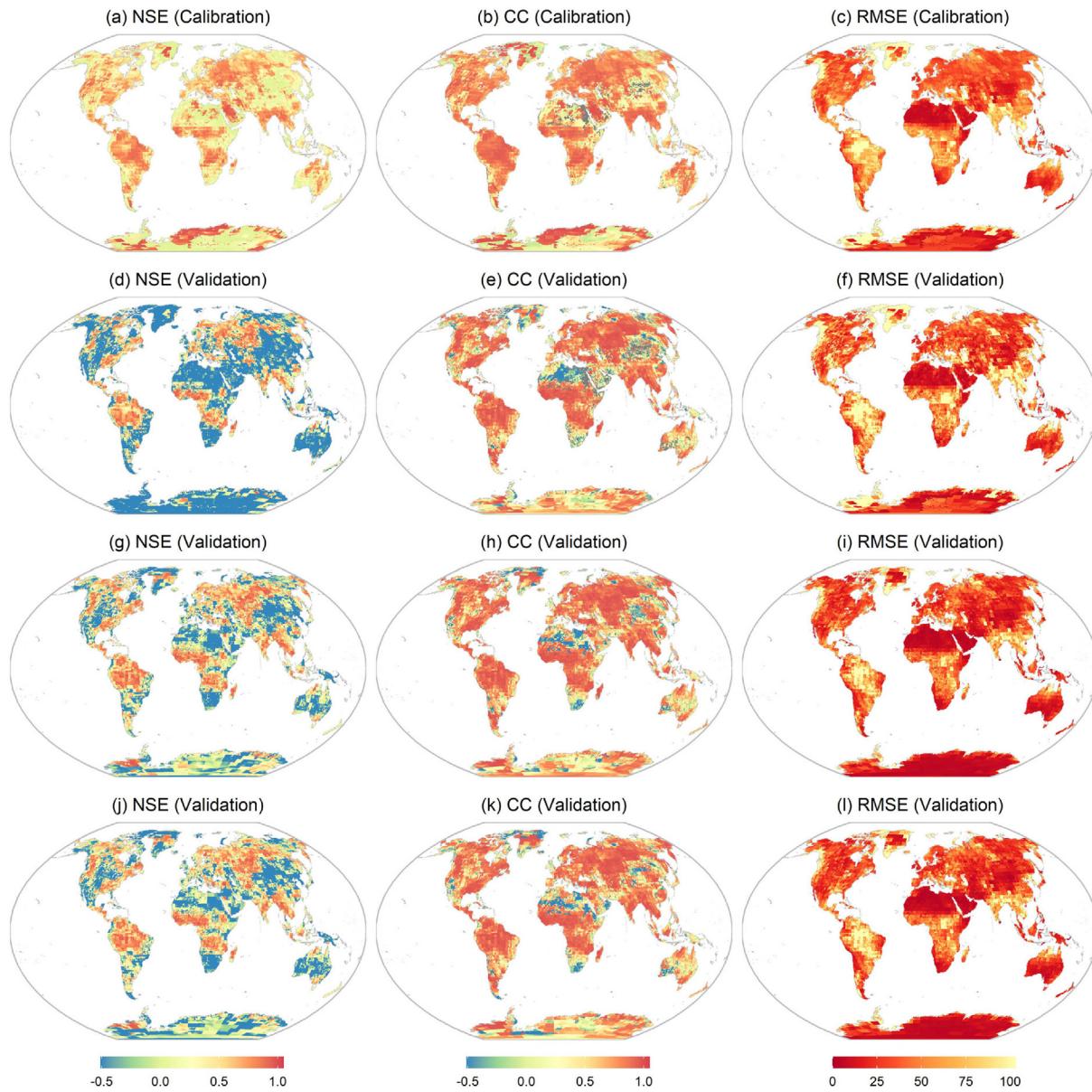


Fig. 2. The NSE, CC, and RMSE of model_no for the artificial gap in the calibration period (a-c), validation period before linear correction (d-f), validation period after linear correction from July 2013 to May 2016 (g-i), and linear correction from July 2012 to May 2017 (j-l).

account for the errors, because this study only uses reanalysis data and may include biases in estimation (Scanlon et al., 2018; Sridhar et al., 2019). Long-term drought may also cause abnormally low TWSA and bring difficulties for water storage simulation (Satish Kumar et al., 2021). Human activities such as dam construction, water diversion project, and agriculture irrigation, may anthropogenically change natural water movement and additionally increase the uncertainty for reconstructions in both arid and humid climates (Li et al., 2019b; Rodell et al., 2018; Zhang et al., 2019b).

3.2. Selection of the best model

The model with the largest NSE in the validation period was selected as the best one for bridging the gap (Fig. 4a, b). The results after linear correction were used because of the improvement in evaluation metrics and the incorporation of trends. In general, the selection is in accord with the properties of hydrological models. For tropical and temperate climates, model_no is mostly used. Model_snow is mainly used for high-latitude regions. The evaluation metrics indicate that the CC is larger than 0.5 in most areas. NSE are larger than 0.5 in low- and mid-latitudes, and larger

than 0 for many arid regions (Fig. 4c-e). However, the performance in regions around Greenland, central Australia, middle-eastern Asia, and southern and northern Africa deserves further improvements.

Time series of reconstructed TWSA from the best model in seven continents are shown in Fig. 5. The hydrological models successfully reconstructed the TWSA in Australia, Southern America, and Europe near the gap (Fig. 5c, e, and g). The linear correction corrects the trend around the gap and achieves good performance in the remaining continents. For example, the decreasing TWSA in Antarctica is mainly caused by ice melt in the region of Amundsen Sea Embayment (Mémin et al., 2015; Tapley et al., 2019). The linear correction uses a time-dependent component to represent ice melt and thus simulate TWSA in Antarctica with NSE = 0.91 (Fig. 5f). The performance in seven continents indicates that our reconstructions are available for large-scale analysis.

3.3. Comparison with previous reconstructions

The performance of our hydrological model is similar to previous global reconstructions using the AI (Li et al., 2021; Sun et al., 2020). The Pearson

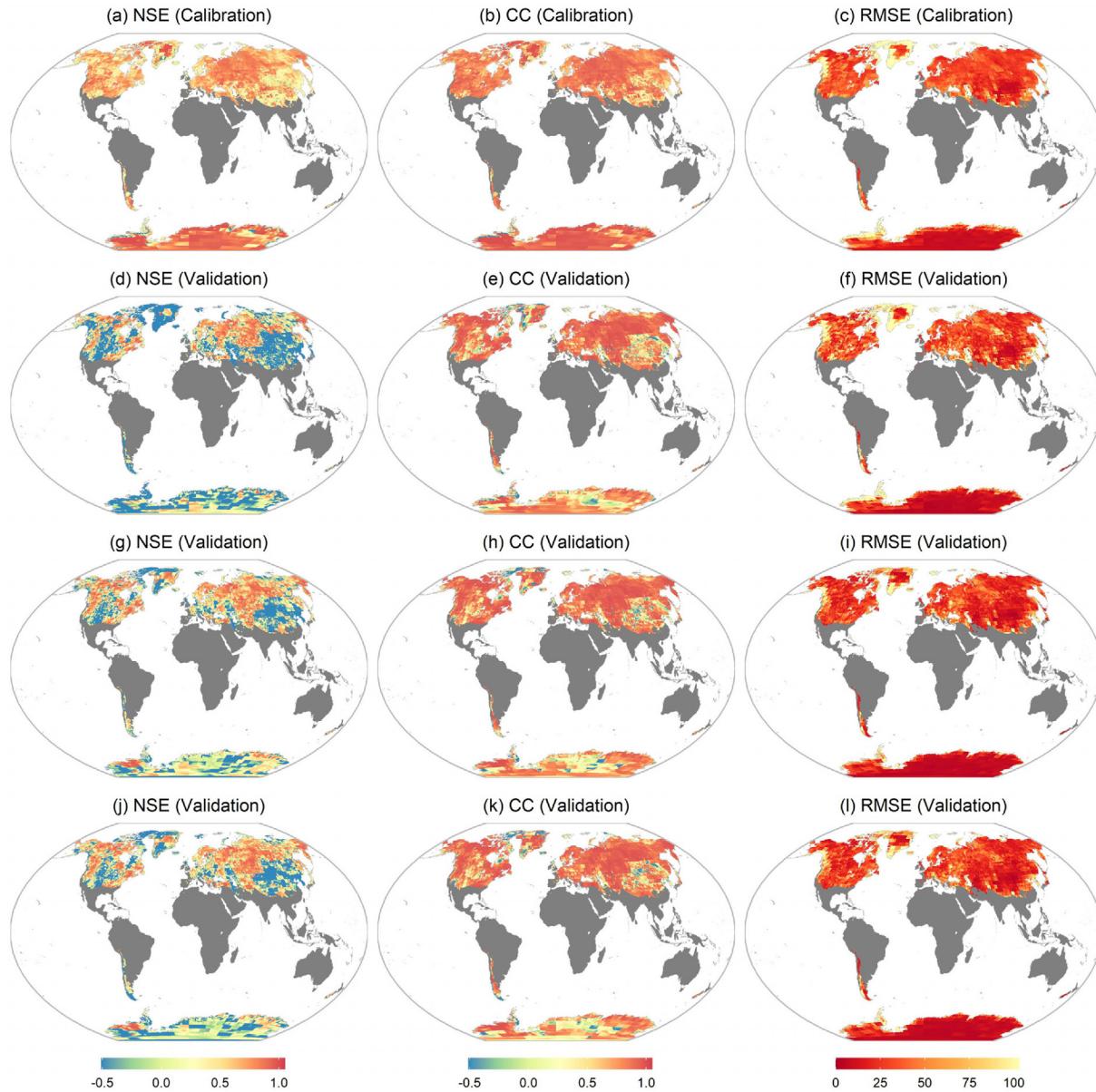


Fig. 3. The NSE, CC, and RMSE of model_snow for the artificial gap in the calibration period (a-c), validation period before linear correction (d-f), validation period after linear correction from July 2013 to May 2016 (g-i), and linear correction from July 2012 to May 2017 (j-l).

correlation coefficients are high in nearly all reconstructions, indicating that the high-frequency variations can be sufficiently expressed from climatic variables (Humphrey and Gudmundsson, 2019; Humphrey et al., 2017; Li et al., 2021; Sun et al., 2020). The NSE of hydrological models is better than the results from multiple linear regression models, but slightly worse than the results from the AI models in the validation period (Sun et al., 2020). About 62.5% of points achieve positive NSE in this study, while they are approximately 75% and 40% using deep neural networks and multiple linear regression, respectively (Sun et al., 2020). The RMSE are as good as the results from machine learning with decomposition techniques in the validation period, with RMSE in ~80% grids less than 47 mm (Li et al., 2021). The reconstruction using a statistical model does not provide evaluation metrics in the validation period, but their performance is very similar to current results in the calibration period (Humphrey and Gudmundsson, 2019). All reconstructions indicate that the arid climates, Greenland, and the Gulf of Alaska are the regions with poor performance, and ice melt, small signal-to-noise ratio, and human activities may account for them (Rodell et al., 2018). Overall, our hydrological model can largely

reproduce the performance from the AI and statistical models while maintaining the physical foundations.

4. Reconstruction for the gap

4.1. Bridging the gap

The gap between GRACE and GRACE-FO was filled using the best model that was selected in Section 3.2 and linearly corrected from July 2016 to May 2019. The performance in the calibration period is shown in Fig. 6. The evaluation metrics are similar to those in the artificial gap, suggesting their similar performance in the validation period. The time series of TWSA in seven continents are shown in Fig. 7. The hydrological models successfully reconstructed the TWSA in Australia and Southern America near the gap, and improves the performance in Northern America, Antarctica, and Europe after linear correction. However, the performance in Africa and Asia deserves further improvement due to the abnormal TWSA that is possibly caused by human activities.

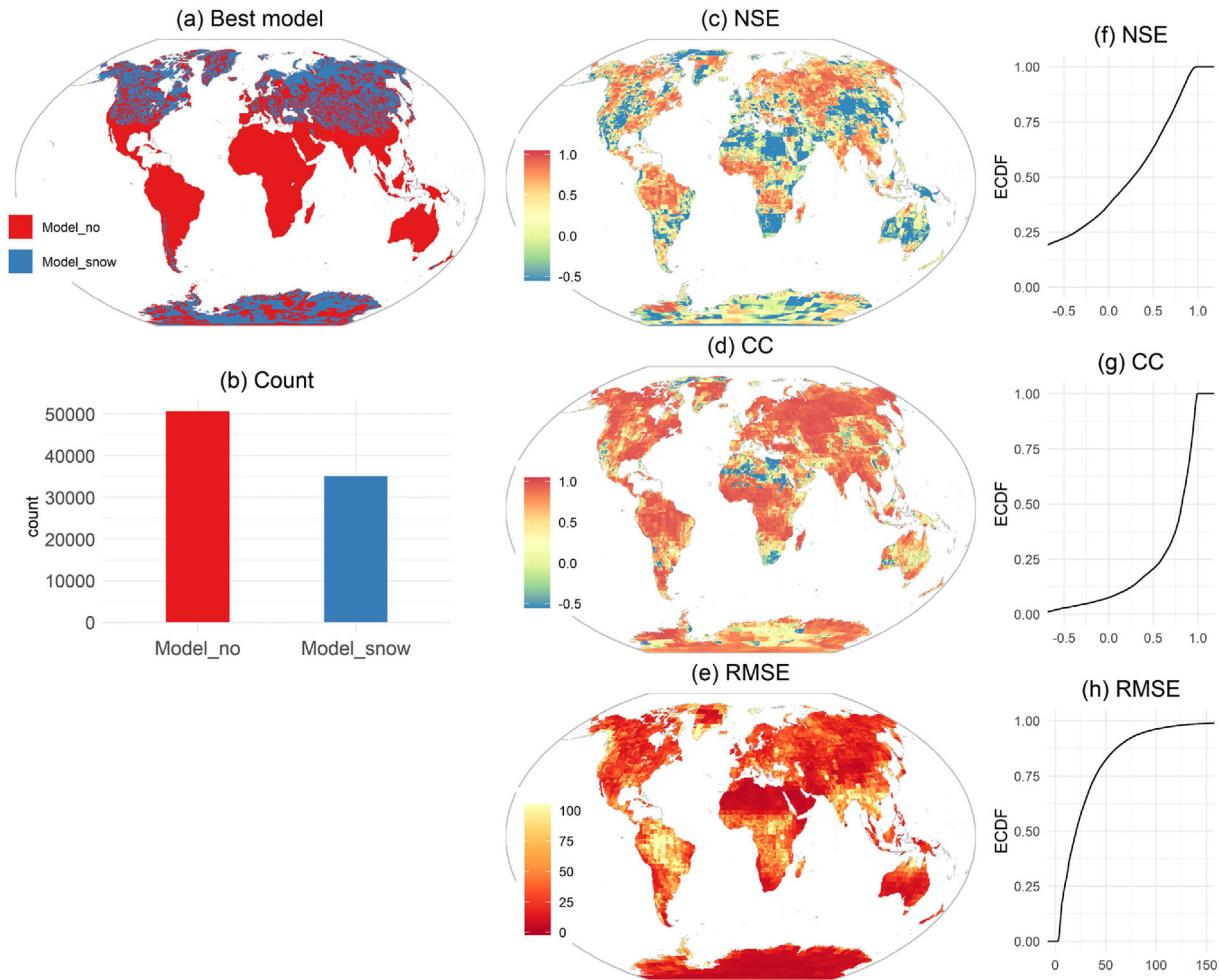


Fig. 4. Selection of the best model based on the NSE in the validation period. (a-b) The best model selected and its count. (c-e) Evaluation metrics of the best model for filling the artificial gap from July 2014 to May 2015. (f-h) Empirical Cumulative Distribution Function (ECDF) of the evaluation metrics.

4.2. Evaluation with sea-level budget

Since the total volume of water on Earth can be viewed as a constant, the changes in TWSA can be oppositely observed in ocean mass (Humphrey and Gudmundsson, 2019). Meanwhile, sea-level budget estimates ocean mass as the difference between global mean sea level and steric expansion, and thus provides an independent way for evaluating the reconstructed TWSA in the gap period (Dieng et al., 2017; Humphrey and Gudmundsson, 2019; Wang et al., 2018a). Here, we used sea level measurements from satellite altimeters (Beckley et al., 2017) and steric expansion from Argo observations (Roemmich and Gilson, 2009), and compared the ocean mass derived from sea-level budget with that from TWSA (Fig. 8). In the past decades, decrease in TWSA accounts for main portions of sea level rise (Jacob et al., 2012; Reager et al., 2016). Both ice melt and decline in endorheic basin water storages contribute to ocean mass increase (Wang et al., 2018a). The abnormal ocean mass decrease in 2010 and 2011 may be attributed to the strong La Niña event, which affects global precipitation and increases TWSA at some continents (Boening et al., 2012; Cazenave et al., 2012). From 2010 to 2020, ocean mass derived from TWSA largely follow the trend that is independently obtained from sea level budget, indicating that TWSA in the gap period should also follow the sea level budget if our reconstruction is reliable. In the gap period, the ocean mass inferred from TWSA decreased in the first four months, and then increased and reached its maximum in January 2018 (Fig. 8). The ocean mass derived from altimetry method largely reproduces this pattern and reached a similar magnitude, showing the fidelity of TWSA reconstruction at both global and decadal scales.

4.3. Evaluation with the TWSA in GLDAS-Noah

The TWSA in the GLDAS-Noah model has a good correlation with the GRACE measurements (Long et al., 2014; Syed et al., 2008; Zhang et al., 2019d), which has been used as a predictor for simulation (Jing et al., 2020; Sun et al., 2020). Although some discrepancies are found between GLDAS and GRACE (Scanlon et al., 2018), the GLDAS-Noah data can partly verify our reconstruction. The Pearson correlation coefficients between GLDAS-derived TWSA and GRACE-derived TWSA are shown in Fig. 9. The GLDAS-derived TWSA includes all layers of soil moisture, snow water equivalent, and canopy interception (Rodell et al., 2004; Sun et al., 2020). In three periods with observations, the GLDAS-derived TWSA reproduces the variations in GRACE-derived TWSA in most areas, while TWSA in Greenland, northern and southern Africa, and eastern Asia was not well simulated (Fig. 9a-c). Human activities largely decrease water storage in northern Africa and eastern Asia (Rodell et al., 2018), while GLDAS-Noah does not incorporate these perturbations. Additionally, ice melt and its coupling with atmosphere and ocean are not fully included in GLDAS-Noah, resulting in low correlation in Greenland. The reconstruction for the gap, to a large extent, shows similar spatial patterns of correlation, indicating the reliability of our reconstruction at gridded resolution and on interannual timescales (Fig. 9d).

5. Discussion

The gap between GRACE and GRACE-FO products was reconstructed based on an empirical model that attempts to simulate hydrological processes

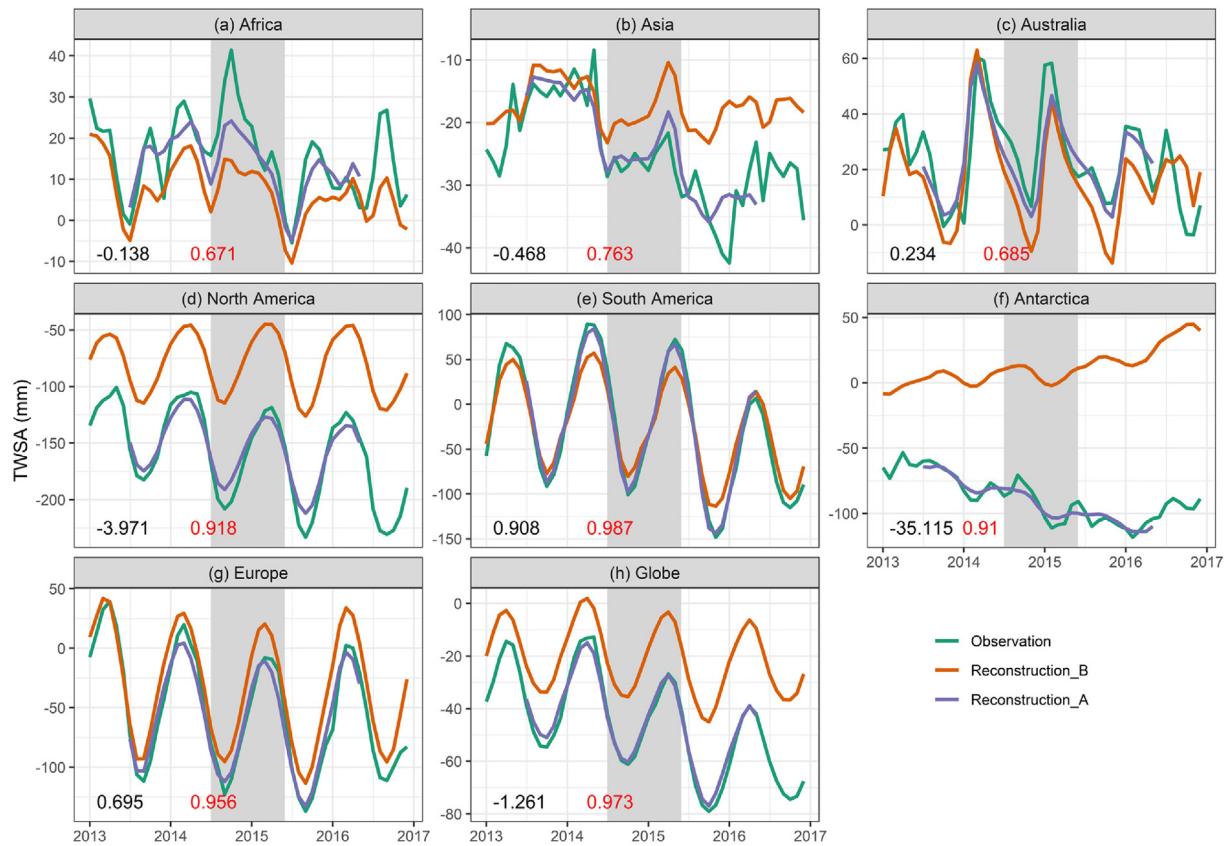


Fig. 5. Time series of observed and reconstructed TWSA from the best model before (Reconstruction_B) and after (Reconstruction_A) linear correction in seven continents and globe. The gray area denotes the artificial period (July 2014 to May 2015). The numbers at the bottom left are the NSE before (black) and after (red) linear correction.

in no-snow and snow-covered regions separately. Compared with previous studies that employed the AI or pure statistical models (Humphrey and Gudmundsson, 2019; Humphrey et al., 2017; Li et al., 2021; Sun et al., 2020), the main advantage of our approach is that physical dynamics are explicitly incorporated, meanwhile our model achieved similar performance at gridded, continental, and global scales. Current GRACE and GRACE-FO products only cover a period of less than 20 years, and many extreme TWSA events, such as 1998 floods in China, did not occur in the observational period and thus are not sufficiently calibrated in AI models. Thus, the error-

correction learning algorithm in AI models is likely to bear large uncertainties for extreme events (Chollet, 2018). The physically-based hydrological models, in contrast, are more robust under such circumstances. In the future, more physically-based hydrological models may be developed and calibrated against GRACE-based TWSA, so as to explore the possibility of reconstructing or predicting TWSA.

Another advantage of this study is that the raw measurements, rather than the de-trend and de-seasonalized time series, were reconstructed, which leads to the direct combination with original GRACE-based observations. Previous

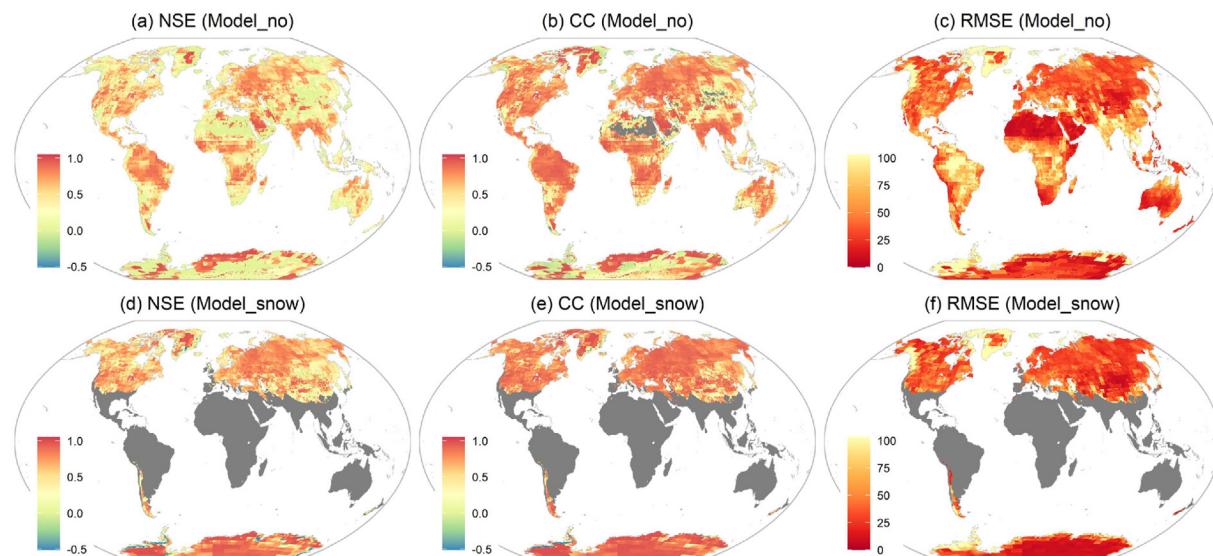


Fig. 6. The NSE, CC, and RMSE of model_no (a-c) and model_snow (d-f) for the real gap in the calibration period.

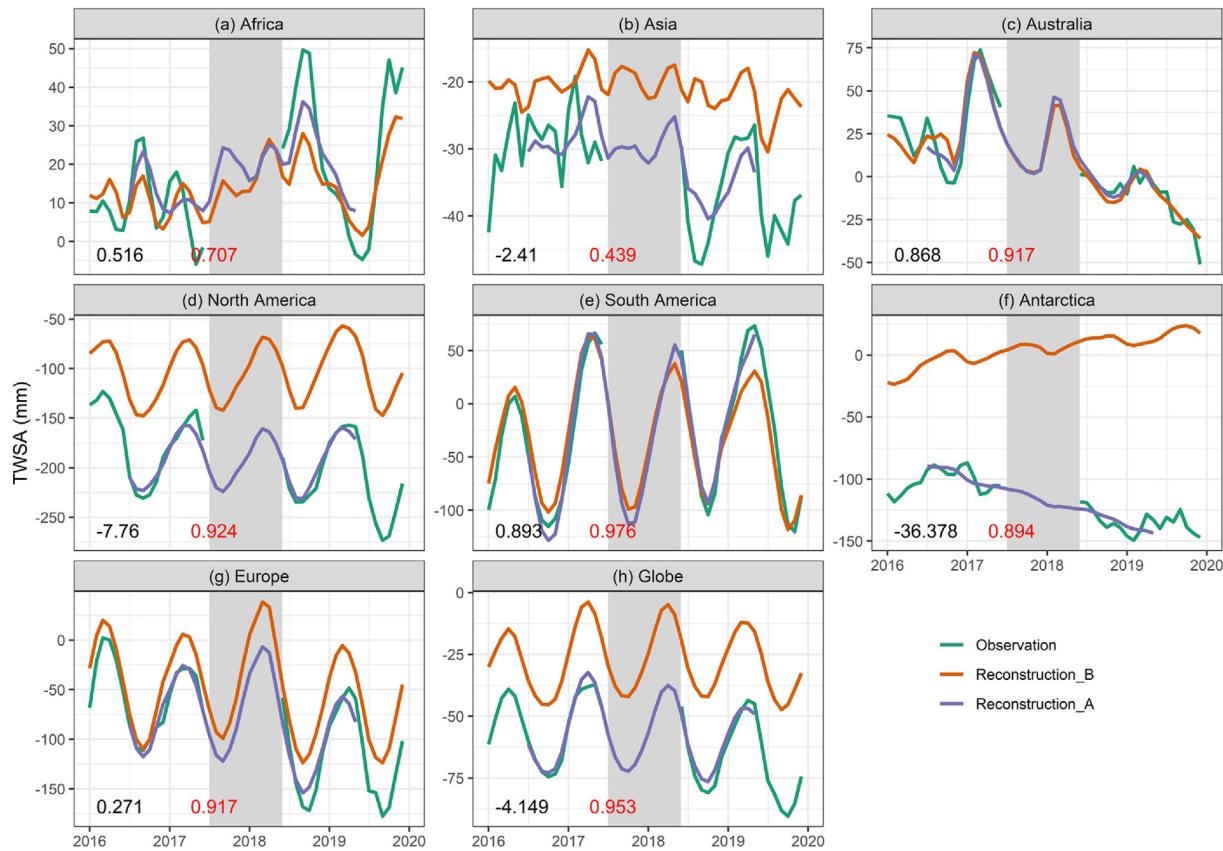


Fig. 7. Time series of observed and reconstructed TWSA before (Reconstruction_B) and after (Reconstruction_A) linear correction in seven continents and globe. The gray area denotes the gap period from July 2017 to May 2018. The numbers at the bottom left are the NSE before (black) and after (red) linear correction.

studies mainly reconstructed and evaluated the performance of de-detrended and de-seasonalized TWSA (Humphrey and Gudmundsson, 2019; Jing et al., 2020; Li et al., 2021). However, the de-detrended transformation lacks of physical justifications because many long-term trends in TWSA are caused by anthropogenic climate change and/or natural environmental factors. In addition, the performance in the validation period after adding the trend

was not rigorously evaluated (Humphrey and Gudmundsson, 2019; Jing et al., 2020; Li et al., 2021), which will cause uncertainties if the reconstructions are directly combined with GRACE and GRACE-FO products. Our results show that the negligence of trend may lead to large errors even when the monthly variations have been well predicted (Fig. 5). Simply extending the trend in the calibration period to the validation period may not be

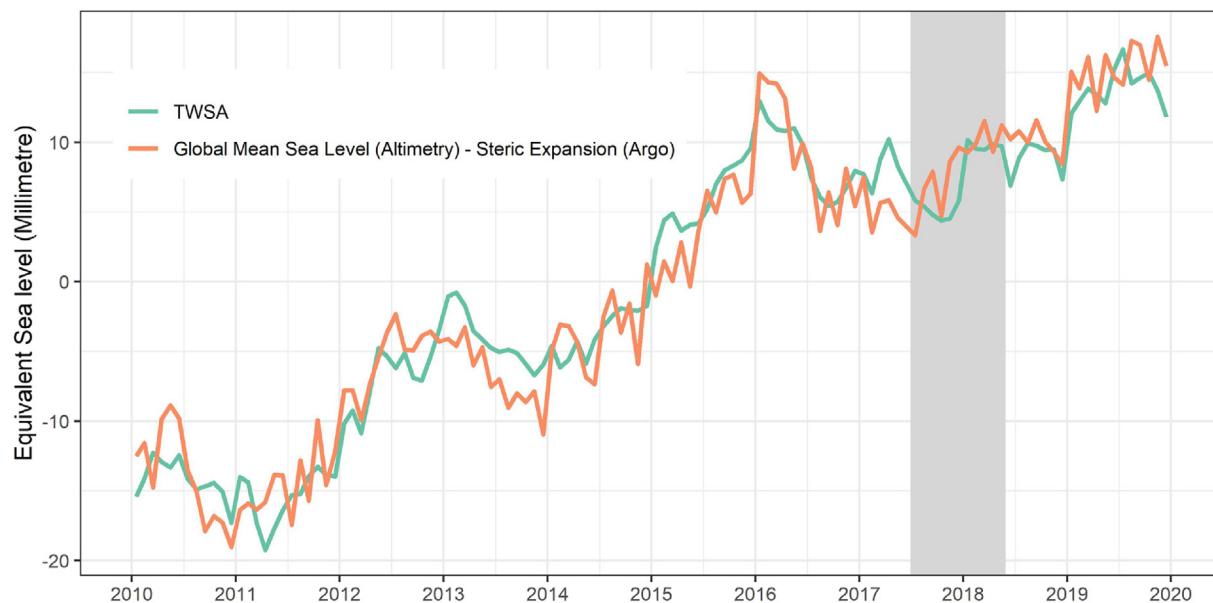


Fig. 8. Comparison of ocean mass derived from TWSA and sea level budget as the difference between global mean sea level height and steric expansion. The values are de-seasonalized and expressed as equivalent sea level in millimeter. The gray area denotes the gap period from July 2017 to May 2018 during which the TWSA is from our reconstruction.

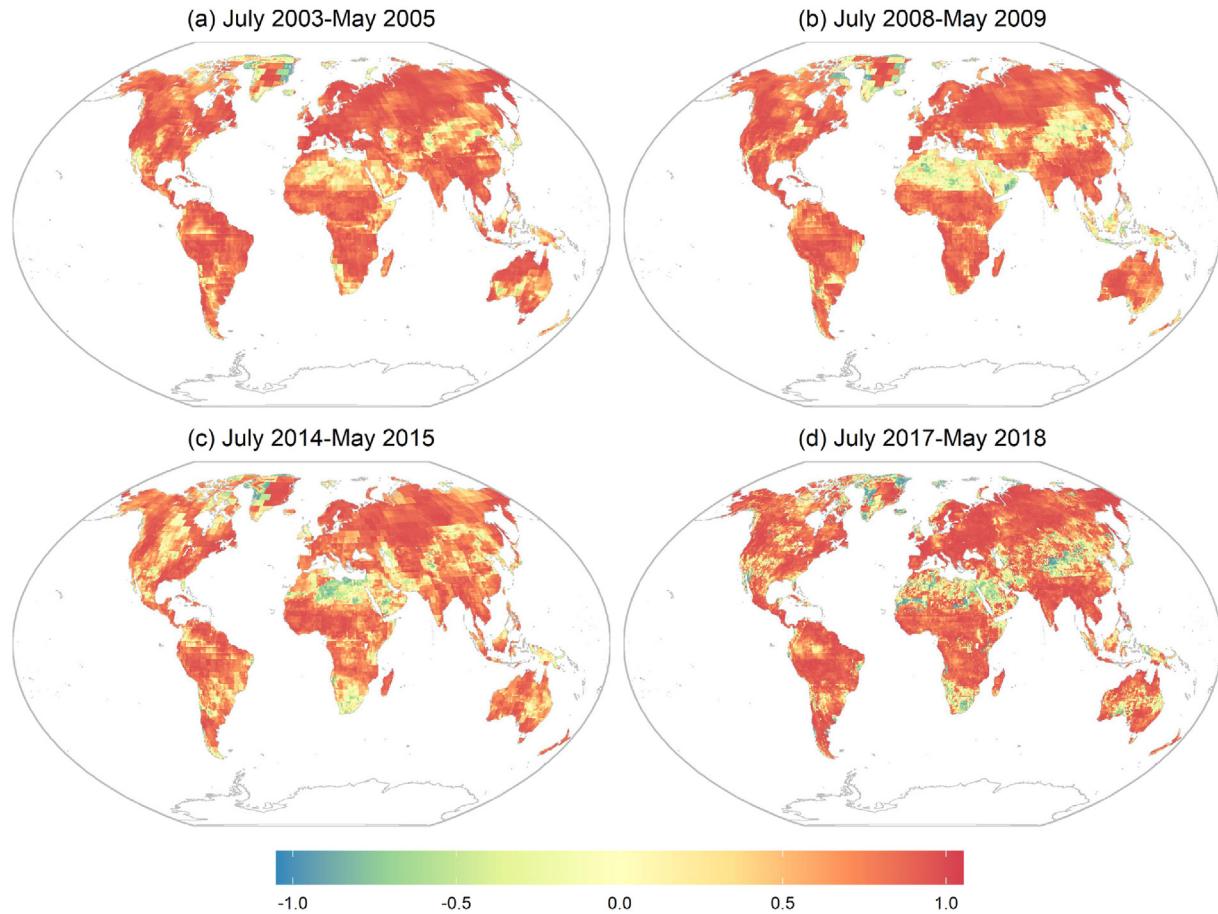


Fig. 9. The Pearson correlation coefficients between GLDAS-derived TWSA and GRACE-derived TWSA in three periods with observed GRACE data (a-c) and the gap period (d).

appropriate because natural factors and human activities may not persistently impact TWSA to the same extent for over 20 years, not to mention till 1979 (Li et al., 2021). Therefore, the reconstruction of TWSA without de-trended and de-seasonalized processes is required for increasing the physical foundations and joint applications of GRACE and GRACE-FO products.

The evaluation for an artificial gap indicates the accuracy and feasibility of our model reconstruction. In general, our model performs well for tropical, temperate, and continental climates where the results show high correlations and NSE values with the measurements. However, our model deserves improvement in central Australia, Sahara Desert, central America, and eastern Asia. Previous AI models show similar drawback in simulating TWSA in these areas (Li et al., 2021; Sun et al., 2020). One major reason is because many factors that largely influence TWSA, such as human water consumption, population, agricultural irrigation, economic development, are not well incorporated in the models (Ahmed et al., 2014; Ramillien et al., 2014; Rodell et al., 2018). For example, the South-North Water Diversion Project and the strict water resources management system in China have changed the groundwater dynamics in eastern Asia (Long et al., 2020; Wang et al., 2018b; Xu et al., 2021), and current hydrological model without these components is unable to accurately simulate TWSA in these areas. Another major reason may be related to the small signal-to-noise ratio in arid climates (Long et al., 2015; Sun et al., 2020). TWSA in Greenland and Antarctic are also not well reconstructed because ice melt is an element that was not incorporated in our hydrological model. Nonetheless, the linear correction helps improve the reconstruction of the trend. Our reconstruction is correlated well with the sea-level budget (Lumpkin et al., 2020) and GLDAS-derived TWSA (Rodell et al., 2004), which validates its overall accuracy.

The rationale behind our hydrological model is similar to some conceptual hydrological models, such as the “abcd” model (Thomas, 1981,

temperature-based snow dynamics (Hay and McCabe, 2002; Martinez and Gupta, 2010; Woods, 2009), and the Budyko-type equations (Fu, 1981; Yang et al., 2008). However, other hydrological components still deserve further attention in order to improve current models. For example, the linear correction method can only be used for filling a gap but cannot be extended into the future or the past decades without known ends. Groundwater accounts for a large portion of TWSA variations but is ignored in current hydrological model. This ignorance brings uncertainty for parameters calibration, and hinders us to incorporate groundwater depletion or replenishment during drought or flood periods (Reager et al., 2014; Zhang et al., 2019a; Zhang et al., 2015). Human water consumption and agriculture irrigation also influence TWSA through groundwater (Felfelani et al., 2017; Xie et al., 2019; Zhang et al., 2019c; Zhang et al., 2020), while current hydrological model simply quantifies this process from surface water movement and linear correction. Future studies may explore the role of groundwater in TWSA variations and further incorporate groundwater movement when groundwater observations become more available (Moeck et al., 2020). Apart from that, applications of hydrological model determine that our reconstruction is dependent on hydroclimatic information, and it is thus inappropriate for precipitation reconstruction, evaporation estimation, and data assimilation in hydrological models (Behrangi et al., 2017; Houborg et al., 2012; Li et al., 2019a; Werth et al., 2009). Nonetheless, it can be applied in the fields that do not require independence, such as drought monitoring (Houborg et al., 2012; Sridhar et al., 2019), flood prediction (Reager et al., 2014), sea level budget (Reager et al., 2016), and detection of human activities (Rodell et al., 2018; Xu et al., 2021). Meanwhile, most of them call for continuous analysis of soil moisture, drought propagation, and flood inundation based on antecedent variables, highlighting the necessity of current reconstruction (Houborg et al., 2012; Reager et al., 2016; Reager et al., 2014; Sridhar et al., 2019).

Various sources of uncertainties are not assessed in this study, despite that they are crucial in global hydrological studies. GRACE and meteorological data include uncertainties because of the measurement errors and processing algorithms. For example, the selection of mascons or spherical harmonics (SH) solutions in GRACE products brings differences for reconstruction (Sun et al., 2020). The estimation of E is also challenging as there are not many independent measurements for correction (Zhang et al., 2018). For GRACE products, the short period of observations, signal leakage errors, human perturbations, and the difference between TWSA and TWS may further increase the uncertainties in parameters estimation (Landerer and Swenson, 2012; Sun et al., 2020). The uncertainties in GRACE also have spatial autocorrelation, which makes it difficult for simulation under an uncertain framework (Humphrey and Gudmundsson, 2019). Systematically quantifying the uncertainties in hydroclimatic variables, model structure and parameters, and GRACE products will help improve the hydrological models and the TWSA simulations.

6. Conclusions

In this study, a hydrological model was explicitly developed to reconstruct TWSA for the gap period between GRACE and GRACE-FO products from July 2017 to May 2018. The atmospheric drivers (precipitation and evaporation), hydrological conditions (maximum water storage), and snow dynamics were incorporated into the model through a combination of “abcd” model, Budyko hypothesis, and a temperature-based snow component. The long-term trend that may be attributed to human activities is simulated from a linear correction with values available at both ends of the gap. The performance of the hydrological model was rigorously tested through an artificial gap, and the reconstruction for the real gap was independently validated against the sea-level budget and GLDAS-derived TWSA. Our hydrological model explicitly incorporates physical processes and uses original data without detrending, which make the TWSA reconstruction consistent with the original GRACE and GRACE-FO products. Our study shows the potential of developing hydrological models to simulate GRACE-based TWSA, which will advance the understanding of the forcing mechanisms of TWSA variations and the application of GRACE products in the related disciplines.

Despite the rigorous evaluation of current reconstruction, this study ignores human activities and many natural hydrological processes such as groundwater, ocean current, reservoir operation, agricultural irrigation and domestic water consumption. The performance in most dry and polar climates also deserves further improvement. The linear correction for simulating the trend is appropriate for filling a gap, but is not suitable for projecting or reconstructing TWSA over a decadal timescale. Future study may further investigate the roles of natural hydrological processes and human activities in TWSA variations, and establish an improved global hydrological model for modeling TWSA with GRACE products.

CRediT authorship contribution statement

Xu Zhang: Conceptualization, Methodology, Investigation, Writing – original draft. **Jinbao Li:** Writing – review & editing, Visualization, Investigation, Funding acquisition. **Qianjin Dong:** Writing – review & editing, Investigation, Validation. **Zifeng Wang:** Writing – review & editing, Validation. **Han Zhang:** Writing – review & editing, Validation. **Xiaofeng Liu:** Writing – review & editing, Validation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was financially supported by the National Key Research and Development Program of China (No. 2018YFA0605601) and Hong

Kong Research Grants Council (No. 17303017). The reconstruction data are available at <https://github.com/870846704/Bridging-the-gap-between-GRACE-and-GRACE-FO-using-a-hydrological-model>. The GRACE dataset is accessed at <https://grace.jpl.nasa.gov/>. The ERA5 dataset is accessed at <https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5>. The sea level budget data for validation are accessed at <https://sealevel.nasa.gov/understanding-sea-level/overview>.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.scitotenv.2022.153659>.

References

- Ahmed, M., Sultan, M., Wahr, J., et al., 2014. The use of GRACE data to monitor natural and anthropogenic induced variations in water availability across Africa. *Earth Sci. Rev.* 136, 289–300.
- Ahmed, M., Sultan, M., Elbayoumi, T., et al., 2019. Forecasting GRACE data over the african watersheds using artificial neural networks. *Remote Sens.* 11 (15), 1769.
- Beckley, B.D., Callahan, P.S., Hancock III, D.W., et al., 2017. On the “Cal-mode” correction to TOPEX satellite altimetry and its effect on the global mean sea level time series. *J. Geophys. Res. Oceans* 122 (11), 8371–8384.
- Behrangi, A., Gardner, A.S., Reager, J.T., et al., 2017. Using GRACE to constrain precipitation amount over cold mountainous basins. *Geophys. Res. Lett.* 44 (1), 219–227.
- Boening, C., Willis, J.K., Landerer, F.W., et al., 2012. The 2011 La Niña: so strong, the oceans fell. *Geophys. Res. Lett.* 39 (19), L19602.
- Budyko, M.I., 1974. *Climate and Life*. Academic Press, New York.
- Byrd, R.H., Lu, P., Nocedal, J., et al., 1995. A limited memory algorithm for bound constrained optimization. *SIAM J. Sci. Comput.* 16 (5), 1190–1208.
- Cazenave, A., Henry, O., Munier, S., et al., 2012. Estimating ENSO influence on the global mean sea level, 1993–2010. *Mar. Geod.* 35 (sup1), 82–97.
- Chen, H., Zhang, W., Nie, N., et al., 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.
- Chen, J., Tapley, B., Rodell, M., et al., 2020. Basin-Scale River runoff estimation from GRACE gravity satellites, climate models, and in situ observations: a case study in the Amazon Basin. *Water Resour. Res.* 56 (10), e2020WR028032.
- Chollet, F., 2018. *Deep Learning With Python*. Manning, New York.
- Deng, H., Chen, Y., 2017. Influences of recent climate change and human activities on water storage variations in Central Asia. *J. Hydrol.* 544, 46–57.
- Dieng, H.B., Cazenave, A., Meyssignac, B., et al., 2017. New estimate of the current rate of sea level rise from a sea level budget approach. *Geophys. Res. Lett.* 44 (8), 3744–3751.
- Eicker, A., Jensen, L., Wöhnke, V., et al., 2020. Daily GRACE satellite data evaluate short-term hydro-meteorological fluxes from global atmospheric reanalyses. *Sci. Rep.* 10 (1), 4504.
- Fatolazadeh, F., Goita, K., 2021. Mapping terrestrial water storage changes in Canada using GRACE and GRACE-FO. *Sci. Total Environ.* 779, 146435.
- Felfelani, F., Wada, Y., Longuevergne, L., et al., 2017. Natural and human-induced terrestrial water storage change: a global analysis using hydrological models and GRACE. *J. Hydrol.* 553, 105–118.
- Forootan, E., Khaki, M., Schumacher, M., et al., 2019. Understanding the global hydrological droughts of 2003–2016 and their relationships with teleconnections. *Sci. Total Environ.* 650, 2587–2604.
- Fu, B., 1981. On the calculation of the evaporation from land surface. *Chin. J. Atmos. Sci.* 5 (1), 23–31.
- Gaughan, A.E., Waylen, P.R., 2012. Spatial and temporal precipitation variability in the Okavango-Kwando-Zambezi catchment, southern Africa. *J. Arid Environ.* 82, 19–30.
- Gnann, S.J., Woods, R.A., Howden, N.J.K., 2019. Is there a baseflow Budyko curve? *Water Resour. Res.* 55 (4), 17.
- Hall, D.K., Riggs, G.A., Salomonson, V.V., et al., 2002. MODIS snow-cover products. *Remote Sens. Environ.* 83 (1), 181–194.
- Hay, L.E., McCabe, G.J., 2002. Spatial variability in water-balance model performance in the conterminous United States. *J. Am. Water Resour. Assoc.* 38 (3), 847–860.
- Hersbach, H., Bell, B., Berrisford, P., et al., 2020. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* 146 (730), 1999–2049.
- Houborg, R., Rodell, M., Li, B., et al., 2012. Drought indicators based on model-assimilated Gravity Recovery and Climate Experiment (GRACE) terrestrial water storage observations. *Water Resour. Res.* 48 (7), W07525.
- Humphrey, V., Gudmundsson, L., 2019. GRACE-REC: a reconstruction of climate-driven water storage changes over the last century. *Earth Syst. Sci. Data* 11 (3), 1153–1170.
- Humphrey, V., Gudmundsson, L., Seneviratne, S.I., 2017. A global reconstruction of climate-driven subdecadal water storage variability. *Geophys. Res. Lett.* 44 (5), 2300–2309.
- Jacob, T., Wahr, J., Pfeffer, W.T., et al., 2012. Recent contributions of glaciers and ice caps to sea level rise. *Nature* 482 (7386), 514–518.
- Jing, W., Zhang, P., Zhao, X., et al., 2020. Extending GRACE terrestrial water storage anomalies by combining the random forest regression and a spatially moving window structure. *J. Hydrol.* 590, 125239.
- Jones, P.W., 1999. First-and second-order conservative remapping schemes for grids in spherical coordinates. *Mon. Weather Rev.* 127 (9), 2204–2210.
- Kornfeld, R.P., Arnold, B.W., Gross, M.A., et al., 2019. GRACE-FO: the gravity recovery and climate experiment follow-on mission. *J. Spacecr. Rocket.* 56 (3), 931–951.

- Landerer, F.W., Swenson, S.C., 2012. Accuracy of scaled GRACE terrestrial water storage estimates. *Water Resour. Res.* 48 (4), W04531.
- Landerer, F.W., Flechtner, F.M., Save, H., et al., 2020. Extending the global mass change data record: GRACE follow-on instrument and science data performance. *Geophys. Res. Lett.* 47 (12) (e2020GL088306).
- Li, X., Gao, X., Chang, Y., et al., 2017. Water storage variations and their relation to climate factors over Central Asia and surrounding areas over 30 years. *Water Supply* 18 (5), 1564–1580.
- Li, B., Rodell, M., Kumar, S., et al., 2019a. Global GRACE data assimilation for groundwater and drought monitoring: advances and challenges. *Water Resour. Res.* 55 (9), 7564–7586.
- Li, J., Xie, S.-P., Cook, E.R., et al., 2019b. Deciphering human contributions to Yellow River flow reductions and downstream drying using centuries-long tree ring records. *Geophys. Res. Lett.* 46 (2), 898–905.
- Li, F., Kusche, J., Rietbroek, R., et al., 2020. Comparison of data-driven techniques to reconstruct (1992–2002) and predict (2017–2018) GRACE-like gridded total water storage changes using climate inputs. *Water Resour. Res.* 56 (5), e2019WR026551.
- Li, F., Kusche, J., Chao, N., et al., 2021. Long-term (1979–present) total water storage anomalies over the global land derived by reconstructing GRACE data. 48 (8), e2021GL093492.
- Löcher, A., Kusche, J., 2020. A hybrid approach for recovering high-resolution temporal gravity fields from satellite laser ranging. *J. Geod.* 95 (1), 6.
- Long, D., Shen, Y., Sun, A., et al., 2014. Drought and flood monitoring for a large karst plateau in Southwest China using extended GRACE data. *Remote Sens. Environ.* 155, 145–160.
- Long, D., Longuevergne, L., Scanlon, B.R., 2015. Global analysis of approaches for deriving total water storage changes from GRACE satellites. *Water Resour. Res.* 51 (4), 2574–2594.
- Long, D., Yang, W., Scanlon, B.R., et al., 2020. South-to-north water diversion stabilizing Beijing's groundwater levels. *Nat. Commun.* 11 (1), 3665.
- Lumpkin, R., Baringer, M., Bif, M.B., et al., 2020. Global oceans [in "state of the climate in 2019"]. *Bull. Am. Meteorol. Soc.* 101 (8), S129–S184.
- Luthcke, S.B., Zwally, H.J., Abdalati, W., et al., 2006. Recent Greenland ice mass loss by drainage system from satellite gravity observations. *Science* 314 (5803), 1286–1289.
- Martinez, G.F., Gupta, H.V., 2010. Toward improved identification of hydrological models: a diagnostic evaluation of the "abcd" monthly water balance model for the conterminous United States. *Water Resour. Res.* 46 (8), W08507.
- Mehrnegar, N., Jones, O., Singer, M.B., et al., 2021. Exploring groundwater and soil water storage changes across the CONUS at 12.5 km resolution by a Bayesian integration of GRACE data into W3RA. *Sci. Total Environ.* 758, 143579.
- Mémin, A., Flament, T., Alizier, B., et al., 2015. Interannual variation of the Antarctic Ice Sheet from a combined analysis of satellite gravimetry and altimetry data. *Earth Planet. Sci. Lett.* 422, 150–156.
- Moeck, C., Grech-Cumbo, N., Podgorski, J., et al., 2020. A global-scale dataset of direct natural groundwater recharge rates: a review of variables, processes and relationships. *Sci. Total Environ.* 717, 137042.
- Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I — a discussion of principles. *J. Hydrol.* 10 (3), 282–290.
- Ramillien, G., Frappart, F., Seoane, L., 2014. Application of the regional water mass variations from GRACE satellite gravimetry to large-scale water management in Africa. *Remote Sens.* 6 (8), 7379–7405.
- Reager, J.T., Thomas, B.F., Famiglietti, J.S., 2014. River basin flood potential inferred using GRACE gravity observations at several months lead time. *Nat. Geosci.* 7 (8), 588–592.
- Reager, J.T., Gardner, A.S., Famiglietti, J.S., et al., 2016. A decade of sea level rise slowed by climate-driven hydrology. *Science* 351 (6274), 699–703.
- Richter, H.M.P., Lück, C., Klos, A., et al., 2021. Reconstructing GRACE-type time-variable gravity from the Swarm satellites. *Sci. Rep.* 11 (1), 1117.
- Rodell, M., Houser, P.R., Jambor, U., et al., 2004. The global land data assimilation system. *Bull. Am. Meteorol. Soc.* 85 (3), 381–394.
- Rodell, M., Famiglietti, J.S., Wiese, D.N., et al., 2018. Emerging trends in global freshwater availability. *Nature* 557 (7707), 651–659.
- Roemmich, D., Gilson, J., 2009. The 2004–2008 mean and annual cycle of temperature, salinity, and steric height in the global ocean from the Argo Program. *Prog. Oceanogr.* 82 (2), 81–100.
- Satish Kumar, K., Venkata Rathnam, E., Sridhar, V., 2021. Tracking seasonal and monthly drought with GRACE-based terrestrial water storage assessments over major river basins in South India. *Sci. Total Environ.* 763, 142994.
- Scanlon, B.R., Zhang, Z., Save, H., et al., 2018. Global models underestimate large decadal declining and rising water storage trends relative to GRACE satellite data. *Proc. Natl. Acad. Sci.* 115 (6), E1080.
- Schlegel, N., Wiese, D., Larour, E., et al., 2016. Application of GRACE to the assessment of model-based estimates of monthly Greenland Ice Sheet mass balance (2003–2012). *Cryosphere* 10, 1965–1989.
- Sridhar, V., Ali, S.A., Lakshmi, V., 2019. Assessment and validation of total water storage in the Chesapeake Bay watershed using GRACE. *J. Hydrol. Reg. Stud.* 24, 100607.
- Sun, Z., Zhu, X., Pan, Y., et al., 2018. Drought evaluation using the GRACE terrestrial water storage deficit over the Yangtze River Basin, China. *Sci. Total Environ.* 634, 727–738.
- Sun, A.Y., Scanlon, B.R., Zhang, Z., et al., 2019. Combining physically based modeling and deep learning for fusing GRACE satellite data: can we learn from mismatch? *Water Resour. Res.* 55 (2), 1179–1195.
- Sun, Z., Long, D., Yang, W., et al., 2020. Reconstruction of GRACE data on changes in total water storage over the global land surface and 60 basins. *Water Resour. Res.* 56 (4), e2019WR026250.
- Swann, A.L.S., Koven, C.D., 2017. A direct estimate of the seasonal cycle of evapotranspiration over the Amazon Basin. *J. Hydrometeorol.* 18 (8), 2173–2185.
- Syed, T.H., Famiglietti, J.S., Rodell, M., et al., 2008. Analysis of terrestrial water storage changes from GRACE and GLDAS. *Water Resour. Res.* 44 (2), W02433.
- Tapley, B.D., Bettadpur, S., Ries, J.C., et al., 2004. GRACE measurements of mass variability in the Earth system. *Science* 305 (5683), 503–505.
- Tapley, B.D., Watkins, M.M., Flechtner, F., et al., 2019. Contributions of GRACE to understanding climate change. *Nat. Clim. Chang.* 9 (5), 358–369.
- Thomas, H.A., 1981. Improved Methods for National Water Assessment, Water Resources Contract, p. 59 WR15249270.
- Velicogna, I., Wahr, J., 2005. Greenland mass balance from GRACE. *Geophys. Res. Lett.* 32 (18), L18505.
- Wang, J., Song, C., Reager, J.T., et al., 2018a. Recent global decline in endorheic basin water storages. *Nat. Geosci.* 11 (12), 926–932.
- Wang, X.-J., Zhang, J.-Y., Gao, J., et al., 2018b. The new concept of water resources management in China: ensuring water security in changing environment. *Environ. Dev. Sustain.* 20 (2), 897–909.
- Wang, J., Chen, Y., Wang, Z., et al., 2020a. Drought evaluation over Yangtze River basin based on weighted water storage deficit. *J. Hydrol.* 591, 125283.
- Wang, X., Xiao, X., Zou, Z., et al., 2020b. Gainers and losers of surface and terrestrial water resources in China during 1989–2016. *Nat. Commun.* 11 (1), 3471.
- Watkins, M.M., Wiese, D.N., Yuan, D.-N., et al., 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.
- Werth, S., Günther, A., Petrovic, S., et al., 2009. Integration of GRACE mass variations into a global hydrological model. *Earth Planet. Sci. Lett.* 277 (1), 166–173.
- Wiese, D.N., Landerer, F.W., Watkins, M.M., 2016. Quantifying and reducing leakage errors in the JPL RL05M GRACE mascon solution. *Water Resour. Res.* 52 (9), 7490–7502.
- Wiese, D.N., Yuan, D.-N., Boening, C., et al., 2019. JPL GRACE and GRACE-FO Mascon Ocean, Ice, and Hydrology Equivalent Water Height Coastal Resolution Improvement (CRI) Filtered Release 06 Version 02. NASA Physical Oceanography DAAC.
- Woods, R.A., 2009. Analytical model of seasonal climate impacts on snow hydrology: continuous snowpacks. *Adv. Water Resour.* 32 (10), 1465–1481.
- Xie, J., Xu, Y.-P., Wang, Y., et al., 2019. Influences of climatic variability and human activities on terrestrial water storage variations across the Yellow River basin in the recent decade. *J. Hydrol.* 579, 124218.
- Xu, Y., Gong, H., Chen, B., et al., 2021. Long-term and seasonal variation in groundwater storage in the North China Plain based on GRACE. *Int. J. Appl. Earth Obs. Geoinf.* 104, 102560.
- Yang, H., Yang, D., Lei, Z., et al., 2008. New analytical derivation of the mean annual water-energy balance equation. *Water Resour. Res.* 44 (3), W034103.
- Zhang, L., Potter, N., Hinkel, K., et al., 2008. Water balance modeling over variable time scales based on the budyko framework – model development and testing. *J. Hydrol.* 360 (1), 117–131.
- Zhang, Z., Chao, B.F., Chen, J., et al., 2015. Terrestrial water storage anomalies of Yangtze River Basin droughts observed by GRACE and connections with ENSO. *Glob. Planet. Chang.* 126, 35–45.
- Zhang, Y., Pan, M., Sheffield, J., et al., 2018. A climate data record (CDR) for the global terrestrial water budget: 1984–2010. *Hydroclim. Earth Syst. Sci.* 22 (1), 241–263.
- Zhang, D., Liu, X., Bai, P., 2019a. Assessment of hydrological drought and its recovery time for eight tributaries of the Yangtze River (China) based on downscaled GRACE data. *J. Hydrol.* 568, 592–603.
- Zhang, X., Dong, Q., Cheng, L., et al., 2019b. A budyko-based framework for quantifying the impacts of aridity index and other factors on annual runoff. *J. Hydrol.* 579, 124224.
- Zhang, X., Dong, Q., Costa, V., et al., 2019c. A hierarchical bayesian model for decomposing the impacts of human activities and climate change on water resources in China. *Sci. Total Environ.* 665, 836–847.
- Zhang, Y., He, B., Guo, L., et al., 2019d. The relative contributions of precipitation, evapotranspiration, and runoff to terrestrial water storage changes across 168 river basins. *J. Hydrol.* 579, 124194.
- Zhang, X., Dong, Q., Zhang, Q., et al., 2020. A unified framework of water balance models for monthly, annual, and mean annual timescales. *J. Hydrol.* 589, 125186.