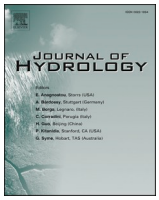
Journal of Hydrology 635 (2024) 131096

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol



Research papers

Quantifying the long-term changes of terrestrial water storage and their driving factors☆

Xiaoying Shi [a,*,1], Yaoping Wang [a,1], Jiafu Mao [a], Peter E. Thornton [a], Daniel M. Riccuito [a], Forrest M. Hoffman [b], Yuefeng Hao [c]

[a] Environmental Sciences Division and Climate Change Science Institute, Oak Ridge National Laboratory, Oak Ridge, TN, USA
[b] Computational Sciences and Engineering Division and Climate Change Science Institute, Oak Ridge National Laboratory, Oak Ridge, TN, USA
[c] Institute for a Secure and Sustainable Environment, The University of Tennessee, Knoxville, TN, USA

ARTICLE INFO

This manuscript was handled by Marco Borga, Editor-in-Chief, with the assistance of Di Long, Associate Editor

Keywords:
TWS
GRACE
ELM v1 model
Climate change
Factorial simulations

ABSTRACT

Global warming is expected to cause changes in terrestrial water storage (TWS) across the land surface, with widespread impacts on ecosystems and society. Although extensive research has been performed to analyze TWS changes and possible drivers during the post-2000 period, longer-term evolution of TWS and associated environmental forcings remain relatively unexplored. In this study, we evaluated the performance of the Energy Exascale Earth System model (E3SM) land model ELM version 1 (ELM v1) in simulating global TWS, and used factorial simulations of ELMv1 to quantify global TWS changes and their drivers during 1948–2012. We found that ELM's agreed best with existing satellites and reconstruction datasets in temperate regions unaffected by irrigation. Biome- and climate zone-averaged TWS mainly increased at rates between 0 and 10 mm/year over 1948–2012, but the second half of that period saw smaller positive trends than the first half or even negative trends. Climate change explained >80 % of the TWS trends across most biomes and climate zones, followed by land use and land cover change. The physiological and phenological effects of CO_2 primarily induced noticeable TWS trends in the more humid biomes and climate zones across different latitudes. In contrast, nitrogen deposition and aerosol deposition generally had smaller and negative impacts across the biomes and climate regions. Among the meteorological drivers analyzed, the long-term average imbalance between precipitation (P), evapotranspiration (E), and runoff (Q) contributed >50 % of the TWS trends in most biomes and climate zones, with nonlinearity being induced by spatially heterogenous changes in E/P and Q/P ratios. The accumulated detrended anomalies in P, E, and Q also often contributed substantially, while the trends difference between P, E, and Q contributed little. Together, these findings unveiled an intensification of the global TWS and its diverse patterns of climate change and different non-withdrawal human-induced alterations, contributing to a more comprehensive understanding and projection of the global water cycle.

1. Introduction

Terrestrial water storage (TWS) is the sum of individual water components (e.g., soil moisture, ground water, snow, canopy water storage) on the land surface and in the subsurface. By monitoring and analyzing changes in TWS over space and time, we can develop a deeper understanding of global and regional hydrological cycles, biogeochemical cycles, ecosystem services, and human sustainability (Haddeland et al., 2014; Jiménez Cisneros et al., 2014; Pokhrel et al., 2021; Scanlon et al., 2018; Strassberg et al., 2009; Syed et al., 2008; Yeh et al., 2006).

Observations from the Gravity Recovery and Climate Experiment (GRACE) and GRACE Follow-On (GRACE-FO) satellite missions and land

☆ This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (https://energy.gov/downloads/doe-public-access-plan).

* Corresponding author.
E-mail address: shix@ornl.gov (X. Shi).

[1] These authors contributed equally: Xiaoying Shi, Yaoping Wang.

https://doi.org/10.1016/j.jhydrol.2024.131096
Received 31 March 2023; Received in revised form 6 February 2024; Accepted 5 March 2024
Available online 22 March 2024

and hydrological modeling results have been used to examine the changes and variability in TWS, especially during the post-2000 period (Asoka et al., 2017; Ju et al., 2023; Tapley et al., 2019). These studies show that TWS decreased in about 30 % of global basins, including endorheic basins, and basins in the northern high and midlatitudes (e.g., Lena, Yenisei and Ob basins and those in northern India), as well as in most low-latitude basins (e.g., Nile, Congo, Murray and Zambezi basins) (Ahmed et al., 2014; Scanlon et al., 2016; Wang et al., 2018; Zhang et al., 2019a). Significant increasing trends in TWS were found over west (e.g., Gambia and Niger) and South Africa (e.g., Zambezi), South America (e.g., Essequibo), North America (e.g., Koksoak and Missouri), central India (e.g., Narmada and Godavari), the north Tibetan plateau (Li et al., 2022), and middle Yangtze River basin (Long et al., 2017). Some studies have attributed changes in TWS to different factors. For example, snow storage was shown to impact TWS anomalies in the high-latitude basins and the Himalayas, while groundwater and soil moisture storage determined TWS variations in other regions, such as Indian and China (Felfelani et al., 2017; Lv et al., 2021; Rodell et al., 2018; Zhang et al., 2019b). Zhang et al., (2019b) investigated the relative contributions of P, E, and Q to TWS changes across 168 river basins; they found hydrological fluxes (P, E, and Q) explained more than 60 % of global TWS changes. Moreover, the impacts of anthropogenic withdrawals on TWS were found to be large, primarily in regions with high irrigation demands (e.g. northern India, North China plain) and dam constructions (e.g. northern Brazil, the Three Gorges dam region in China) (Asoka et al., 2021; Lv et al., 2021; Rodell et al., 2018).

The limitations of using GRACE/GRACE-FO data to investigate TWS changes include the relatively short time period of approximately 20 years and the observational nature of the data, which means decadal to multidecadal variabilities may not be well-captured. Furthermore, the data doesn't allow for the disentanglement of the contributions of different TWS components or the differing influences from various climate and anthropogenic drivers on TWS changes (An et al., 2021; Rodell et al., 2018). Therefore, global hydrological and land surface models are typically used to achieve those purposes. However, it is known that they underestimate decadal TWS trends compared to the GRACE observations due to missing components such as snow water and ground water storage, as well as inadequate representations of human processes such as land use and land cover change (LULCC), irrigation, and impoundment (Asoka and Mishra, 2020; Jing et al., 2019; Scanlon et al., 2018). For example, hydrological and land surface modeling was used to analyze the contributions of hydrological components to TWS changes by Chao et al. (2021) and Guo et al. (2022). They revealed that human-induced ground water changes accounted for approximately half of the total ground water storge of the Yangtze River basin and ground water storage changes also contributed more than 50 % of TWS changes in Inner Mongolia.

Based on the above literature review, previous studies have mainly focused on the impacts of human withdrawals and trends in meteorological drivers on TWS changes. However, there has been limited exploration of the specific roles played by non-withdrawal natural or human forcings (e.g., atmospheric CO_2 concentration, LULCC, nitrogen deposition and aerosol deposition), which may be important for understanding TWS changes in the less disturbed natural ecosystems. Those forcings have been shown to notably modulate the land hydrological cycle (Almendra-Martín et al., 2022; Forbes et al., 2019, 2018; Gentine et al., 2019; Mao et al., 2015; Naha et al., 2021; Piao et al., 2007; Shi et al., 2013, 2011). Factorial experiments have been widely used to quantify the contributions of those forcings on the changes of hydrological cycle (Cui et al., 2021; Forbes et al., 2019, 2018; Mao et al., 2015; Padrón et al., 2020; Shi et al., 2013, 2011; Yang et al., 2019; Zhu et al., 2021).

As part of the Land Surface, Snow and Soil Moisture Model Intercomparison project (LS3MIP [van den Hurk et al., 2016]), ELM v1 conducted a unique set of 36 factorial simulations. These simulations effectively separated the effects of individual environmental forcings (including climate change, atmospheric CO_2 concentration, LULCC, nitrogen deposition and aerosol deposition) on land surface variables and sampled different model configurations and meteorological forcing datasets. In offline land model simulations, it's worth noting that the climate change factor is based on observational meteorological datasets, encompassing the impacts of increasing atmospheric CO_2 concentration and anthropogenic aerosol emissions on the climate. The atmospheric CO_2 concentration factor here refers specifically to both physiological (e.g., stomatal opening and closure) and phenological (e.g., leaf area index) effects of CO_2 (Zhu et al., 2021; Fatichi et al., 2016). The aerosol deposition factor refers specifically to the effect on snow albedo, which further affects the hydrological cycle. In this study, we took advantage of those factorial simulations to answer questions such as how global TWS evolved before the satellite period (1948–2012), how individual environmental forcings contributed to annual TWS trends, how the TWS trends related to the meteorological factors P, E, and Q, and the uncertainty therein. We note that ELM v1 does not simulate human water withdrawal processes. Therefore, the value of this paper mainly lies in illustrating how the environmental forcings affect TWS changes in natural ecosystems, which are still important to understanding global TWS changes.

To ascertain the quality of the ELM simulations, we evaluated the modeled TWS with respect to trend using GRACE satellites observations during the short overlapping period (2003–2012), and with respect to de-trended anomalies using two GRACE-based TWS reconstruction datasets, GRACE-REC (Humphrey and Gudmundsson, 2019) and Li et al. (2021) (hereafter called Li-REC). Although the reconstruction datasets cannot be used to evaluate TWS trends due to the setup of the methodologies (Humphrey and Gudmundsson, 2019; Li et al., 2021), they were useful supplements to the GRACE data. This is attributed to their long period of availability (GRACE-REC: 1902–2014, Li-REC: July 1979–June 2020) and their utility in evaluation on interannual anomalies. Such evaluations can still reveal any large errors in ELM hydrological processes that affect TWS. Past studies have used the GRACE-REC dataset in attributing changes in global dry-season water availability to human-induced climate change (Padrón et al., 2020), and the Li-REC dataset was also used in a few hydrological studies (Mo et al., 2022; Xiong et al., 2022a; Zou et al., 2022). The evaluation methods and results are reported in Sections 2.1–2.2 and 3.1. Based on the evaluation results, we found ELM performance was mainly satisfactory in temperate biomes and climate zones not heavily affected by irrigation, but unsatisfactory in heavily irrigated, high-latitude, and tropical regions. Section 4.1 discusses the region-specific potential sources of error.

We then used the factorial simulations to decompose historical regional trends in TWS into contributions from the individual environmental forcings and contributions from different hydrological components (climatology, trends, and variability in P, E, and Q). The methods and results of the decomposition are reported in Section 2.3, 3.2, and 3.3. We found globally increasing trends in TWS during 1948–2012, but weakened increasing trends or decreasing trends in many temperate regions since the 70 s or 80 s. Both the increasing and decreasing trends were mainly driven by climate forcing, followed by LULCC (mainly conversion from natural vegetation to croplands) in the temperate regions and CO_2 effects in the boreal forests. Among the hydrological components, the increasing or decreasing trends in TWS were always mainly driven by climatological imbalance between P, E, and Q and very little by the trends in P, E, and Q. This finding implies TWS trends should be better understood in terms of changes in the partitioning of precipitation instead of changes in the component fluxes (Section 4.2).

2. Methodology

2.1. Model configurations and simulations

ELM v1 used here is based on the Community Land Model version 4.5 (CLM 4.5), including terrestrial hydrological processes such as

interception of precipitation by the vegetation canopy, throughfall, infiltration, surface and subsurface runoff, snow and soil moisture evolution, evaporation from soil and vegetation and transpiration (Oleson et al., 2013). Subsurface hydrology in ELM v1 consists of a 10-layer soil column spanning the uppermost 3.8 m of the land surface coupled to an additional "aquifer" layer representing the time-varying volume of soil between the deepest layer of the static soil column and the water table (Lawrence et al., 2011; Niu et al., 2007). The aquifer layer represents changes in saturated subsurface water storage, i.e., groundwater.

The TWS in ELM v1 consists of the following components: canopy water, snow water equivalent, surface water storage, soil liquid water and ice contents, and groundwater. Each component is simulated separately and added up to give the TWS value. Those components capture main components of TWS in the real world, but miss glacier water storage and the direct impacts from human land–water management activities (e.g., water withdrawals, mining, and dam constructions (Chao et al., 2021; Guo et al., 2022), and minor biological storages such as in tree stems. Since ELM v1 does not include glacier storage in its TWS term, it cannot be expected to capture TWS changes in glaciated regions. Therefore, we used the land cover input to the ELM v1 simulations to exclude the maximum region with glacier cover from all the model evaluations and analysis in this paper, resulting in some blank regions in the global maps (e.g., Greenland, southern coast of Alaska, southern part of the Andes mountains, the Himalayas). The lack of human land–water management activities may introduce some degree of errors in TWS trend over regions where such activities are intensive (e.g., India and north China Plain) (Fig. S3).

The 36 ELM offline simulations were driven by atmospheric and other environmental forcings. ELM was first spun up for preindustrial conditions in 1850, followed by historical transient simulations with time-varying atmospheric CO_2 concentration, LULCC, nitrogen deposition, aerosol deposition, and/or climate forcings (Table 1). Three climate forcings were used, including GSWP3 (Kim, 2017), CRUNCEP (Viovy, 2009), and PRINCETON (Sheffield et al., 2006). Details about the data sources for other forcings, such as CO_2 concentration, LULCC, nitrogen deposition, and aerosol deposition, are described in Burrows et al. (2020). For each climate forcing setup, we used two model configurations, carbon–nitrogen-phosphorus (CNP) (Yang et al., 2019) and carbon–nitrogen only (CN). The CN configuration was executed with the same model code as CNP but assuming phosphorus saturation. At the time of the simulations, the GSWP3 forcing spans 1901–2014, the CRUNCEP forcing spans 1901–2016, and the PRINCETON forcing spans 1901–2012. However, we found the simulations prior to 1920 have degraded quality compared to later years. We also considered the fact that meteorological observations that feed into the construction of those global forcings are relatively sparse during the first half of the 20th century, especially in the Middle East, Asia, Africa, and South America (Harris et al., 2020, 2014; Sheffield et al., 2006). Therefore, we decided to limit the time period of analysis to 1948–2012.

All the simulations were conducted at global 0.5° spatial resolution. All the non-TWS outputs (we used P, E, and Q in this study) were saved as monthly averages and we further averaged them to annual level in subsequent analysis. ELM saves two different types of TWS outputs: the monthly averages, and the instantaneous values at the end of each month. The monthly averages are comparable to observational data and we averaged them to annual level for model evaluation. The instantaneous TWS values satisfy the water balance equation by virtue of mass conservation in ELM: $TWS_t = TWS_{t-1} + (P_t - E_t - Q_t)$, where t can be any time step, P, E, and Q are, respectively, the ELM-simulated total P, E, and Q between time step *t-1* and *t*. We converted the end-of-month TWS values to end-of-year values by retaining only the December value of each year, whereupon the total P, E, Q in the water balance equation would be the annual average rates (mm/day) multiplied by number of days in a year. The water balance equation facilitates decomposition of the contributions to TWS changes by meteorological drivers. For the purpose of consistency, we therefore used the end-of-year values for decomposition analysis on both environmental forcings and meteorological drivers.

To attribute TWS changes to climate change and other environmental forcings, we used differences between two simulations that differed in only one forcing, which is a common approach in attributing the effects of environmental forcings on climate variables (Langenbrunner et al., 2019; Zhao et al., 2019). We interpreted S6–S1 as the total effect of all forcings, which is equal to the sum of the effects of changes in climate change (S2–S1), LULCC (S3–S2), atmospheric CO_2 concentration (S4–S3), nitrogen deposition (S5–S4) and aerosol deposition (S6–S5) (Table 1). A forcing contributed more to the total effect if the magnitude of its effect was greater relative to the total effect. We also note that the effects of the individual forcings, here evaluated by subtractions, include incremental interactive effects. For example, the CO_2 concentration's effect, evaluated by S4–S3, includes both the effect of CO_2 concentration alone and the interactive effects between CO_2 concentration and all the previously included forcings in S3 (transient climate and LULCC).

Table 1
Global factorial (S1 to S6) offline simulations of ELMv1.0 with CN or CNP dynamics driven by the GSWP3v2 (1850–2014), CRUNCEPv8 (1850–2016), and PRINCETON (1850–2012) meteorology forcings (36 simulations).

Simulations	Climate forcing and two configurations		
	GSWP3 (CN and CNP)	CRUNCEP (CN and CNP)	PRINCETON (CN and CNP)
S1	1901–1920 climate cycling and 1850 conditions		
S2	Transient climate only		
S3	Transient climate and LULCC		
S4	Transient climate, LULCC and CO_2		
S5	Transient climate, LULCC, CO_2 and nitrogen deposition		
S6	Transient climate, LULCC, CO_2, nitrogen and aerosol depositions		
S2-S1	Climate only		
S3-S2	LULCC only		
S4-S3	CO_2 only		
S5-S4	Nitrogen deposition only		
S6-S5	Aerosol deposition only		

2.2. Observational data

We employed three widely used mass concentration (mascon) GRACE solutions to evaluate the trend in ELM-simulated annual TWS over the longest overlapping period (2003–2012) (Hasan and Tarhule, 2021; Humphrey and Gudmundsson, 2019) (Table 2). Note the overlapping period was before the launch of the GRACE-FO mission. The mascon GRACE solutions are more accurate than traditional spherical harmonic solutions because of reduced leakage errors and no need for empirical smoothing (Hasan and Tarhule, 2021). We bilinearly interpolated the original GRACE data to 0.5° and averaged the monthly values to annual averages before the comparison with ELM v1. We calculated the linear least squares trend separately for each GRACE dataset and each ELM v1 S6 simulation, and then compared between the average, maximum, and minimum trends over all three GRACE datasets and the average, maximum, and minimum trends over the ELM v1 S6 simulations driven by all the combinations of model configurations and meteorological forcings.

In order to understand potential causes of disagreement between the ELM v1 trends and the GRACE trends, we also evaluated the ELM v1 trends in selected major components of TWS: soil moisture, snow water equivalent, and the sum of soil moisture and groundwater. We used the MERRA and GLDAS Noah2.0 reanalysis to evaluate ELM-simulated soil moisture and snow water equivalent because they were well-accepted data sources for the two variables (Beck et al., 2021; Mortimer et al., 2020) and because they provide column-integrated soil moisture beyond the root zone, which were comparable to the definition of soil moisture in ELM v1. We also found an existing groundwater dataset from G3P

Table 2
List of non-ELM datasets used in this study.

Variable	Time period	Source resolution	Data source	Reference
GRACE TWS observations	April 2022–June 2023	0.25°	CSR RL06 v2	Himanshu, 2020; Save et al., 2016
	April 2022–present	0.5°	JPL RL06.1 v3 CRI	Watkins et al., 2015; Wiese et al., 2023
	April 2022–May 2023	0.5°	GSFC RL06 v2	Loomis et al., 2019
GRACE-REC TWS reconstructions	1979–2016	0.5°	JPL-MSWEP, GSFC-MSWEP	Humphrey and Gudmundsson, 2019
	1979–2019	0.5°	JPL-ERA5, GSFC-ERA5	
	1901–2014	0.5°	JPL-GSWP3, GSFC-GSWP3	
Li-REC TWS reconstructions	June 1979–June 2020	0.5°	–	Li et al., 2021
Column total soil moisture	1980–present	0.5° × 0.625°	MERRA2	Gelaro et al., 2017
	1948–2014	0.25°	GLDAS Noah 2.0	Beaudoing et al., 2019
Snow water equivalent	1980–present	0.5° × 0.625°	MERRA2	Gelaro et al., 2017
	1948–2014	0.25°	GLDAS Noah 2.0	Beaudoing et al., 2019
Sum of soil moisture and groundwater	April 2022–May 2023	0.5°	G3P	G3P, 2023

(https://www.g3p.eu/). This groundwater dataset was derived by subtracting the other components from the GRACE/GRACE-FO observations, and the soil moisture component was separated from groundwater at 1 m depth, which is shallower than in ELM v1 (G3P, 2023). As a result, the definition of groundwater in this dataset is not comparable to that in ELM v1. Instead, the sum of soil moisture and groundwater has comparable definitions between this dataset and ELM. Therefore, we compared the sum of groundwater and soil moisture between G3P and ELM. We converted all these variables into the same unit as TWS, i.e. mm, or kg water equivalent/m^2, before calculating the trend, to make the results comparable to TWS trends.

Since the overlapping period between GRACE and ELM is short, we supplemented the trend evaluation with evaluation of de-trended anomalies using GRACE and two longer reconstruction datasets, GRACE-REC (Humphrey and Gudmundsson, 2019) and Li-REC (Li et al., 2021). We calculated the de-trended TWS anomalies by subtracting the fitted linear regression equation between year and the TWS values. The evaluation metrics were (1) Pearson correlation between ELM and the GRACE observations/reconstructions, and (2) the root mean squared error between ELM and the GRACE observations/reconstructions, subtracted by the standard deviation of the ELM detrended anomalies ("RMSE – ELM std"). When the metric is negative, it means the detrended ELM anomalies are more similar to the observations/reconstructions than to a vanilla prediction (zero), i.e., ELM has predictive power of the observed/reconstructed anomalies. We chose to subtract the standard deviation of ELM, instead of the standard deviation of the observations/reconstructions, because during preliminary analysis, we found the detrended interannual variabilities in the reconstruction datasets to almost always underestimate the GRACE observed detrended interannual variabilities, while the ELM did not. This might reflect the inherent inability of statistical regression to capture all the sources of variability in their predictands (Humphrey and Gudmundsson, 2019; Li et al., 2021), but its investigation is beyond the scope of this study. Both reconstruction datasets were based on calibrated relationships between climate variables and the GRACE/GRACE-FO-based TWS observations. Such relationships were unsuitable for estimating long-term trends in TWS, but performed better than hydrological and land surface models in capturing the interannual TWS anomalies (Humphrey and Gudmundsson, 2019; Li et al., 2021). The reconstruction method for GRACE-REC was an exponential filter-like statistical model that predicts TWS anomalies from precipitation and near surface temperature (Humphrey and Gudmundsson, 2019). To consider the uncertainty in input fields, the reconstruction method was applied on combinations of two sources of GRACE/GRACE-FO satellites data (JPL mascons RL06 with CRI and GSFC mascons v2.4 ICE6G) and three sources of meteorological data (MSWEP, ERA5, and GSWP3). The resulting six GRACE-REC datasets are JPL-MSWEP, JPL-ERA5, JPL-GSWP3, GSFC-MSWEP, GSFC-ERA5, and GSFC-GSWP3. Due to different lengths of the meteorological data, JPL-MSWEP and GSFC-MSWEP span 1979–2016, JPL-ERA5 and GSFC-ERA5 span 1979–2019, and JPL-GSWP3 and GSFC-GSWP3 span 1901–2014. We therefore calculated the evaluation metrics separately between each ELM S6 simulation and each GRACE-REC dataset, using the longest reliable overlapping period (i.e., 1979–2012 or 1948–2012). We then averaged the evaluation metrics individually for each meteorological input of GRACE-REC (MSWEP, ERA5, or GSWP3), over all combinations of the six ELM S6 simulations and two GRACE inputs of GRACE-REC (JPL, GSFC). The reconstruction method for Li-REC involved decomposing the TWS fields into spatial modes and time series, and then predicting the time series using traditional statistical or neural network models and the best predictor set selected from hundreds to thousands of potential predictors (Li et al., 2021, 2020). The resulting single dataset span June 1979–June 2020. We therefore calculated the evaluation metrics between the dataset and each ELM S6 simulation over the longest feasible overlapping period (1980–2012), and averaged the results over all six combinations of ELM configuration and meteorological forcings.

In order to understand the spatial variations in ELM performance, we calculated all the evaluation metrics at the levels of individual 0.5° grids, biome average TWS, and climate-zone average TWS. The grid-level comparison between ELM simulations and GRACE observations may be impacted by the coarse native spatial resolutions of the GRACE satellites (~200,000 km^2, implying nearly one hundred 0.5° grids near the tropics and much more in the polar regions) (Long et al., 2015). Therefore, performance evaluation at more aggregated levels may provide a fairer view of ELM performance. The biome classification was based on the MODIS MCD12C1 product, and the spatial extent of the biomes used in this study is shown in Fig. S1 (Friedl and Sulla-Menashe, 2015). The climate zone classification was based on a recent high resolution Köppen-Geiger climate classification product (Beck et al., 2018) and the spatial extent of the climate zones is shown in Fig. S2. We excluded a few climate zones smaller than or close to the GRACE footprint from comparison because those would face high uncertainty (Cfc – 9.81 * 10^4 km, Dsa – 2.17 * 10^5 km^2, Dsb–5.43 * 10^5 km^2, Dsd – 5.21 * 10^4 km^2, Dwd – 2.26 * 10^5 km^2, Dfd – 6.18 * 10^5 km^2) (Long et al., 2015).

2.3. Decomposition of the contributions to TWS trends

As noted in Section 2.1, we used annual average TWS in assessing the performance of ELM v1, but used end-of-year TWS values to decompose the contributions from environmental forcings or meteorological drivers because of the water balance equation and for consistency. To account for uncertainty in ELM v1 configurations and meteorological forcings, we performed the decomposition separately over different ways of averaging the ELM v1 simulations: average over all the meteorological forcings for the CN/CNP configuration only (CN-met/CNP-met), average over all the configurations for the CRUNCEP/PRINCETON/GSWP3 meteorological forcings only (elm-CRUNCEP/elm-PRINCETON/elm-GSWP3), and average over all the ensemble members (Avg).

We performed the decomposition at both biome and climate zone levels. The time period 1948–2012 is long and visual inspection of the evolution of TWS during this period suggested that nonlinear trends existed in at least some biomes and climate zones (Fig. S10). We accounted for this fact by searching for breakpoints in linear trends using the dynamic programming method implemented in the "ruptures" package in python (Truong et al., 2020). The resulting breakpoints were generally floating point numbers that fell between two integer years. We then fitted separate least squares linear regression before and after the breakpoint, and compared whether such two-piece linear regression results better described the data than simple linear regression using the Bayesian Information Criterion (BIC), a metric for model selection that balances model complexity and performance (Hastie et al., 2009). Table S1 displays the comparison results for the biome- and climate zone-average trends under all forcing (S6–S1) and shows that piecewise linear regression was indeed sometimes superior. Therefore, for those biomes and climate zones, we performed the decomposition by forcing and by meteorological drivers separately for before and after the breakpoints. We identified the breakpoints separately for different ways of averaging the ELM simulations (CNP-met, CN-met, elm-CRUNCEP, elm-PRINCETON, elm-GSWP3, Avg).

As noted in Section 2.1, we decomposed the effects of all environmental forcings on TWS (S6–S1) as the sum of the effects of individual forcings (S2–S1, S3–S2, S4–S3, S5–S4, S6–S5). The formula of the slope of least squares linear regression is a linear transformation of the dependent variable using a matrix that is purely based on the independent variable (DeGroot and Schervish, 2018). Since the independent variable was the same (the years 1948–breakpoint or breakpoint–2012) in calculating the trends of any ELM v1 simulations, the trends of the individual forcing ELM v1 simulations were additive. In other words, the TWS trends due to the effects of all forcings (i.e., TWS trends calculated on the TWS series obtained by S6–S1) are also equal to the sum of the TWS trends due to the individual forcings (S2–S1, S3–S2, S4–S3, S5–S4, S6–S5). We therefore interpreted the TWS trends of the individual forcing simulations as the contributions of individual forcings to the total TWS trends.

To decompose the TWS trends due to the effects of all forcings (S6–S1) into contributions by meteorological variables, we used the water balance equation $TWS_t = TWS_{t-1} + (P_t - E_t - Q_t)$ and wrote both the TWS and each meteorological driver as the sum of its mean value over the time period of interest ($t = 1, 2, \cdots, T$, T being the length of the time period), a linear trend, and remainder anomalies. Here, we also denote $D_t = P_t - E_t - Q_t$ for brevity. Thus

$$D_t = \overline{D} + \beta_D t' + \varepsilon_t = (\overline{P} - \overline{E} - \overline{Q}) + (\beta_P - \beta_E - \beta_Q)t' + (\varepsilon_{P,t} - \varepsilon_{E,t} - \varepsilon_{Q,t}) \tag{1}$$

where the overbar denotes mean value over the time period, and the conversion $t' = t - \frac{T+1}{2}$ ensures that $\overline{P} + \beta_P t' + \varepsilon_{P,t}$, $\overline{E} + \beta_{ET} t' + \varepsilon_{ET,t}$, and $\overline{Q} + \beta_Q t' + \varepsilon_{Q,t}$ were simply re-written forms of least squares linear regression between P_t, E_t, Q_t, respectively, and t. Substituting Eq. (1) into the balance equation and taking the cumulative sum yields:

$$TWS_t = TWS_0 + t\overline{D} + \beta_D \frac{(t-T)t}{2} + \sum_{i=1}^{t} \varepsilon_i \tag{2}$$

Since the formula of linear regression gives the trend in TWS as

$$\widehat{a} = \frac{\sum_{t=1}^{T} (TWS_t - \overline{TWS}) \left(t - \frac{T+1}{2}\right)}{\sum_{t=1}^{T} \left(t - \frac{T+1}{2}\right)^2} \tag{3}$$

One can substitute (2) into (3), and after some simplifications obtain

$$\begin{aligned} \widehat{a} = {} & \overline{D} + \frac{\beta_D}{2} + \frac{\sum_{t=1}^{T} \left(\sum_{i=1}^{t} \varepsilon_i\right) \left(t - \frac{T+1}{2}\right)}{\sum_{t=1}^{T} \left(t - \frac{T+1}{2}\right)^2} = (\overline{P} - \overline{E} - \overline{Q}) \\ & + \frac{\beta_P - \beta_E - \beta_Q}{2} + \left[\frac{\sum_{t=1}^{T} \left(\sum_{i=1}^{t} \varepsilon_{P,i}\right) \left(t - \frac{T+1}{2}\right)}{\sum_{t=1}^{T} \left(t - \frac{T+1}{2}\right)^2} \right. \\ & \left. - \frac{\sum_{t=1}^{T} \left(\sum_{i=1}^{t} \varepsilon_{E,i}\right) \left(t - \frac{T+1}{2}\right)}{\sum_{t=1}^{T} \left(t - \frac{T+1}{2}\right)^2} - \frac{\sum_{t=1}^{T} \left(\sum_{i=1}^{t} \varepsilon_{Q,i}\right) \left(t - \frac{T+1}{2}\right)}{\sum_{t=1}^{T} \left(t - \frac{T+1}{2}\right)^2} \right] \end{aligned} \tag{4}$$

Based on Eq. (4), the trend in TWS was influenced by the climatological imbalance between P, E, and Q ($\overline{P} - \overline{E} - \overline{Q}$), the trends in P, E, and Q ($\frac{\beta_P - \beta_E - \beta_Q}{2}$), and accumulated random anomalies. Note the "random" here only means not accounted by the climatology or trend, and may still be accounted by non-random physical factors like teleconnections (Guo et al., 2021).

3. Results

3.1. Evaluation of ELM performance on TWS

3.1.1. Grid-level evaluation

The broad patterns of ELM-simulated and GRACE observed trends in annual average TWS over 2003–2012 were similar (Fig. 1). Both had positive trends in the center of North America, central and northern Europe, western and southern India, the northern half of South America, the Sahel region, southern Africa, and northeastern Australia (Fig. 1ad). Both had negative trends in eastern Europe, Mongolia, southeastern U. S., the Sahara, the Middle East, the eastern coast of Africa, western Australia, and southern South America (Fig. 1ad). The ranges of uncertainty across the ELM v1 simulations and GRACE observations were large, which may be partially due to the short length of feasible comparison between ELM v1 and GRACE data (10 years), but the broad spatial patterns of relative positivity and negativity of the trends held (Fig. 1bcefhi). The main regions of discrepancy between ELM v1 and GRACE were the northern high latitudes (Alaska, western Canada, and much of Siberia), central Asia and southwestern Russia, Syria-Iraq-Iran, northern India, the North China Plain, and the center of Africa, where ELM v1 underestimated the magnitude of negative trends beyond 10 mm/year (Fig. 1ad). When averaged latitudinally, ELM v1 reproduced the observed trends fairly well, with some overestimation towards the high norther latitudes and the southernmost tip of South America.

To supplement the evaluation of ELM v1 performances on short-term trends, we evaluated ELM on its ability to simulate detrended interannual anomalies in TWS using GRACE and various GRACE/GRACE-FO based reconstruction data (Section 2.2). The Pearson correlations of ELM were the highest against the GSWP3 forced GRACE-REC, intermediate against GRACE observations and the MSWEP and ERA5 forced GRACE-REC, and worst against Li-REC (Fig. S4a–e). In all the datasets, the lower Pearson correlations occurred over northern Canada, the grassland regions of South America, and northern to central Africa. In

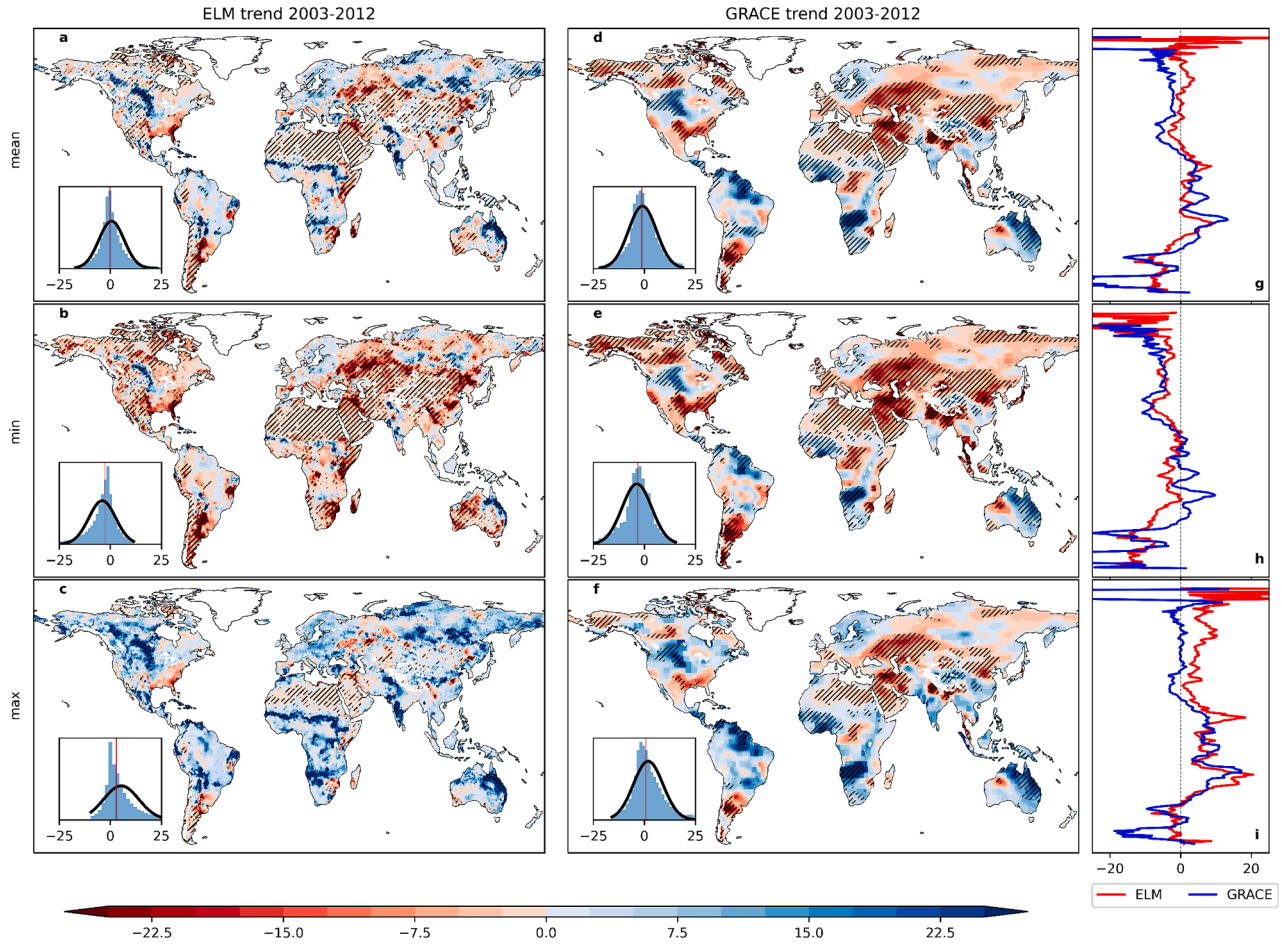


Fig. 1. Spatial distribution of linear trends in annual average TWS during 2003–2012 (mm/year). (a–c) Gridded distribution of ELM-simulated linear trends under the S6 forcing. (d–f) Gridded distribution of GRACE observed linear trends. (g–i) Latitudinal average trends of the ELM simulations and GRACE observations. In (a–f), inset plots show the probability density distribution of the displayed gridded values over the globe, with vertical red line indicating the median. In (a, d), the trends were calculated on the simple arithmetic average ("mean") of the ELM (CN-GSWP3, CN-CRUNCEP, CN-PRINCETON, CNP-GSWP3, CNP-CRUNCEP, CNP-PRINCETON) or GRACE ensemble members (CSR RL06 v2, JPL RL06.1 v3 CRI, GSFC RL06 v2), and stippling means where the trends were significant at $p \leq 0.05$. In (b, c, e, f), the trends were calculated on the individual ELM or GRACE members, and the minimum ("min") or maximum ("max") displayed, alone with whether the minimum or maximum values were statistically significant at $p \leq 0.05$.

addition to Pearson correlations, we evaluated ELM v1 using root mean squared error minus the standard deviation of the ELM detrended anomalies ("RMSE – ELM std", second column of Fig. S4). When the metric is negative, it means the detrended ELM anomalies is more similar to the observation than to its average, and we interpret such as satisfactory ELM v1 performance. ELM v1 performance is more satisfactory compared to GRACE-REC than GRACE observations, and least satisfactory compared to Li-REC (Fig. S4d–h). The worst performance was seen in the South American tropical forests, central Africa, and parts of northern high latitudes. In the arid regions of northern Africa and central Asia, the metric was around zero. In most of the mid-latitude regions and the high-latitude regions of Eurasia, the metric was generally negative.

3.1.2. Biome and climate zone level evaluation of TWS

In addition to the grid-level evaluation, we performed the evaluation at biomes (Fig. S1) and Köppen-Geiger climate zones (Fig. S2) levels in order to better match the spatial scales of the ELM v1 simulations and GRACE observations (Section 2.2).

Upon examining biome-average TWS (Fig. 2a), both ELM v1 simulations and GRACE observations showed significant positive trends in the South American tropical forests (except for elm-CRUNCEP) and Tropical savannahs & grasslands, insignificant positive trends in Asian tropical forests (except for elm-GSPW3), and significant negative trends in Temperate grasslands & shrublands. In Wetlands, ELM simulations exhibited significant positive trends, and the GRACE trend was insignificant but had comparable magnitude. ELM simulations had insignificant negative trends in Temperate forests, while GRACE has significant negative trends that were stronger. ELM simulations demonstrated insignificant positive trends but were in contrast to GRACE's insignificant negative trends in African tropical forests and Boreal forests. ELM simulations displayed partially insignificant or significant trends in Croplands, Deserts and Tundra, whereas GRACE indicated significant negative trends in these biome regions (Fig. 2a). For the climate zone-average TWS, the signs and significances of the trends of the ELM simulations and GRACE observations were fairly consistent in all the tropical climate zones, three of the arid climate zones (BWk, BSh and BSk), three out of six temperate climate zones (Csb, Cfa, Cfb), and two out of eight cold climate zones (Dwa, Dwb) (Fig. 2b). It's noteworthy that ELM v1 simulations tended to exhibit more positive trends in colder climate zones, while GRACE observations revealed more negative trends (Fig. 2b). Overall, the aggregation of data at biome and climate zone levels helped mitigate model errors in the tropical and temperate regions. However, model errors were still large in the crop-dominated and boreal regions, supporting the missing human water withdrawals and missing glaciers/simple snow processes explanations (Section 4.1).

The Pearson correlations between the detrended biome-average and climate zone-average TWS anomalies of ELM simulations and GRACE

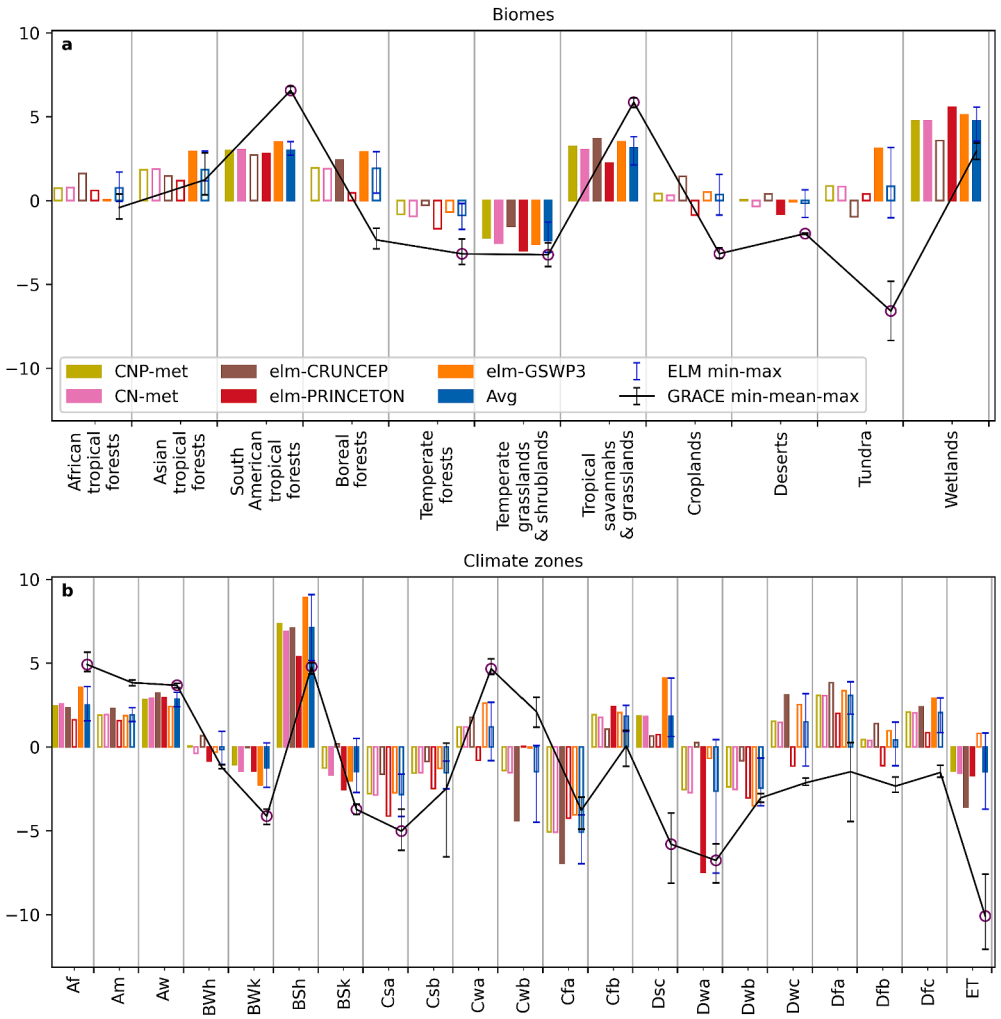


Fig. 2. Comparison of ELM-simulated and GRACE observed trends in biome- and climate zone-level annual average TWS over 2003–2012. The ELM simulations were under the S6 forcing. The trends were calculated on the simple arithmetically averaged TWS time series over different subsets of ELM ensemble members (CNP-met, CN-met, elm-CRUNCEP, elm-PRINCETON, elm-GSWP3, Avg; see Section 2.3 for their definitions) or over all three GRACE members (CSR RL06 v2, JPL RL06.1 v3 CRI, GSFC RL06 v2). Solid bars for ELM and purple circles for GRACE indicate statistically significant trends at $p \leq 0.05$. Error bars show the minimum and maximum range over the ELM or GRACE ensemble members. Spatial patterns of the biomes and climate zones are in Figs. S1 and S2.

observations and reconstructions were mostly statistically significant and >0.6 (Figs. S5 and S7). The low correlations in the Deserts biome probably reflect the low correlations in the Sahara region (Fig. S4). The correlations between ELM and Li-REC were systematically lower than the other evaluation datasets, notable examples of which include the Temperate forests and Temperate grasslands & grasslands biomes (Fig. S5) and the BSk (Arid, steppe, cold) climate zone (Fig. S7). Globally, the Pearson correlations between ELM and Li-REC were ≥0.4 (≥0.6) in only ~40 % (<20 %) of the grids, compared to ≥68 % (>40 %) for the other evaluation datasets (Fig. S4a–e). The problem may be caused by low interannual variability in the Li-REC dataset, which is lower by about 2–5 mm than GRACE, ELM, or GRACE-REC across all biomes and climate zones (Fig. S9). The root mean squared errors between the ELM-simulated detrended TWS anomalies and the various observed or reconstruction datasets including GRACE, Li-REC and GRACE-REC were smaller than the interannual variability of ELM over most of the biome regions and climate zones, while were greater in South American tropical forests and Deserts (except for Li-REC) and arid climate zone BWk (Figs. S6 and S8).

3.2. Regional relative contributions to TWS trends by different forcings

Given the results of trend- and Pearson correlation-evaluation of ELM at biomes and climate zone levels, we deemed the performance of ELM satisfactory if the trends agree in sign, the Pearson correlations were always higher than 0.3, and the root mean squared error was always approximately equal to or smaller than ELM's standard deviation. We chose a low value for the Pearson correlation criterion mainly to accommodate the evaluation results of Li-REC (see last paragraph above). Based on this criterion, ELM performance was satisfactory in four of the biomes (Temperate forests, Temperate grasslands & shrublands, Tropical savannahs & grasslands, and Wetlands) and seven climate zones (Aw, BSh, Csa, Csb, Cfa, Dwa, Dwb). We still performed the decomposition analysis for all the biomes and climate zones to reveal the global pattern but used shading to denote where ELM performed less well, i.e., the results were less reliable. ELM may be satisfactory for the purpose of the study in slightly more regions than the evaluation results suggest, since some of the discrepancies on trend would be caused by human land–water management practices, which this study is not investigating. Good ability to reproduce the detrended anomalies per se lends confidence to ELM's precipitation-evapotranspiration-runoff processes, which are the same for the detrended anomalies and the climate-driven long-term trends, i.e., the S2–S1 component (Table 1). We also note the TWS in this section means end-of-year TWS from the ELM simulations, for reasons given in the Methods Section 2.1.

During 1948–2012, the total effects of all the forcings was an

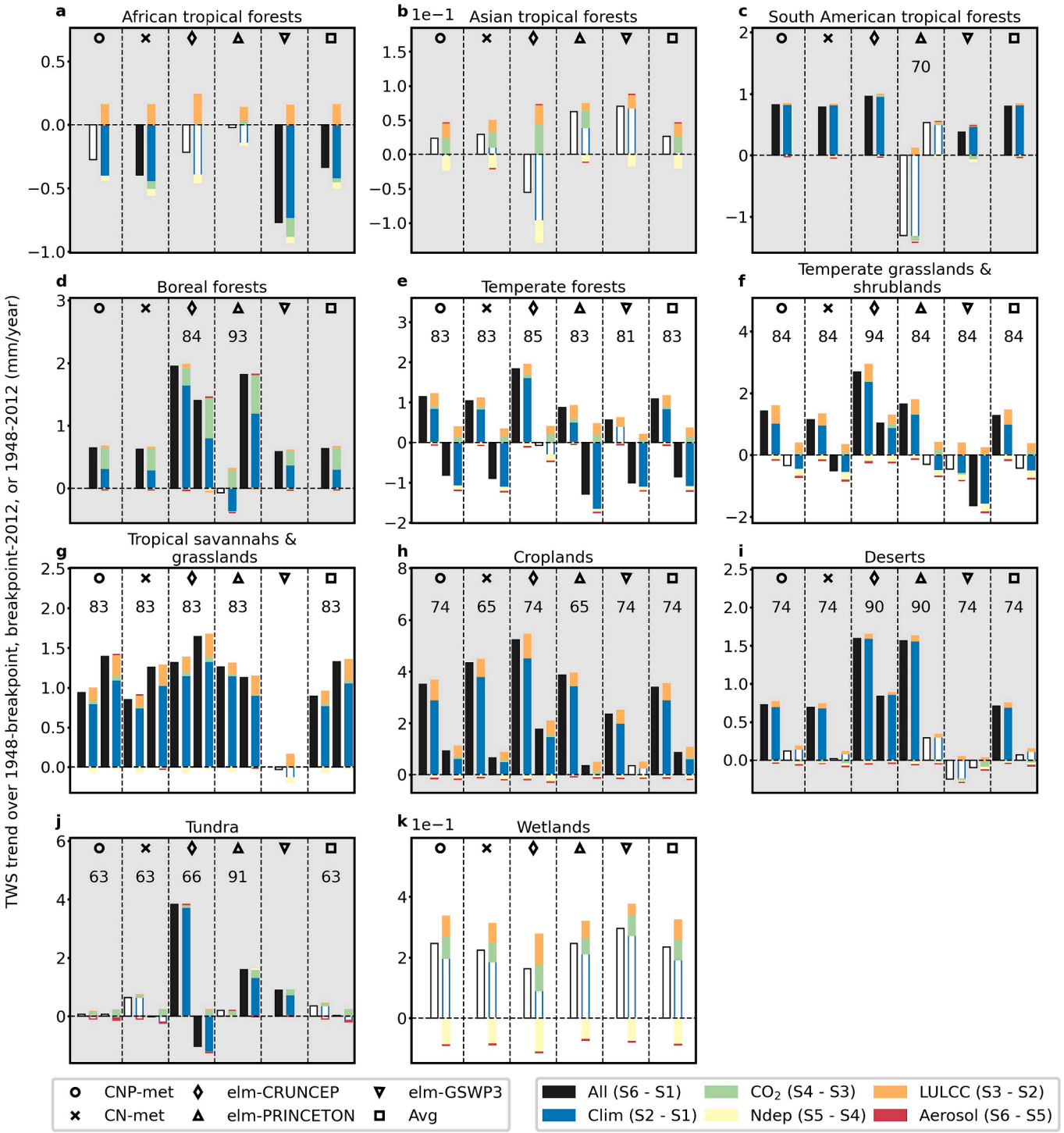


Fig. 3. Trends (unit: mm/year) in biome-average TWS contributed by various forcings. Vertical dashed lines in each subplot divide up different ways of averaging the ELM simulations (CNP-met, CN-met, elm-CRUNCEP, elm-PRINCETON, elm-GSWP3, Avg). If a breakpoint was detected in the trends of a biome according to one way of averaging, the year of the breakpoint is indicated by two digits "XX" (meaning the year 19XX) above the bars, and trends are displayed for both 1948–breakpoint (the left two bars) and breakpoint–2012 (the right two bars). If breakpoint was not detected, the displayed trends are for 1948–2012. Note the y-scales are different between the subplots. Filled bars indicate that the trends were statistically significant ($p \leq 0.05$), and hollow bars indicate that the trends were insignificant.

increase in TWS in all the biomes and climate zones except the African tropical forests (Fig. 3, Fig. S11). In the Temperate forests, the total forcings induced increase in TWS from 1948 to the 1980s, and then induced decrease (Fig. 3), but the overall effect was still to increase (Fig. S10e). Nonlinear trends existed in many biomes (Fig. 3dfghij) and climate zones (Fig. S11defghjopt), with the breakpoints generally occurring between 1960 and 1990, and the trends being more positive before the breakpoints than afterwards. Fig. S12 further illustrates those nonlinear trends via global maps of TWS trends calculated over multiple periods that were selected based on the breakpoints in Fig. 3 and Fig. S11. Hotspots of nonlinear trends can be seen over (1) the western-southeastern U.S. and the croplands of South America, where the trends turned from slightly positive during 1948–2012 to strongly negative in 1993–2012, (2) eastern Australia and the croplands-grasslands in east

Asia, where the trends became more negative over time, (3) the savannahs of Africa and the boreal forests in Siberia, where the trends turned from slightly negative during 1948–2012 to strongly positive in 1993–2012. Those hotspots no doubt drove the biome and climate-zone level changes seen in Fig. 3 and Fig. S11.

Among the individual environmental forcings, climate change (S2–S1) had the largest effects on the TWS trends in all the biomes and climate zones except perhaps the Asian tropical forests and a few smaller climate zones (Csb, Cwb, Dsc, and Dwc) (Fig. 3, Fig. S11; see Fig. S2 for the full name and locations of the climate zones). The climate change effects generally had the same signs as the total effects of all the forcings. LULCC (S3–S2) was the second most influential factor for TWS trends in eight out of eleven biomes (excluding South American tropical forest, Boreal forests, and Tundra) and sixteen out of twenty-one climate zones (excluding Dsc, Dwc, Dfa, Dfc and Polar Tundra [ET]), and the induced trends were generally positive. Those positive contributions coincided with a general trend of croplands replacing trees across over 1948–2012 (Fig. S13). In the three exception biomes, the land cover changes were either a replacement of trees by grass or negligible (Fig. S13a). In the five exception climate zones, the land cover changes were negligible, a replacement of croplands by trees, grasses, or shrub, or a replacement of mainly grasses by crops (Fig. S13b). The physiological and phenological effect of CO_2 (S4–S3) tended to induce non-negligible trends in TWS in the more humid biomes and climate zones at all latitudes. Those trends were positive in the Asian tropical forests, Temperate forests, Boreal forests, Wetlands biomes (Fig. 3), and the Aw, Cwb, Cfb, Dsc, Dwc, Dfb, and Dfc climate zones (Fig. S11), and were negative in the African tropical forests biome (Fig. 3), and the Am and BSk climate zones (Fig. S11). Moreover, atmospheric CO_2 concentration was the second most influential factor for TWS trends over these regions. Nitrogen deposition (S5–S4) and aerosol deposition (S6–S5) generally had small and negative impacts across the biomes and climate zones and their relative contributions were regional dependent as well. However, nitrogen deposition still had non-negligible impacts in four out of eleven biomes: African tropical forests, Asian tropical forests, Temperate grasslands & shrublands (Fig. 3) and four out of twenty-one climate regions (Aw, BSk, Cwb and Dwc, Fig. S11). Aerosol deposition has the smallest negative impacts across the biomes and climate zones. The aerosol deposition factor effects on simulated TWS driven by meteorological and other environmental are mainly due to its influences on the snow albedos, which we can see it has more impacts on the cold climate zones than other regions (Fig. S11, BSk, BWk and Polar Tundra [ET]).

We examined the uncertainty in model configurations and meteorological forcings by performing the decomposition separately on the average ELM simulations over all the meteorological forcings for each configuration (CNP-met, CN-met), the average ELM simulations over the two configurations for each meteorological forcing (elm-CRUNCEP, elm-PRINCETON, elm-GSWP3), and the overall average ELM simulations (Avg). The two model configurations generated trends with the same signs and similar magnitudes except in the tundra region (the Tundra biome and the Polar Tundra [ET] climate zone) (Fig. 3, Fig. S11), where plant growth is slightly more limited by phosphorus than nitrogen (Yang et al., 2023). The uncertainty in meteorological forcings were larger. The CRUNCEP and PRINCETON forcings resulted in insignificant trends in Asian tropical forests and South American tropical forests, respectively (Fig. 3bc), and the GSWP3 forcing resulted in insignificant or negative trends in the Tropical savannahs & grasslands, Temperate grasslands & shrublands, and Deserts biomes (Fig. 3fgi), as opposed to positive trends by other forcings. Similarly, large differences across the meteorological forcings were seen in one tropical forest (Am), two arid desert (BWh, Bwk), two small temperate (Cwa, Cwb), and one cold and dry (Dsc) climate zones (Fig. S11). These results indicate meteorological forcing uncertainties were greater in the tropical and dry climates.

3.3. Regional relative contributions to TWS trends by different meteorological drivers

We used the end-of-year TWS to decompose the trends into contributions from climatological imbalance between P, E, Q during the study period, the linear trends in P, E, and Q, and accumulated random anomalies following Eq. (4) in Methods Section 2.3. In displaying the decomposition results, we focused on the climatological imbalance term $\overline{P}-\overline{E}-\overline{Q}$, the trend imbalance term $\frac{\beta_P-\beta_E-\beta_Q}{2}$, and the accumulated random anomalies term, $\frac{\sum_{t=1}^{T}\left(\sum_{i=1}^{t}\varepsilon_i\right)\left(t-\frac{T+1}{2}\right)}{\sum_{t=1}^{T}\left(t-\frac{T+1}{2}\right)^2}$, rather than the individual P-, E-, and Q-related terms in Eq. (4) for a few reasons. The P-related terms generally contributed to TWS trends in the opposite direction to the E- and Q-related individual terms, and an imbalance term between the three were generally a few magnitudes smaller than the individual terms. As such, it is impossible to tell whether the effect of P or the effects of E or Q were dominant from a plot showing only the individual terms. On the other hand, displaying only the imbalance term did not result in much loss of information. In the term $\overline{P}-\overline{E}-\overline{Q}$, it is clear that $\overline{P}$ could only contribute positively, and $\overline{E}$ and $\overline{Q}$ negatively, so that a positive imbalance means $\overline{P}$ was the dominant factor, and a negative imbalance $\overline{E}$ and $\overline{Q}$. In the term $\frac{\beta_P-\beta_E-\beta_Q}{2}$, we found β_P, β_E, and β_Q to be generally statistically insignificant across the biomes and climate zones (not shown), and the trend imbalance itself generally unimportant compared to the climatological imbalance or accumulated random anomalies (Fig. 4, Fig. S14). For the term $\frac{\sum_{t=1}^{T}\left(\sum_{i=1}^{t}\varepsilon_i\right)\left(t-\frac{T+1}{2}\right)}{\sum_{t=1}^{T}\left(t-\frac{T+1}{2}\right)^2}$, we simply interpret it as resulting from factors not accounted by climatology or linear trends, which may include not only random noise but also teleconnection-induced low frequency variabilities. The mechanisms behind the individual contributions from P, E, and Q to the accumulated random anomalies term cannot be distinguished without further analysis beyond the scope of this study.

The contributions from the climatological imbalance term were positive in most biomes, meaning the water inputs via P were greater than water losses via E and Q During the entire time period of 1948–2012. The contributions were small or uncertain in the South American tropical forests, meaning P, ET, and Q approximately balance out each other. The contributions were negative in the Temperate forests and Temperate grasslands & shrublands biomes during the post-breakpoint periods, and negative in the Wetlands biome, meaning water losses via E and Q were greater (Fig. 4). The signs of the contributions from the climatological imbalance term were consistent with the signs of TWS trends over the same biome and time period except in the African tropical forests, South American tropical forests, and Wetlands (Fig. 4).. It is notable that the nonlinearity in TWS trends were matched by changing signs or magnitudes of the $\overline{P}-\overline{E}-\overline{Q}$ term (Fig. 4d–j). To further understand the changes in $\overline{P}-\overline{E}-\overline{Q}$, we compared the ratios $(E+Q)/P$, E/P, and Q/P between the post-breakpoint periods (1965–2012, 1974–2012, 1983–2012, 1993–2012) and the whole-period average (1948–2012) (Fig. S15). Among the hotspots of nonlinear changes (see the description in section 3.2 and Fig. S12), the western-southeastern U. S., croplands of South America, eastern Australia, and croplands-grasslands in east Asia all saw increasing $(E+Q)/P$, and the savannahs of Africa and the boreal forests of Siberia saw decreasing $(E+Q)/P$. The two components E/P and Q/P evolved in opposite directions globally (Fig. S15). As such, the increasing $(E+Q)/P$ in the western U.S., eastern Australia, and the croplands-grasslands in east Asia were driven by increasing E/P. The increasing $(E+Q)/P$ in the southeastern U.S. and the croplands of South America were driven by increasing Q/P. The decreasing $(E+Q)/P$ in the savannahs of Africa were driven by decreasing Q/P, and in the boreal forests of Siberia by decreasing E/P.

The accumulated random anomalies term had opposite contributions to that of the climatological imbalance, i.e., negative contributions to

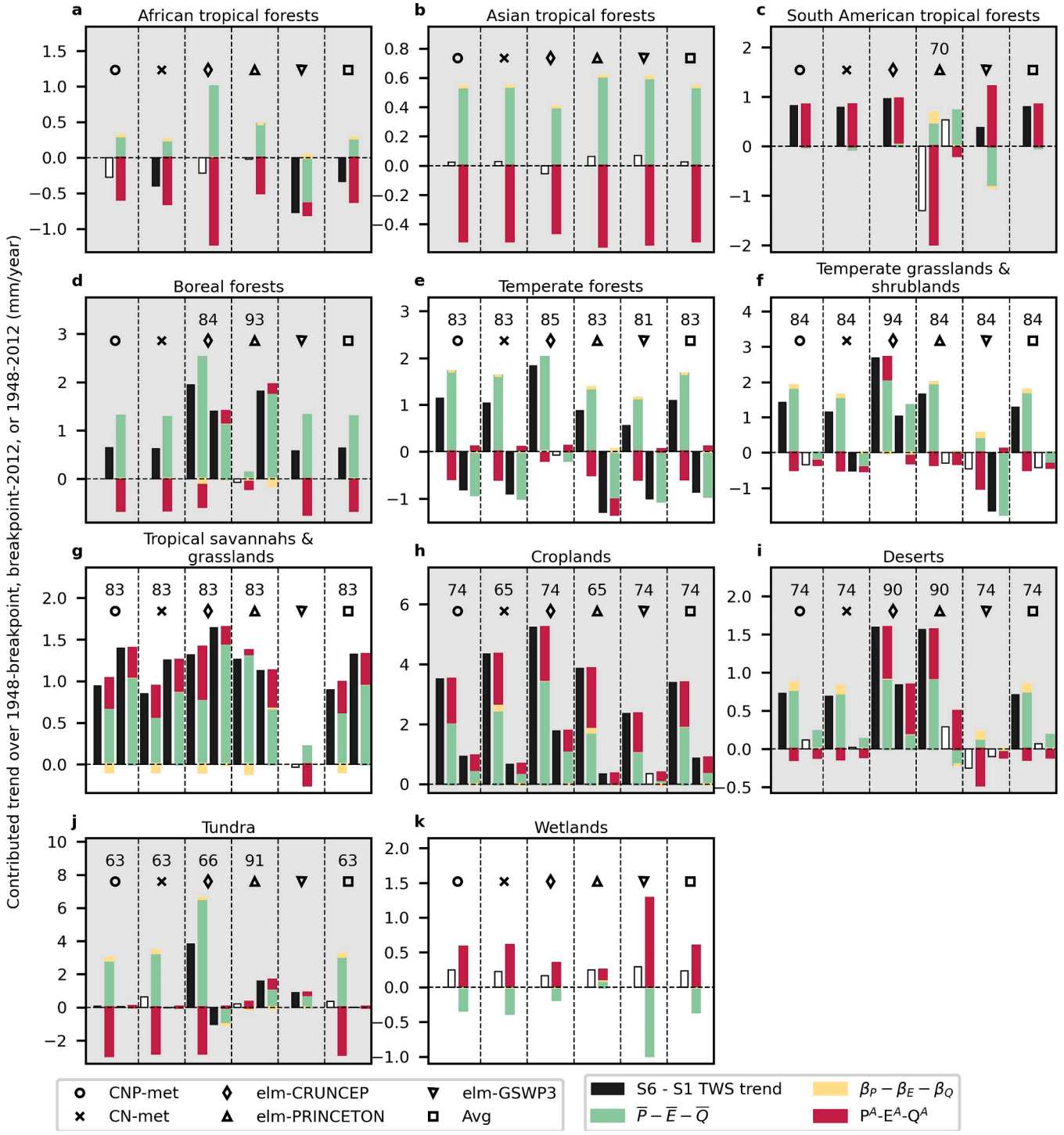


Fig. 4. Decomposition of biome-average TWS trends driven by all forcings ("S6–S1 TWS trend") into contributions by the climatological imbalance between P, E, and Q ($\overline{P}-\overline{E}-\overline{Q}$), the trends in P, E, and Q ($\beta_P-\beta_E-\beta_Q$), and accumulated random anomalies ($P^A-E^A-Q^A$). Vertical dashed lines in each subplot divide up different ways of averaging the ELM simulations (CNP-met, CN-met, elm-CRUNCEP, elm-PRINCETON, elm-GSWP3, Avg). If a breakpoint was detected in the TWS trends of a biome according to one way of averaging, the year of the breakpoint is indicated by two digits "XX" (meaning the year 19XX) above the bars, and trends are displayed for both 1948–breakpoint (the left two bars) and breakpoint–2012 (the right two bars). If breakpoint was not detected, the displayed trends are for 1948–2012. Note the y-scales are different between the subplots.

TWS trends, in most of the forest biomes, Temperate grasslands & shrublands, Tundra, and the Deserts with some uncertainty (Fig. 4). Among those biomes, only African tropical forests saw negative TWS trends. Therefore, the climatological imbalance was the dominant driver of TWS trends in most biomes except African tropical forests. In the South American tropical forests and Wetlands, the accumulated random anomalies term was the dominant driver of TWS trends. In the Tropical savannahs & grasslands and Croplands biomes, contributions from the accumulated random anomalies were in the same direction as the climatological imbalance and positive, and both terms contributed similar magnitudes.

For the climate zone-average TWS trends, the accumulated random

anomalies term was the dominant driver of positive TWS trends in the three tropical climate zones (Af, Am, Aw) (Fig. S14, Fig. S3). In the non-tropical climate zones, the climatological imbalance almost always contributed positively to TWS trends, and among those climate zones, the accumulated random anomalies term mostly contributed in the opposite direction (BWh, Bwk, Csb, Cwb, Cfa, Cfb, Dsc, Dwc, Dfc, and Polar Tundra [ET]) (Fig. S14). In the cold arid steppe (BSk) and a cold humid climate zone (Dfb), the accumulated random anomalies term had synergistic but smaller contribution than the climatological imbalance (Fig. S14gs). In the remaining climate zones (BSh, Csa, Cwa, Dwa, Dwb, Dfa), the accumulated random anomalies had synergistic and similar or greater magnitudes of contributions to the climatological imbalance (Fig. S14). Overall, the dominant and primarily positive contributions from $\overline{P}-\overline{E}-\overline{Q}$ and the changing contributions before and after breakpoints at the climate zone level were consistent with the biome-level results.

The large contributions from the accumulated random anomalies in both biome and climate zone level results may be partially due to chance, since unit root processes can exhibit unpredictable large trends (Dickey and Fuller, 1979), and partially due to persistent imbalances between P, E, and Q at the decadal to multidecadal scale caused by teleconnections.

4. Discussion

4.1. ELM performance on TWS trends and variabilities

The agreement between ELM simulated and GRACE observed trends in TWS were low in the northern high latitudes, a few arid or heavily irrigated mid-latitude regions, and a small region in central Africa (Fig. 1). To better understand the causes of these differences, we compared the trends in a few major components of TWS (soil moisture, snow water equivalent, the sum of soil moisture and groundwater) between ELM and reanalysis sources (Section 2.2). In the northern high latitudes (Alaska, western Canada, and eastern Siberia), both positive and negative trends existed in the reanalysis soil moisture and snow water equivalent, and there was no definite evidence that ELM was biased (Fig. S16 and 17). However, ELM predicted snow water equivalent had very small contributions to TWS trends (Fig. S17), which may be caused there is no active ice-sheet, and its snowpack depth was limited to one meter snow-water-equivalent in ELM v1. Schneider et al. (2022) reported that ELM v1 was exaggerated by too quick runoff of surface melt due to the absence of pore-space for meltwater storage in the shallow snowpack, and by unrealistic snow-darkening. In central Asia and southeastern Russia, the reanalysis soil moisture trends were more strongly negative than ELM trends (Fig. S16). Therefore, inaccurate soil moisture simulations may have contributed to ELM biases in TWS in these regions (Fig. 1). In Syria-Iraq-Iran, northern India, the North China Plain, and the center of Africa, ELM simulated trends were only clearly different from the reanalysis in the sum of soil moisture and groundwater (Fig. S18), and not in soil moisture (Fig. S16). In the first three regions, ELM biases (Fig. 1) were likely caused by a lack of representation of dams and reservoirs operations and human withdrawal of groundwater, especially for irrigation (Fig. S3; FAO, 2013; Voss et al., 2013). A new irrigation scheme simulates irrigation water demand and applies irrigation water in E3SM Land Model version 2 (ELM v2), which is coupled to a river routing model and a water management model (MOSART-WM) that simulate streamflow, reservoir operations, and irrigation water supply. With this new two-way coupling, surface water irrigation is constrained by the available runoff, streamflow, and reservoir storage, ELM is more realistically accounting for the interactions between human water use and the terrestrial water cycle in E3SM (Zhou et al., 2020). Addressing the lack of the representation of water withdrawal and dam construction in the ELM model requires a multi-faceted approach involving data integration, model coupling, scenario analysis, and interdisciplinary collaboration. Future studies should aim to improve the accuracy of TWS trend simulations by considering these factors and their interactions with climate and land surface processes. In central Africa, there were few dams and reservoirs or irrigated area (Fig. S3). Therefore, the errors may be caused by similar reasons to previous findings in the Amazon rainforests, such as, the lack of deep roots, inaccurate soil texture and pedotransfer function, inaccurate meteorological forcing, inaccurate groundwater dynamics (Tang et al., 2015).

The disagreement in detrended TWS anomalies between ELM and the GRACE observations or various GRACE-based reconstructions occurred mainly in the South American tropical forests and a large area spanning the Sahara and central Africa (Fig. S4). One general source of error in ELM may be excessive response to large wetting events that is caused by inaccurate soil depth; spatially varying soil depth has not been introduced into ELMv1.0 (Swenson and Lawrence, 2015). Additionally, in the arid Sahara and semi-arid Sahel regions, the low Pearson correlations between ELM simulations and the observations/reconstructions may be because the uncertainty in meteorological drivers is high (Swenson and Lawrence, 2015; Wang et al., 2022), and low signal-to-noise ratios (Long et al., 2015). This argument is supported by the fact that ELM compared the best against GSWP3 forced GRACE-REC, which shared meteorological forcing with some of the ELM v1 simulations (Section 2.1). Another reason may be ELM behaves like its predecessor CLM version 4.5 which exhibits biases in evapotranspiration and total water storage. These biases are consistent with excessive soil evaporation when canopy is sparse or absent, which reduced moisture inputs into the ground (Swenson and Lawrence, 2014). They improved the model simulated soil evaporation by replacing CLM's existing empirical soil resistance parameterization with a more mechanistically based formulation in which soil evaporation is controlled by the rate of diffusion of water vapor through a dry surface layer. This improvement hasn't been implemented to the ELM v1. In the African and South American tropical forests, the errors may be caused by similar reasons to the errors in trend (see the end of above paragraph) (Tang et al., 2015). In addition, ELM v1 uses soil moisture stress (SMS) parameterization to determine the regulation of stomatal conductance in response to soil water dynamic which may overestimate dry season water stress in tropical forests and contribute to the bad performance of these regions (Kennedy et al., 2019). A more mechanistic plant hydraulics (PHS) representation of stress and vegetation water use dynamic scheme has been implemented into the ELM v2, and the dry season water stress issue has been improved (Fang and Leung, 2022).

It is also clear from Fig. 1, Fig. S4, and Fig. S16–S18 that observational, reconstruction, and reanalysis data still contained considerable uncertainty in TWS and its components. The sources of uncertainty in GRACE observations included lower signal-to-noise ratios in the drier regions, leakage errors alone coastlines, parameterization errors, and large spatial footprints (Hasan and Tarhule, 2021; Long et al., 2015; Wiese et al., 2016). A previous assessment of uncertainty in GRACE/GRACE-FO TWS trends were as high as >14 mm/year along the Gulf of Alaska, and 4–10 mm/year in the Middle East and northern India regions (Hasan and Tarhule, 2021), suggesting that observational uncertainty may have contributed partially to bad ELM performances in those regions (Fig. 1). The sources of uncertainty in GRACE/GRACE-FO based TWS reconstruction include lack of explicit handling of glacier melting effects and human land–water management impacts, which are similar to the drawbacks of models, and potential underestimation of interannual variabilities (Humphrey and Gudmundsson, 2019; Li et al., 2021). Since reanalysis data involve the use of land surface models, they may be affected by similar process deficiencies and uncertainties in the input data as ELM (Gelaro et al., 2017; Rui and Beaudoing, 2020).

4.2. Contributions to TWS trends by forcings and meteorological drivers

Factorial simulations (Table 1) indicated that the climate forcing was

the most dominant contributor to the TWS trends in all the biomes and climate zones except for the Asian tropical forests and a few smaller climate zones (Fig. 4 and Fig. S11). In regions with intense human water withdrawal or dam constructions, the human land–water management practices may be a more important factor than the climate forcing, but comparison between that and the environmental forcings (climate, CO_2, LULCC, nitrogen deposition, aerosol deposition) are beyond the scope of this study. Although the TWS trends were mainly positive over the entire study period of 1948–2012, there were tendencies towards less positive or even negative trends during the second half of the study period (post-1960s to post-1980s), which may be related to declining TWS under climate change (Pokhrel et al., 2021).

We also decomposed contributions to TWS trends by each meteorological driver as the sum of its mean value over the time period of interest, a linear trend, and remainder anomalies (Eq. (4)). Interestingly, climatological imbalance between P, E, Q and accumulated random anomalies contributed more to the long-term trends in TWS than the linear trends in P, E, and Q (Fig. 4 and Fig. S14). The failure of the trends in P and atmospheric evaporative demands to explain the trends in TWS is further supported by previous studies (Chang et al., 2020; Rodell et al., 2018). This finding is easily understood from the fact that P, E, and Q are fluxes, and TWS is a storage term that reflects the cumulative effects of these fluxes. For example, when there is a negative trend in P-E-Q, the regional TWS can still increase if the P-E-Q at the beginning of the time period is a large positive value, and the negative trend has not caused the P-E-Q to become negative by the end of the time period. Over the study period, the ratios of $(E+Q)/P$ mainly increased, explaining the mainly negative turns in TWS trends, but the increases were not uniformly driven by E or Q across the globe (Fig. S15). Those heterogenous changes in E/P and Q/P may be mainly related to transpiration changes, followed by surface water, ground water, and soil moisture (Xiong et al., 2022b).

In the eight biomes where LULCC is the second largest contributor to TWS trends after the climate forcing, trees, grasses, and shrubs have been replaced by crops during our study period. These regions correspond well to regions that experienced LULCC during the late 20th-early 21st century (Winkler et al., 2021). Deforestation, conversion to croplands, and urbanization can directly affect the canopy component of TWS and alter the soil moisture and groundwater components of TWS by changing evapotranspiration. Interestingly, the net contributions of LULCC to TWS trends are generally positive (Fig. 3 and Fig. S11), perhaps due to a reduction in E and an increase in storage when deep-rooted forests are converted to shallow-rooted crops. However, if the anthropogenic water withdrawals that typically accompany these LULCC were to be represented in ELM, the effect of LULCC is likely to become more negative. In the biomes and climate zones where LULCC were not the second largest contributor, the land cover changes were either a replacement of crops by natural vegetation, replacement between vegetation of equal heights (grasses and crops), or negligible (Fig. S13).

The physiological and phenological effect of CO_2 (S4-S3) tended to induce non-negligible trends in TWS in the more humid biomes and climate zones at all latitudes and contributed to positive TWS trends in the most regions, except for African tropical forests and the Deserts biomes (Fig. 3), and the Am and BSk climate zones (Fig. S11). There are multiple opposing mechanisms through which increasing atmospheric CO_2 concentration can affect TWS. The physiological effects of elevated CO_2 concentrations induce the stomatal closure which reduces E per leaf area. However, the excess water may be drained as runoff instead of being retained as TWS. Conversely, CO_2 fertilization (phenological effect) generally increase E by increasing leaf area index (Gentine et al., 2019; Shi et al., 2013; Yang et al., 2021; Zhu et al., 2021). In most of the high- and mid-latitude regions, increasing CO_2 concentration generally increases evapotranspiration and decreases runoff. In the tropical regions, increasing CO_2 concentration generally decreases evapotranspiration and increases runoff. However, the effect of atmospheric CO_2 concentration on TWS ultimately depends on the local relative changes (Gentine et al., 2019; Yang et al., 2021).

Nitrogen deposition can benefit the growth of vegetation in nitrogen limited ecosystems, thereby increasing evapotranspiration and decreasing TWS (Mao et al., 2015; Zhu et al., 2021). This can help explain the general negative contribution of nitrogen deposition to TWS trends (Fig. 3, Fig. S11). However, tropical forests are more nitrogen-abundant than mid- and high-latitude ecosystems, and anthropogenic nitrogen deposition is less intense in the high-latitudes than the mid-latitudes (Ackerman et al., 2019; LeBauer and Treseder, 2008; Schulte-Uebbing and de Vries, 2018.). In addition, Mao et al. (2015) reported that nitrogen deposition induced higher LAI, leading to a rise in canopy evaporation within tropical evergreen broadleaf forest. This increase in canopy evaporation effectively counterbalanced the reduction in soil evaporation resulting from the shading effects brought about by the higher LAI and even countered the nitrogen-induced increment in total E. These may explain the relatively stronger contribution of nitrogen deposition to TWS trends in temperate grasslands and shrublands and tropical savannahs, tropical forest and tropical Savanah climate zone (Aw) and three mid-latitude climate zones (Fig. 3, Fig. S11).

Aerosol forcing mainly affects the hydrological cycle in coupled Earth system model simulations by affecting radiative forcing and cloud properties (Liu et al., 2018). However, in the land-only simulations in this study, the climate forcing already encompasses these effects. In land-only simulations using ELM model, the aerosol deposition has been shown to induce a reduction in snow albedo, resulting in an increase in surface energy absorption. This leads to increased snow surface temperature, subsequently leading to increased rates of snowmelt, which may enhance the evaporation and runoff. Those effects may have impacts on TWS, particularly in regions characterized by extensive snow cover, such as Tibetan Plateau areas (Hao et al., 2023). These findings may explain the stronger aerosol deposition impacts over several climate zones (BWk, BSk and Polar Tundra [ET]), which include the Tibetan Plateau. However, it's important to note that the specific impacts of aerosol deposition on the snowpack and the surrounding environment can be influenced by various factors, including the type and concentration of aerosols, the size and shape of snow grains, and the local climate conditions. The effects can vary in different regions and under different circumstances (Dang et al., 2019; Hao et al., 2023, 2021). Overall, the aerosols contributed smallest to TWS trends compared to other factors as shown in Fig. 3, Fig. S11.

We found that long-term imbalance between P, E, and Q and time-varying E/P and Q/P ratios lead to long-term nonlinear trends in TWS (Fig. 4, Fig. S14-S15).Previous analyses of the effects of climate forcings on TWS trends have not extensively addressed those facts, although cumulative fluxes and lagged responses were sometimes considered (Rodell et al., 2018; Chang et al., 2020; Zhang et al., 2019b). In future studies, there is a need to better characterize and understand the mechanisms of long-term imbalance in P, E, and Q. For example, global warming or teleconnections-driven P changes are typically accompanied by changes in atmospheric evaporative demand, and can drive vegetation changes (Brandt et al., 2017). Snowmelt-supplied Q may increase at the early stages of warming, and then decrease as the snowpack is depleted. Anthropogenic factors like the CO_2 fertilization effects on vegetation growth and urbanization also affect the partitioning of P (Kumar et al., 2018; Yang et al., 2021). In the tropical regions, the greater influences from the anomalies term than elsewhere in the globe (Fig. 4, Fig. S14) may be because the influences from teleconnections were more prominent in those otherwise relatively stable ecosystems (Rifai et al., 2018).

There are several sources of uncertainty in the ELM-simulated TWS trends. The uncertainty in meteorological forcings is found to be important for TWS in this study (Fig. 3, Fig. 4, Fig. S12, Fig. S14). The signs of TWS trends varied between the three forcings datasets across some of biomes and climate zones. The large uncertainties occurred over Asian tropical forests, South American tropical forests, Tropical savannahs & grasslands, Temperate grasslands & shrublands, and Deserts

biomes (Fig. 3bcfgi). Moreover, large differences were also seen in one tropical forest (Am), two arid desert (BWh, Bwk), two small temperate (Cwa, Cwb), and one cold and dry (Dsc) climate zones (Fig. S12). These results indicate meteorological forcing uncertainties were greater in the less well-observed regions. The temperature variable in the CRUNCEP, PRINCETON, and GSWP3 forcings were all calibrated against the CRU meteorological dataset (Harris et al., 2014), but precipitation was calibrated against the CRU (Harris et al., 2014), the CRU and TRMM (Huffman et al., 2007), and the GPCC and GPCP (Huffman et al., 2009; Schamm et al., 2014) datasets, respectively (Kim et al., 2014; Viovy, 2009). The use of TRMM by the PRINCETON dataset may have made it more accurate than the other two datasets over the mission period of the TRMM satellite (1997–2015) (Huffman et al., 2007). The GPCC dataset appears to use more rain gauges than the CRU dataset over the African and tropical South American regions during 2000–2009 (Deutscher Wetterdienst, 2022; Harris et al., 2020). Therefore, despite the better agreement between the CRUNCEP and PRINCETON forcings, the trends under the GSWP3 forcing might be more accurate. Our two model configurations generated TWS trends with the same signs and similar magnitudes except in the tundra region, where plant growth is slightly more limited by phosphorus than nitrogen (Yang et al., 2023). They also reported that ELMv1-CNP had similar performance to ELMv1-CN on evapotranspiration, performed better on ecosystem respiration when evaluated using the GBAF dataset. ELM v1-CNP improves carbon and nutrient cycles, but more work is necessary to improve hydrology, especially in tropical systems where P is limiting (Yang et al., 2023). Therefore, we may investigate the water cycle using ELM CNP configuration with PHS scheme over the tropical regions for the future potential study. Accounting for human water withdrawals, especially surface and groundwater withdrawals for irrigation, will also be a valuable future pursuit, considering their importance to TWS (Rodell et al., 2018) and their ability to affect evapotranspiration, which is a vital variable in land surface modeling.

5. Conclusions

In summary, this study has comprehensively examined changes in TWS from 1948 to 2012. We utilized GRACE, GRACE-based reconstruction datasets, and 36 ensemble land surface model simulations to quantify the drivers of TWS changes and explore the influences of uncertainty in in meteorological forcings and model biogeochemistry configurations. Our analysis revealed that ELM effectively captured the global patterns of GRACE-observed TWS trends. It performed particularly well in four biomes – Temperate forests, Temperate grasslands & shrublands, Tropical savannahs & grasslands, and Wetlands – as well as in seven climate zones (Aw, BSh, Csa, Csb, Cfa, Dwa, Dwb), meeting criterions such as trends agreement, Pearson correlations exceeding 0.3 and root mean squared error within ELM's standard deviation. The TWS generally increased during the whole study period but increased less or decreased towards the second half of the study period (after the 1960–1980s). Hotspots of nonlinear TWS trends included the western-southeastern U.S., croplands of South America, eastern Australia, croplands-grasslands in east Asia, the savannahs of Africa, and the boreal forests of Siberia. Decomposition of the TWS trends by forcings showed that climate change had the largest effects on the TWS trends across all the biomes and climate zones, with some exceptions like the Asian tropical forests and a few smaller climate zones (Csb, Cwb, Dsc, and Dwc). LULCC had a significant impact on TWS trends in eight of the eleven biomes and sixteen of the 21 climate zones. These effects were generally positive and corresponded to a trend of croplands replacing trees over the study period. The influence of atmospheric CO_2 concentration was notable in more humid biomes and climate zones across different latitudes. It led to positive trends in Asian tropical forests, Temperate forests, Boreal forests, Wetlands biomes, and certain climate zones (Aw, Cwb, Cfb, Dsc, Dwc, Dfb, and Dfc), and negative trends in African tropical forests biome, and the specific climate regions (Am and BSk). Nitrogen deposition and aerosol deposition generally had smaller and negative impacts across the biomes and climate regions compared to climate change, LULCC and CO_2 concentration. However, they were not negligible in four of the eleven biomes and four of the twenty-one climate regions. Aerosol deposition had the smallest effect primarily on cold climate zones (BSk, BWk and Polar Tundra [ET]). we further found that the long-term average imbalance between P, E, and Q, along with the accumulated de-trended anomalies in P, E, and Q, were major contributors to TWS trends in all regions, while the trend-related term had a less significant impact. The changing imbalance between P, E, and Q were related to heterogeneous changes in E/P and Q/P. These findings underscore the need to consider precipitation partitioning and accumulated random anomalies, when explaining the effects of climate forcing on TWS trends. This study offers a valuable long-term perspective, revealing that TWS trends were less affected by internal climate variability than post-2002 GRACE satellite data (Rodell et al., 2018). Additionally, by explicitly decomposing the trends according to different forcings and the water balance equation, we contribute to a deeper understanding of TWS changes and their drivers.

6. Data availabXXXility

All the data used in this study are available from open-source repositories.

GRACE satellites TWS data: CSR RL06 v2: https://www2.csr.utexas.edu/grace/RL06_mascons.html, JPL RL06.1 v3 CRI: https://podaac.jpl.nasa.gov/dataset/TELLUS_GRAC-GRFO_MASCON_CRI_GRID_RL06.1_V3#, GSFC RL06 v2: https://earth.gsfc.nasa.gov/geo/data/grace-mascons.

GRACE-REC TWS data: https://doi.org/10.6084/m9.figshare.7670849.

Li-REC TWS data: https://datadryad.org/stash/dataset/doi:10.5061/dryad.z612jm6bt.

E3SM simulated dataset: CMIP6 repositories such as https://esgf-node.llnl.gov/projects/cmip6/, search for ELM simulations under the LS3MIP activity.

CRediT authorship contribution statement

Xiaoying Shi: Formal analysis, Supervision, Writing – original draft, Writing – review & editing. **Yaoping Wang:** Formal analysis, Methodology, Visualization, Writing – review & editing. **Jiafu Mao:** Conceptualization, Writing – review & editing. **Peter E. Thornton:** Writing – review & editing. **Daniel M. Riccuito:** Writing – review & editing. **Forrest M. Hoffman:** Resources, Writing – review & editing. **Yuefeng Hao:** Data curation, Writing – review & editing.

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.

Data availability

Data will be made available on request.

Acknowledgements

This work was supported by the Terrestrial Ecosystem Science Scientific Focus Area (TES SFA) project, the Reducing Uncertainties in Biogeochemical Interactions through Synthesis and Computing Scientific Focus Area (RUBISCO SFA) project, and the Energy Exascale Earth System Model (E3SM) project, funded through the Earth and Environmental Systems Sciences Division of the Biological and Environmental Research Office in the US Department of Energy (DOE), Office of

Science, Office of Biological and Environmental Research. Oak Ridge National Laboratory is supported by the Office of Science of the DOE under Contract DE-AC05-00OR22725.

Appendix A. Supplementary material

Supplementary material to this article can be found online at https://doi.org/10.1016/j.jhydrol.2024.131096.

References

Ackerman, D., Millet, D.B., Chen, X., 2019. Global estimates of inorganic nitrogen deposition across four decades. Glob. Biogeochem. Cycles 33, 100–107. https://doi.org/10.1029/2018GB005990.

Ahmed, M., Sultan, M., Wahr, J., Yan, E., 2014. The use of GRACE data to monitor natural and anthropogenic induced variations in water availability across Africa. Earth Sci. Rev. 136, 289–300. https://doi.org/10.1016/j.earscirev.2014.05.009.

Almendra-Martín, L., Martínez-Fernández, J., Piles, M., González-Zamora, Á., Benito-Verdugo, P., Gaona, J., 2022. Analysis of soil moisture trends in Europe using rank-based and empirical decomposition approaches. Glob. Planet. Change 215, 103868. https://doi.org/10.1016/j.gloplacha.2022.103868.

An, L., Wang, J., Huang, J., Pokhrel, Y., Hugonnet, R., Wada, Y., Cáceres, D., Müller Schmied, H., Song, C., Berthier, E., Yu, H., Zhang, G., 2021. Divergent causes of terrestrial water storage decline between drylands and humid regions globally. Geophys. Res. Lett. 48 https://doi.org/10.1029/2021GL095035.

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

Asoka, A., Mishra, V., 2020. Anthropogenic and climate contributions on the changes in terrestrial water storage in India. J. Geophys. Res. Atmos. 125 https://doi.org/10.1029/2020JD032470.

Asoka, A., Wardlow, B., Tsegaye, T., Huber, M., Mishra, V., 2021. A satellite-based assessment of the relative contribution of hydroclimatic variables on vegetation growth in global agricultural and nonagricultural regions. Geophys. Res. Atmos. 126 https://doi.org/10.1029/2020JD033228.

Beaudoing, H., Rodell, M., NASA/GSFC/HSL, 2019. GLDAS Noah land surface model L4 3 hourly 0.25 × 0.25 degree V2.0. Goddard Earth Sciences Data and Information Services Center (GES DISC), Greenbelt, MD, USA.

Beck, H.E., Zimmermann, N.E., McVicar, T.R., Vergopolan, N., Berg, A., Wood, E.F., 2018. Present and future Köppen-Geiger climate classification maps at 1-km resolution. Sci. Data 5, 180214. https://doi.org/10.1038/sdata.2018.214.

Beck, H.E., Pan, M., Miralles, D.G., Reichle, R.H., Dorigo, W.A., Hahn, S., Sheffield, J., Karthikeyan, L., Balsamo, G., Parinussa, R.M., Du, J., Kimball, J.S., Vergopolan, N., Wood, E.F., 2021. Evaluation of 18 satellite- and model-based soil moisture products using in situ measurements from 826 sensors. Hydrol. Earth Syst. Sci. 25, 17–40. https://doi.org/10.5194/hess-25-17-2021.

Brandt, M., Rasmussen, K., Peñuelas, J., Tian, F., Schurgers, G., Verger, A., Mertz, O., Palmer, J.R.B., Fensholt, R., 2017. Human population growth offsets climate-driven increase in woody vegetation in sub-Saharan Africa. Nat. Ecol. Evol. 1, 0081. https://doi.org/10.1038/s41559-017-0081.

Burrows, S.M., Maltrud, M., Yang, X., Zhu, Q., Jeffery, N., Shi, X., Ricciuto, D., Wang, S., Bisht, G., Tang, J., Wolfe, J., Harrop, B.E., Singh, B., Brent, L., Baldwin, S., Zhou, T., Cameron-Smith, P., Keen, N., Collier, N., Xu, M., Hunke, E.C., Elliott, S.M., Turner, A.K., Li, H., Wang, H., Golaz, J.-C., Bond-Lamberty, B., Hoffman, F.M., Riley, W.J., Thornton, P.E., Calvin, K., Leung, L.R., 2020. The DOE E3SM v1.1 biogeochemistry configuration: description and simulated ecosystem-climate responses to historical changes in forcing. J. Adv. Model Earth Syst. 12 https://doi.org/10.1029/2019MS001766.

Chang, L., Yuan, R., Gupta, H.V., Winter, C.L., Niu, G., 2020. Why is the terrestrial water storage in dryland regions declining? a perspective based on gravity recovery and climate experiment satellite observations and Noah land surface model with multiparameterization schemes model simulations. Water Resour. Res. 56 https://doi.org/10.1029/2020WR027102.

Chao, N., Jin, T., Cai, Z., Chen, G., Liu, X., Wang, Z., Yeh, P.-J.-F., 2021. Estimation of component contributions to total terrestrial water storage change in the Yangtze river basin. J. Hydrol. 595, 125661 https://doi.org/10.1016/j.jhydrol.2020.125661.

Cui, J., Yang, H., Huntingford, C., Kooperman, G.J., Lian, X., He, M., Piao, S., 2021. Vegetation response to rising CO_2 amplifies contrasts in water resources between global wet and dry land areas. Geophys. Res. Lett. 48, e2021GL094293 https://doi.org/10.1029/2021GL094293.

Dang, C., Zender, C.S., Flanner, M.G., 2019. Intercomparison and improvement of two-stream shortwave radiative transfer schemes in earth system models for a unified treatment of cryospheric surfaces. The Cryosphere 13, 2325–2343. https://doi.org/10.5194/tc-13-2325-2019.

DeGroot, M., Schervish, M., 2018. Probability and Statistics, fourth ed. Pearson Modern Classics for Advanced Statistics Series, Pearson.

Deutscher Wetterdienst, 2022. GPCC visualizer [WWW Document]. URL https://kunden.dwd.de/GPCC/Visualizer (accessed 6.28.22).

Dickey, D.A., Fuller, W.A., 1979. Distribution of the estimators for autoregressive time series with a unit root. J. Am. Stat. Assoc. 74, 427–431. https://doi.org/10.2307/2286348.

Fang, Y., Leung, L.R., 2022. Relative controls of vapor pressure deficit and soil water stress on canopy conductance in global simulations by an earth system model. Earth's Future 10, e2022EF002810. https://doi.org/10.1029/2022EF002810.

FAO, 2013. AQUASTAT – FAO's Global Information System on Water and Agriculture.

Fatichi, S., Leuzinger, S., Paschalis, A., Langley, J.A., Donnellan Barraclough, A., Hovenden, M.J., 2016. Partitioning direct and indirect effects reveals the response of water-limited ecosystems to elevated CO $_2$. Proc. Natl. Acad. Sci. USA 113, 12757–12762. https://doi.org/10.1073/pnas.1605036113.

Felfelani, F., Wada, Y., Longuevergne, L., Pokhrel, Y.N., 2017. Natural and human-induced terrestrial water storage change: A global analysis using hydrological models and GRACE. J. Hydrol. 553, 105–118. https://doi.org/10.1016/j.jhydrol.2017.07.048.

Forbes, W.L., Mao, J., Jin, M., Kao, S.-C., Fu, W., Shi, X., Riccuito, D.M., Thornton, P.E., Ribes, A., Wang, Y., Piao, S., Zhao, T., Schwalm, C.R., Hoffman, F.M., Fisher, J.B., Ito, A., Poulter, B., Fang, Y., Tian, H., Jain, A.K., Hayes, D.J., 2018. Contribution of environmental forcings to US runoff changes for the period 1950–2010. Environ. Res. Lett. 13, 054023 https://doi.org/10.1088/1748-9326/aabb41.

Forbes, W.L., Mao, J., Ricciuto, D.M., Kao, S., Shi, X., Tavakoly, A.A., Jin, M., Guo, W., Zhao, T., Wang, Y., Thornton, P.E., Hoffman, F.M., 2019. Streamflow in the Columbia River basin: Quantifying changes over the period 1951–2008 and determining the drivers of those changes. Water Resour. Res. 55, 6640–6652. https://doi.org/10.1029/2018WR024256.

Friedl, M., Sulla-Menashe, D., 2015. MCD12C1 MODIS/Terra+Aqua land cover type L3 global 0.05Deg CMG V006 . NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/MODIS/MCD12C1.006.

G3P, 2023. Global gravity-based groundwater product [WWW Document]. URL https://www.g3p.eu/ (accessed 10.2.23).

Gelaro, R., McCarty, W., Suárez, M.J., Todling, R., Molod, A., Takacs, L., Randles, C.A., Darmenov, A., Bosilovich, M.G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A.M., Gu, W., Kim, G.-K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J.E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S.D., Sienkiewicz, M., Zhao, B., 2017. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). J. Clim. 30, 5419–5454. https://doi.org/10.1175/JCLI-D-16-0758.1.

Gentine, P., Green, J.K., Guérin, M., Humphrey, V., Seneviratne, S.I., Zhang, Y., Zhou, S., 2019. Coupling between the terrestrial carbon and water cycles—a review. Environ. Res. Lett. 14, 083003 https://doi.org/10.1088/1748-9326/ab22d6.

Guo, Y., Gan, F., Yan, B., Bai, J., Xing, N., Zhuo, Y., 2022. Evaluation of terrestrial water storage changes and major driving factors analysis in Inner Mongolia. China. Sensors 22, 9665. https://doi.org/10.3390/s22249665.

Guo, L., Li, T., Chen, D., Liu, J., He, B., Zhang, Y., 2021. Links between global terrestrial water storage and large-scale modes of climatic variability. J. Hydrol. 598, 126419 https://doi.org/10.1016/j.jhydrol.2021.126419.

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

Hao, D., Bisht, G., Gu, Y., Lee, W.-L., Liou, K.-N., Leung, L.R., 2021. A parameterization of sub-grid topographical effects on solar radiation in the E3SM land model (version 1.0): implementation and evaluation over the Tibetan Plateau. Geosci. Model Dev. 14, 6273–6289. https://doi.org/10.5194/gmd-14-6273-2021.

Hao, D., Bisht, G., Rittger, K., Bair, E., He, C., Huang, H., Dang, C., Stillinger, T., Gu, Y., Wang, H., Qian, Y., Leung, L.R., 2023. Improving snow albedo modeling in the E3SM land model (version 2.0) and assessing its impacts on snow and surface fluxes over the Tibetan Plateau. Geosci. Model Dev. 16, 75–94. https://doi.org/10.5194/gmd-16-75-2023.

Harris, I., Jones, P.D., Osborn, T.J., Lister, D.H., 2014. Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 dataset. Int. J. Climatol. 34, 623–642. https://doi.org/10.1002/joc.3711.

Harris, I., Osborn, T.J., Jones, P., Lister, D., 2020. Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. Sci. Data 7, 109. https://doi.org/10.1038/s41597-020-0453-3.

Hasan, E., Tarhule, A., 2021. Comparison of decadal water storage trends from common GRACE releases (RL05, RL06) using spatial diagnostics and a modified triple collocation approach. J. Hydrol. X 13, 100108. https://doi.org/10.1016/j.hydroa.2021.100108.

Hastie, T., Tibshirani, R., Friedman, J., 2009. Model Assessment and Selection. In: Hastie, T., Tibshirani, R., Friedman, J. (Eds.), The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer Series in Statistics. Springer, New York, NY, pp. 219–259. https://doi.org/10.1007/978-0-387-84858-7_7.

Himanshu, S., 2020. CSR GRACE and GRACE-FO RL06 Mascon Solutions v02 [WWW Document]. URL https://www2.csr.utexas.edu/grace/RL06_mascons.html (accessed 9.20.23).

Huffman, G.J., Adler, R.F., Bolvin, D.T., Gu, G., Nelkin, E.J., Bowman, K.P., Hong, Y., Stocker, E.F., Wolff, D.B., 2007. The TRMM Multisatellite Precipitation Analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. J. Hydrometeorol. 8, 38–55. https://doi.org/10.1175/JHM560.1.

Huffman, G.J., Adler, R.F., Bolvin, D.T., Gu, G., 2009. Improving the global precipitation record: GPCP version 2.1. Geophys. Res. Lett. 36, L17808. https://doi.org/10.1029/2009GL040000.

Humphrey, V., Gudmundsson, L., 2019. GRACE-REC: a reconstruction of climate-driven water storage changes over the last century. Earth Syst. Sci. Data 11, 1153–1170. https://doi.org/10.5194/essd-11-1153-2019.

Jiménez Cisneros, B.E., Oki, T., Arnell, N.W., Benito, G., Cogley, J.G., Döll, P., Jiang, T., Mwakalila, S.S., 2014. Freshwater resources. In: Field, C.B., Barros, V.R., Dokken, D.

J., Mach, K.J., Mastrandrea, M.D., Bilir, T.E., Chatterjee, M., Ebi, K.L., Estrada, Y.O., Genova, R.C., Girma, B., Kissel, E.S., Levy, A.N., MacCracken, S., Mastrandrea, P.R., White, L.L. (Eds.), Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 229–269.

Jing, W., Yao, L., Zhao, X., Zhang, P., Liu, Y., Xia, X., Song, J., Yang, J., Li, Y., Zhou, C., 2019. Understanding terrestrial water storage declining trends in the Yellow River Basin. J. Geophys. Res. Atmos. 124, 12963–12984. https://doi.org/10.1029/2019JD031432.

Ju, J., Wu, C., Li, J., Yeh, P., Hu, B.X., 2023. Global evaluation of model agreement and uncertainty in terrestrial water storage simulations from ISIMIP 2b framework. J. Hydrol. 617, 129137 https://doi.org/10.1016/j.jhydrol.2023.129137.

Kennedy, D., Swenson, S., Oleson, K.W., Lawrence, D.M., Fisher, R., Lola Da Costa, A.C., Gentine, P., 2019. Implementing plant hydraulics in the community land model, version 5. J. Adv. Model. Earth Syst. 11, 485–513. https://doi.org/10.1029/2018MS001500.

Kim, H., Watanabe, S., Chang, E.-C., Yoshimura, K., Hirabayashi, Y., Famiglietti, J., Oki, T., 2014. Development of a new global dataset for offline terrestrial simulations for Global Soil Wetness Project Phase 3.

Kim, H., 2017. Global Soil Wetness Project Phase 3 Atmospheric Boundary Conditions (Experiment 1) [Data set]. https://doi.org/10.20783/DIAS.501.

Kumar, S., Moglen, G.E., Godrej, A.N., Grizzard, T.J., Post, H.E., 2018. Trends in water yield under climate change and urbanization in the US mid-Atlantic region. J. Water Resour. Plann. Manage. 144, 05018009 https://doi.org/10.1061/(ASCE)WR.1943-5452.0000937.

Langenbrunner, B., Pritchard, M.S., Kooperman, G.J., Randerson, J.T., 2019. Why does Amazon precipitation decrease when tropical forests respond to increasing CO_2 ? Earths Future 7, 450–468. https://doi.org/10.1029/2018EF001026.

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

LeBauer, D.S., Treseder, K.K., 2008. Nitrogen limitation of net primary productivity in terrestrial ecosystems is globally distributed. Ecol. 89, 371–379. https://doi.org/10.1890/06-2057.1.

Li, F., Kusche, J., Rietbroek, R., Wang, Z., Forootan, E., Schulze, K., Lück, C., 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, e2019WR026551 https://doi.org/10.1029/2019WR026551.

Li, F., Kusche, J., Chao, N., Wang, Z., Löcher, A., 2021. Long-term (1979-present) total water storage anomalies over the global land derived by reconstructing GRACE data. Geophys. Res. Lett. 48 https://doi.org/10.1029/2021GL093492.

Li, X., Long, D., Scanlon, B.R., Mann, M.E., Li, X., Tian, F., Sun, Z., Wang, G., 2022. Climate change threatens terrestrial water storage over the Tibetan Plateau. Nat. Clim. Chang. 12, 801–807. https://doi.org/10.1038/s41558-022-01443-0.

Liu, L., Shawki, D., Voulgarakis, A., Kasoar, M., Samset, B.H., Myhre, G., Forster, P.M., Hodnebrog, Ø., Sillmann, J., Aalbergsjø, S.G., Boucher, O., Faluvegi, G., Iversen, T., Kirkevåg, A., Lamarque, J.-F., Olivié, D., Richardson, T., Shindell, D., Takemura, T., 2018. A PDRMIP multimodel study on the impacts of regional aerosol forcings on global and regional precipitation. J. Clim. 31, 4429–4447. https://doi.org/10.1175/JCLI-D-17-0439.1.

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, 2574–2594. https://doi.org/10.1002/2014WR016853.

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

Loomis, B.D., Luthcke, S.B., Sabaka, T.J., 2019. Regularization and error characterization of GRACE mascons. J. Geod. 93, 1381–1398. https://doi.org/10.1007/s00190-019-01252-y.

Lv, M., Ma, Z., Yuan, N., 2021. Attributing terrestrial water storage variations across China to changes in groundwater and human water use. J. Hydrometeorol. 22, 3–21. https://doi.org/10.1175/JHM-D-20-0095.1.

Mao, J., Fu, W., Shi, X., Ricciuto, D.M., Fisher, J.B., Dickinson, R.E., Wei, Y., Shem, W., Piao, S., Wang, K., Schwalm, C.R., Tian, H., Mu, M., Arain, A., Ciais, P., Cook, R., Dai, Y., Hayes, D., Hoffman, F.M., Huang, M., Huang, S., Huntzinger, D.N., Ito, A., Jain, A., King, A.W., Lei, H., Lu, C., Michalak, A.M., Parazoo, N., Peng, C., Peng, S., Poulter, B., Schaefer, K., Jafarov, E., Thornton, P.E., Wang, W., Zeng, N., Zeng, Z., Zhao, F., Zhu, Q., Zhu, Z., 2015. Disentangling climatic and anthropogenic controls on global terrestrial evapotranspiration trends. Environ. Res. Lett. 10, 094008 https://doi.org/10.1088/1748-9326/10/9/094008.

Mo, S., Zhong, Y., Forootan, E., Mehrnegar, N., Yin, X., Wu, J., Feng, W., Shi, X., 2022. Bayesian convolutional neural networks for predicting the terrestrial water storage anomalies during GRACE and GRACE-FO gap. J. Hydrol. 604, 127244 https://doi.org/10.1016/j.jhydrol.2021.127244.

Mortimer, C., Mudryk, L., Derksen, C., Luojus, K., Brown, R., Kelly, R., Tedesco, M., 2020. Evaluation of long-term Northern Hemisphere snow water equivalent products. The Cryosphere 14, 1579–1594. https://doi.org/10.5194/tc-14-1579-2020.

Naha, S., Rico-Ramirez, M.A., Rosolem, R., 2021. Quantifying the impacts of land cover change on hydrological responses in the Mahanadi river basin in India. Hydrol. Earth Syst. Sci. 25, 6339–6357. https://doi.org/10.5194/hess-25-6339-2021.

Niu, G.-Y., Yang, Z.-L., Dickinson, R.E., Gulden, L.E., Su, H., 2007. Development of a simple groundwater model for use in climate models and evaluation with gravity recovery and climate experiment data. J. Geophys. Res.: Atmos. 112 https://doi.org/10.1029/2006JD007522.

Oleson, K.W., Lawrence, D.M., Bonan, G.B., Drewniak, B., Huang, M., Levis, S., Li, F., Riley, W.J., Swenson, S.C., Thornton, P.E., Bozbiyik, A., Fisher, R., Heald, C.L., Kluzek, E., Lamarque, F., Lawrence, P.J., Leung, L.R., Muszala, S., Ricciuto, D.M., Sacks, W., Sun, Y., Tang, J., Yang, Z.-L., 2013. Technical Description of Version 4.5 of the Community Land Model (CLM) (No. NCAR/TN-503+STR NCAR Technical Note). National Center for Atmospheric Research, Boulder, CO, USA.

Padrón, R.S., Gudmundsson, L., Decharme, B., Ducharne, A., Lawrence, D.M., Mao, J., Peano, D., Krinner, G., Kim, H., Seneviratne, S.I., 2020. Observed changes in dry-season water availability attributed to human-induced climate change. Nat. Geosci. 13, 477–481. https://doi.org/10.1038/s41561-020-0594-1.

Piao, S., Friedlingstein, P., Ciais, P., de Noblet-Ducoudré, N., Labat, D., Zaehle, S., 2007. Changes in climate and land use have a larger direct impact than rising CO_2 on global river runoff trends. Proc. Natl. Acad. Sci. USA 104, 15242–15247. https://doi.org/10.1073/pnas.0707213104.

Pokhrel, Y., Felfelani, F., Satoh, Y., Boulange, J., Burek, P., Gädeke, A., Gerten, D., Gosling, S.N., Grillakis, M., Gudmundsson, L., Hanasaki, N., Kim, H., Koutroulis, A., Liu, J., Papadimitriou, L., Schewe, J., Müller Schmied, H., Stacke, T., Telteu, C.-E., Thiery, W., Veldkamp, T., Zhao, F., Wada, Y., 2021. Global terrestrial water storage and drought severity under climate change. Clim. Change Nat. https://doi.org/10.1038/s41558-020-00972-w.

Rifai, S.W., Girardin, C.A.J., Berenguer, E., del Aguila-Pasquel, J., Dahlsjö, C.A.L., Doughty, C.E., Jeffery, K.J., Moore, S., Oliveras, I., Riutta, T., Rowland, L.M., Murakami, A.A., Addo-Danso, S.D., Brando, P., Burton, C., Ondo, F.E., Duah-Gyamfi, A., Amézquita, F.F., Freitag, R., Pacha, F.H., Huasco, W.H., Ibrahim, F., Mbou, A.T., Mihindou, V.M., Peixoto, K.S., Rocha, W., Rossi, L.C., Seixas, M., Silva-Espejo, J.E., Abernethy, K.A., Adu-Bredu, S., Barlow, J., da Costa, A.C.L., Marimon, B.S., Marimon-Junior, B.H., Meir, P., Metcalfe, D.B., Phillips, O.L., White, L.J.T., Malhi, Y., 2018. ENSO drives interannual variation of forest woody growth across the tropics. Philos. Trans. R. Soc. B: Biol. Sci. 373, 20170410. https://doi.org/10.1098/rstb.2017.0410.

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

Rui, H., Beaudoing, H., 2020. README Document for NASA GLDAS Version 2 Data Products. NASA Goddard Space Flight Center (GES DISC), Greenbelt, MD, USA.

Save, H., Bettadpur, S., Tapley, B.D., 2016. High-resolution CSR GRACE RL05 mascons. JGR Solid Earth 121, 7547–7569. https://doi.org/10.1002/2016JB013007.

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

Scanlon, B.R., Zhang, Z., Save, H., Sun, A.Y., Müller Schmied, H., van Beek, L.P.H., Wiese, D.N., Wada, Y., Long, D., Reedy, R.C., Longuevergne, L., Döll, P., Bierkens, M. F.P., 2018. Global models underestimate large decadal declining and rising water storage trends relative to GRACE satellite data. Proc. Natl. Acad. Sci. USA 115. https://doi.org/10.1073/pnas.1704665115.

Schamm, K., Ziese, M., Becker, A., Finger, P., Meyer-Christo, A., Stender, P., 2014. Global gridded precipitation over land: a description of the new GPCC first guess daily product. Earth Syst. Sci. Data 6, 49–60. https://doi.org/10.5194/essd-6-49-2014.

Schneider, A.M., Zender, C.S., Price, S.F., 2022. More realistic intermediate depth dry firn densification in the energy exascale earth system model (E3SM). J. Adv. Model Earth Syst. 14, e2021MS002542 https://doi.org/10.1029/2021MS002542.

Schulte-Uebbing, L., de Vries, W., 2018. Global-scale impacts of nitrogen deposition on tree carbon sequestration in tropical, temperate, and boreal forests: A meta-analysis. Glob. Change Biol. 24, e416–e431. https://doi.org/10.1111/gcb.13862.

Sheffield, J., Goteti, G., Wood, E.F., 2006. Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. J. Clim. 19, 3088–3111. https://doi.org/10.1175/JCLI3790.1.

Shi, X., Mao, J., Thornton, P.E., Hoffman, F.M., Post, W.M., 2011. The impact of climate, CO2, nitrogen deposition and land use change on simulated contemporary global river flow. Geophys. Res. Lett. 38, L08704. https://doi.org/10.1029/2011GL046773.

Shi, X., Mao, J., Thornton, P.E., Huang, M., 2013. Spatiotemporal patterns of evapotranspiration in response to multiple environmental factors simulated by the community land model. Environ. Res. Lett. 8, 024012 https://doi.org/10.1088/1748-9326/8/2/024012.

Strassberg, G., Scanlon, B.R., Chambers, D., 2009. Evaluation of groundwater storage monitoring with the GRACE satellite: case study of the High Plains aquifer, Central United States. Water Resour. Res. 45 https://doi.org/10.1029/2008WR006892.

Swenson, S.C., Lawrence, D.M., 2014. Assessing a dry surface layer-based soil resistance parameterization for the community land model using GRACE and FLUXNET-MTE data. J. Geophys. Res. 119, 10299–10312. https://doi.org/10.1002/2014JD022314.

Swenson, S.C., Lawrence, D.M., 2015. A GRACE-based assessment of interannual groundwater dynamics in the community land model. Water Resour. Res. 51, 8817–8833. https://doi.org/10.1002/2015WR01758.

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

Tang, J., Riley, W.J., Niu, J., 2015. Incorporating root hydraulic redistribution in CLM4.5: effects on predicted site and global evapotranspiration, soil moisture, and water storage. J. Adv. Model. Earth Syst. 7, 1828–1848. https://doi.org/10.1002/2015MS000484.

Tapley, B.D., Watkins, M.M., Flechtner, F., Reigber, C., Bettadpur, S., Rodell, M., Sasgen, I., Famiglietti, J.S., Landerer, F.W., Chambers, D.P., Reager, J.T., Gardner, A. S., Save, H., Ivins, E.R., Swenson, S.C., Boening, C., Dahle, C., Wiese, D.N., Dobslaw, H., Tamisiea, M.E., Velicogna, I., 2019. Contributions of GRACE to understanding climate change. Nat. Clim. Chang. 9, 358–369. https://doi.org/10.1038/s41558-019-0456-2.

Truong, C., Oudre, L., Vayatis, N., 2020. Selective review of offline change point detection methods. Signal Process. 167, 107299 https://doi.org/10.1016/j.sigpro.2019.107299.

van den Hurk, B., Kim, H., Krinner, G., Seneviratne, S.I., Derksen, C., Oki, T., Douville, H., Colin, J., Ducharne, A., Cheruy, F., Viovy, N., Puma, M.J., Wada, Y., Li, W., Jia, B., Alessandri, A., Lawrence, D.M., Weedon, G.P., Ellis, R., Hagemann, S., Mao, J., Flanner, M.G., Zampieri, M., Materia, S., Law, R.M., Sheffield, J., 2016. LS3MIP (v1.0) contribution to CMIP6: the land surface, snow and soilmoisture model intercomparison project – aims, setup and expected outcome. Geosci. Model Dev. 9, 2809–2832. https://doi.org/10.5194/gmd-9-2809-2016.

Viovy, N., 2009. CRUNCEP Version 7 - Atmospheric forcing data for the Community Land Model. https://doi.org/10.5065/PZ8F-F017.

Voss, K.A., Famiglietti, J.S., Lo, M., de Linage, C., Rodell, M., Swenson, S.C., 2013. Groundwater depletion in the Middle East from GRACE with implications for transboundary water management in the Tigris-Euphrates-Western Iran region. Water Resour. Res. 49, 904–914. https://doi.org/10.1002/wrcr.20078.

Wang, Y., Mao, J., Hoffman, F.M., Bonfils, C.J.W., Douville, H., Jin, M., Thornton, P.E., Ricciuto, D.M., Shi, X., Chen, H., Wullschleger, S.D., Piao, S., Dai, Y., 2022. Quantification of human contribution to soil moisture–based terrestrial aridity. Nat. Commun. 13, 6848. https://doi.org/10.1038/s41467-022-34071-5.

Wang, J., Song, C., Reager, J.T., Yao, F., Famiglietti, J.S., Sheng, Y., MacDonald, G.M., Brun, F., Schmied, H.M., Marston, R.A., Wada, Y., 2018. Recent global decline in endorheic basin water storages. Nat. Geosci. 11, 926–932. https://doi.org/10.1038/s41561-018-0265-7.

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

Wiese, D.N., Yuan, D.-N., Boening, C., Landerer, F.W., Watkins, M.M., 2023. JPL GRACE and GRACE-FO Mascon Ocean, Ice, and hydrology equivalent water height CRI filtered. https://doi.org/10.5067/TEMSC-3JC63.

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, 7490–7502. https://doi.org/10.1002/2016WR019344.

Winkler, K., Fuchs, R., Rounsevell, M., Herold, M., 2021. Global land use changes are four times greater than previously estimated. Nat. Commun. 12, 2501. https://doi.org/10.1038/s41467-021-22702-2.

Xiong, J., Yin, J., Guo, S., He, S., Chen, J., Abhishek, 2022b. Annual runoff coefficient variation in a changing environment: a global perspective. Environ. Res. Lett. 17, 064006 https://doi.org/10.1088/1748-9326/ac62ad.

Xiong, J., Abhishek, Guo, S., Kinouchi, T., 2022a. Leveraging machine learning methods to quantify 50 years of dwindling groundwater in India. Sci. Total Environ. 835, 155474 https://doi.org/10.1016/j.scitotenv.2022.155474.

Yang, Y., McVicar, T.R., Yang, D., Zhang, Y., Piao, S., Peng, S., Beck, H.E., 2021. Low and contrasting impacts of vegetation CO_2 fertilization on global terrestrial runoff over 1982–2010: accounting for aboveground and belowground vegetation–CO_2 effects. Hydrol. Earth Syst. Sci. 25, 3411–3427. https://doi.org/10.5194/hess-25-3411-2021.

Yang, X., Ricciuto, D.M., Thornton, P.E., Shi, X., Xu, M., Hoffman, F., Norby, R.J., 2019. The effects of phosphorus cycle dynamics on carbon sources and sinks in the Amazon region: a modeling study using ELM V1. J. Geophys. Res. Biogeosci. 124, 3686–3698. https://doi.org/10.1029/2019JG005082.

Yang, X., Thornton, P., Ricciuto, D., Wang, Y., Hoffman, F., 2023. Global evaluation of terrestrial biogeochemistry in the energy exascale earth system model (E3SM) and the role of the phosphorus cycle in the historical terrestrial carbon balance. Biogeosciences 20, 2813–2836. https://doi.org/10.5194/bg-20-2813-2023.

Yeh, P.-J.-F., Swenson, S.C., Famiglietti, J.S., Rodell, M., 2006. Remote sensing of groundwater storage changes in Illinois using the gravity recovery and climate experiment (GRACE). Water Resour. Res. 42 https://doi.org/10.1029/2006WR005374.

Zhang, Y., He, B., Guo, L., Liu, D., 2019a. Differences in response of terrestrial water storage components to precipitation over 168 global river basins. J. Hydrometeorol. 20, 1981–1999. https://doi.org/10.1175/JHM-D-18-0253.1.

Zhang, Y., He, B., Guo, L., Liu, J., Xie, X., 2019b. The relative contributions of precipitation, evapotranspiration, and runoff to terrestrial water storage changes across 168 river basins. J. Hydrol. 579, 124194 https://doi.org/10.1016/j.jhydrol.2019.124194.

Zhao, A., Stevenson, D.S., Bollasina, M.A., 2019. Climate forcing and response to greenhouse gases, aerosols, and ozone in CESM1. J. Geophys. Res. Atmos. 124, 13876–13894. https://doi.org/10.1029/2019JD030769.

Zhou, T., Leung, L.R., Leng, G., Voisin, N., Li, H., Craig, A.P., Tesfa, T., Mao, Y., 2020. Global irrigation characteristics and effects simulated by fully coupled land surface, river, and water management models in E3SM. J. Adv. Model. Earth Syst. 12 https://doi.org/10.1029/2020MS002069.

Zhu, B., Huang, M., Cheng, Y., Xie, X., Liu, Y., Bisht, G., Chen, X., 2021. Impact of vegetation physiology and phenology on watershed hydrology in a semiarid watershed in the Pacific northwest in a changing climate. Water Resour. Res. 57, e2020WR028394 https://doi.org/10.1029/2020WR028394.

Zou, Y., Kuang, X., Feng, Y., Jiao, J.J., Liu, J., Wang, C., Fan, L., Wang, Q., Chen, J., Ji, F., Yao, Y., Zheng, C., 2022. Solid water melt dominates the increase of Total groundwater storage in the Tibetan Plateau. Geophys. Res. Lett. 49, e2022GL100092 https://doi.org/10.1029/2022GL100092.