Global evaluation of the dry gets drier and wet gets wetter paradigm from terrestrial water storage changes perspective

Jinghua Xiong[1], Shenglian Guo[1], Abhishek[2], Jie Chen[1], and Jiabo Yin[1]

[1] State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan, 430072, China
[2] School of Environment and Society, Tokyo Institute of Technology, Yokohama 226-8503, Japan

Correspondence to: S. Guo (slguo@whu.edu.cn)

Abstract. The "dry gets drier and wet gets wetter" (DDWW) paradigm has been widely used to summarise the expected trends of the global hydrologic cycle under climate change. However, the paradigm is largely conditioned by choice of different metrics and datasets used and is still unexplored from the perspective of terrestrial water storage anomaly (TWSA). Considering the essential role of TWSA in wetting and drying of the land system, here we built upon a large ensemble of TWSA datasets, including satellite-based products, global hydrological models, land surface models, and global climate models to evaluate the DDWW hypothesis during the historical (1985-2014) and future (2071-2100) periods under various scenarios with a 0.05 significance level. We find that 11.01%-40.84% (range by various datasets) of global land confirms the DDWW paradigm, while 10.21%-35.43% of the area shows the opposite pattern during the historical period. In the future, the DDWW paradigm is still challenged with the percentage supporting the pattern lower than 18%, and both the DDWW-validated and DDWW-opposed proportion increase along with the intensification of emission scenarios. We show that the different choices of data sources can reasonably influence the test results up to a four-fold difference, while the varying significance levels (0.01-0.1) have subtle influences on the evaluation results of the DDWW paradigm. Our findings will provide insights and implications for global wetting and drying trends from the perspective of TWSA under climate change.

1 Introduction

The global hydrological cycle has experienced considerable changes due to climate change and anthropogenic interventions, exerting a tremendous impact on agriculture, ecological environment, and freshwater availability globally (Shugar et al., 2020; Perera et al., 2020; Gampe et al., 2021). Assessing the variations of constituent components of the water cycle, namely, precipitation (P), evapotranspiration (E), runoff (R), and storage change, are therefore crucial in understanding the systematic hydrological response and dealing with water-related issues in the context of global change (Moreno-Jimenez et al., 2019; Zhao et al., 2021). Under these circumstances, the 'dry gets drier and wet gets wetter' (DDWW) paradigm, firstly introduced by Held and Soden (2006), has become one of the most widely used hypotheses to summarise the long-term trends in the global hydrological cycle (Roderick et al., 2014; Yang et al., 2019). Initially, it was developed based on the deficit between precipitation and evapotranspiration ($P - E$), which is expected to increase due to the enhancement of atmospheric water vapour in humid regions (i.e., convergence zones) under a warming climate, and decrease over arid regions (i.e., divergence zones) (Durack et al., 2012). The DDWW paradigm has been used to represent the historical and future trends in various constituent components of the hydrologic cycle on regional (Chou et al., 2009; Allan et al., 2010; Hu et al., 2019; Zeng et al., 2019) and global scales (Held and Soden, 2006; Donat et al., 2016). However, the rationale and validity of the DDWW mechanism are recently questioned at different levels through the growing number of datasets, model simulations, and indicators (Polson and Hegerl, 2017; Yang et al., 2019; Y. Li et al., 2021a). Byrne and Gorman (2015) used simulations from 10 climate models to reveal an ocean-land contrast pattern in the response of $P - E$ to global warming in historical (1976-2005) and future (2071-2099) periods, highlighting the DDWW as a more suitable mechanism over ocean than over land. Given the fact that historical evaluation of the DDWW paradigm was mainly based on oceanic observations, Greve et al. (2014) adopted 2142 possible combinations of $P - E$ to assess the trends in wetting and drying over global land and discovered merely 10.8% of the area following the DDWW pattern during the period 1948-2005. Roderick et al. (2014) revisited the DDWW paradigm, cautioned about its interpretation owing to the different behavior of land and ocean with respect to the water cycle, and showed that the paradigm does not hold true in terms of projected changes in the mean annual water balance over land. Alternatively, Yang et al. (2019) integrated an ensemble of six hydro-climatic indicators for the global assessment of the DDWW paradigm between 1982 and 2012, suggesting the catchphrase only occurred over 20% of the global land. In a nutshell, there are great uncertainties still remaining in the assessments and subsequent interpretation of global trends in dryness and wetness under climate change (Dai, 2011; Trenberth et al., 2014).

The uncertainties within previous studies are mainly sourced from different choices of metrics adopted and datasets used for evaluating the changes in dryness and wetness (Vicente-Serrano et al., 2010; Feng and Zhang, 2015; Huang et al., 2016). Specifically, the widely used metric $P - E$ over the ocean has been proven overwhelmingly positive over land based on both observations and simulations, revealing an ocean-dominated behavior (Greve et al., 2014; Byrne and O'Gorman, 2015; Greve and Seneviratne, 2015). Moreover, some meteorological indices derived from precipitation and evapotranspiration, such as the standardized precipitation evapotranspiration index (SPEI), aridity index (AI), and standardised precipitation/evapotranspiration index (SPI/SETI), do not capture the integrated response of the land system due to the trade-off between the simplicity of meteorological factors and computational requirements of process-based variables (Huntington, 2006; Dai, 2011; Slette et al., 2020; Barnard et al., 2021). A few indexes like the standardised soil moisture index (SSI), standardised groundwater index (SGI), and standardised runoff index (SRI), however, focus on a single aspect of the water cycle and do not describe the integrated status of the terrestrial water storage (TWS) (AghaKouchak, 2014; Wu et al., 2018; Guo et al., 2021). In the coupled human-natural systems, where the synergistic impacts of natural and anthropogenic drivers are exceedingly difficult to disentangle, an integrated representation of the land systems is of paramount importance for policymakers (Rodell et al., 2018).

TWS, consisting of water storage in surface water, soil moisture, groundwater, snow and ice, and canopies, can physically provide integrated information about the overall status of the land, whose changes are closely linked to the terrestrial wetting and drying tendency (Tapley et al., 2019; Pokhrel et al., 2021). Apart from the societal and economic importance, TWS plays a vital role in Earth system processes, including climate, weather, and biogeochemical cycles (Seyoum and Milewski, 2017). Change in storage, i.e., the difference between the consecutive TWS values, is a key variable of the hydrological cycle. Therefore, understanding the spatiotemporal dynamics of past and future TWS is not only essential for human life but also crucial for assessing the water cycle, planning, policymaking, and other management strategies for water resources in a changing climate and for a continuously increasing population (Abhishek et al., 2021). There are several studies dealing with TWS or derived indicators to assess freshwater availability (Rodell et al., 2018), water storage dynamics (Scanlon et al., 2018), and droughts and floods monitoring (Long et al., 2014), among others. Divergent patterns of TWS changes have been reported over arid and humid regions under the combined effects of climate change (e.g., global warming), climatic variability (e.g., ENSO), and human activity (e.g., groundwater pumping) (Chang et al., 2020; An et al., 2021; Hu et al., 2021). However, there is no study to examine the global variability and validity of DDWW paradigm in the past and future in terms of TWS changes. Furthermore, divergent data sets produce different trends in TWS due to distinctive internal variability and external forcing (from satellites and meteorological stations), especially from precipitation and evapotranspiration (Chen et al., 2020). For example, Scanlon et al. (2018) conducted comprehensive comparisons between decadal trends in TWS from seven global models and three Gravity Recovery and Climate Experiment (GRACE) satellite solutions over major basins globally and showed a large underestimation of the increasing and decreasing trends of models primarily due to human water use and forcing climate variations.

Therefore, to bridge the aforesaid research gap, we conduct a systematic evaluation of the DDWW paradigm from the perspective of terrestrial water storage anomalies (TWSA) using an ensemble of five different TWS datasets, including one GRACE reconstruction, two global hydrological models (GHMs), and two land surface models (LSMs) between 1985 and 2014. Subsequently, an alternative ensemble of eight global climate models (GCMs) from the Coupled Model Intercomparison Project 6 (CMIP6) is used to further test the paradigm under various scenarios during the future period (2071-2100). Utilizing the data from these models and observation-based products, we further establish the metric "$P - E - R$" in terms of the water balance equation for intercomparisons with the test results from the aspect of TWSA and for highlighting the governing mechanisms of the estimated disparities.

2 Data and Methods

2.1 Data pre-processing

We perform the assessment of the DDWW paradigm over global land at both gridded $1° \times 1°$ cell and regional scales excluding Greenland and Antarctica. One of the global hotspots with significant changes in hydroclimatological conditions (e.g., precipitation and air temperature) (Liu et al., 2006; Zhang et al., 2017) and TWSA (Meng et al., 2019) over the last two decades, i.e., the Qinghai-Tibetan Plateau (QTP), is selected as a typical region for regional analysis. The QTP and its surroundings which are called the world's "the Third Pole" play a crucial role in the freshwater availability of more than 1.4

billion people (Immerzeel et al., 2010). The QTP is mainly covered by polar tundra and cold and arid steppe climate region (Figure S2), causing the sparse distribution of in-situ networks there (Wan et al., 2014). Thus, using alternative methods such as remote sensing (e.g., GRACE) and global model outputs (e.g., GHMs, LSMs, and GCMs) to study the hydrological variations in the QTP is of much importance.

We use an ensemble of five TWSA data sets to evaluate the DDWW paradigm during the historical period 1985-2014, which includes one GRACE reconstruction, two global hydrological models (GHMs), and two global land surface models (LSMs) (see Table 1 and next sections). Please note that some studies may use the term GHMs to represent both global hydrological and water resource models (GHWRMs) and LSMs together (Scanlon et al., 2018), while we use it only for the former one for distinction and simplicity. Since no dataset presents the absolutely 'true' value, we demonstrate the individual results of each member to avoid the uncertainty derived from different TWSA definitions in various models/products (Table S1). The missing months (12% of the months, i.e., June 2002, July 2002, June 2003, January 2011, June 2011, May 2012, October 2012, March 2013, August 2013, September 2013, February 2014, July 2014, December 2014) of GRACE measurements have been filled using a linear interpolation method. In addition, an ensemble of eight TWSA simulations from CMIP6 GCMs is used to examine the DDWW paradigm in the future period (2071-2100). The members of the CMIP6 ensemble and all of the historical datasets have been resampled to $1° \times 1°$ scale using a bilinear interpolation approach for consistency and better comparison in the spatial domain. The ensemble mean of CMIP6 models has been estimated using simple averaging because they have the same simulation objects (Table S1). All the historical datasets and CMIP6 members, as well as their ensemble, are represented as the long-term anomaly relative to the baseline between 1985 and 2014. We also calculate the metric P-E-R based on the water balance equation for cross-comparison with the test results from the TWSA perspective. This metric is estimated using P, ET, and R from the same models as those of TWSA (e.g., GHMs, LSMs, and GCMs) for consistency. Moreover, an observation-based combination is also derived as benchmarking subset based on precipitation (P) from the Climatic Research Unit gridded Time Series (CRU TS-v4.06, Harris et al., 2020), evapotranspiration (E) from the Global Land Evaporation Amsterdam Model (GLEAM-v3.6, Martens et al., 2017), and runoff (R) from the G-RUN ensemble (Ghigg et al., 2021) (Table 1).

Table 1. Datasets used in this study.

Type	Data	URL	Selected period	Raw temporal resolution	Raw spatial resolution (longitude×latitude)
GRACE reconstructions	Li et al., 2021b	https://doi.org/10.1029/2021GL093492	1985-2014	Monthly	0.5°×0.5°
GRACE observations	GRACE CSR RL06 mascons-v02	http://www2.csr.utexas.edu/grace/	2002-2014	Monthly	0.25°×0.25°
GHMs	WGHM-v2.2d	https://gmd.copernicus.org/articles/14/1037/2021/	1985-2014	Monthly	0.25°×0.25°
	GLDAS2.0-VIC	https://ldas.gsfc.nasa.gov/gldas	1985-2014	Monthly	1°×1°
LSMs	GLDAS2.0-Noah	https://ldas.gsfc.nasa.gov/gldas	1985-2014	Monthly	1°×1°
	GLDAS2.0-CLSM	https://ldas.gsfc.nasa.gov/gldas	1985-2014	Monthly	1°×1°
GCMs	ACCESS-CM2	https://esgf-node.llnl.gov/projects/cmip6/	1985-2100	Monthly	1.25°×1.875°
	ACCESS-ESM1-5	https://esgf-node.llnl.gov/projects/cmip6/	1985-2100	Monthly	1.24°×1.875°
	CanESM-5	https://esgf-node.llnl.gov/projects/cmip6/	1985-2100	Monthly	2.8125°×2.8125°
	GFDL-ESM4	https://esgf-node.llnl.gov/projects/cmip6/	1985-2100	Monthly	1°×1.25°
	IPSL-CM6A-LR	https://esgf-node.llnl.gov/projects/cmip6/	1985-2100	Monthly	1.2587°×2.5°
	MIROC6	https://esgf-node.llnl.gov/projects/cmip6/	1985-2100	Monthly	1.4063°×1.4063°
	MPI-ESM1-2-HR	https://esgf-node.llnl.gov/projects/cmip6/	1985-2100	Monthly	0.9375°×0.9375°
	MPI-ESM1-2-LR	https://esgf-node.llnl.gov/projects/cmip6/	1985-2100	Monthly	1.875°×1.875°
In-situ-derived precipitation and potential evapotranspiration	CRU TS-v4.06	https://crudata.uea.ac.uk/cru/data/hrg/cru_ts_4.06/	1985-2014	Monthly	0.5°×0.5°
In-situ-derived runoff	G-RUN Ensemble	https://doi.org/10.1029/2020WR028787	1985-2014	Monthly	0.5°×0.5°
Satellite-based evapotranspiration	GLEAM-v3.6a	https://www.gleam.eu/	1985-2014	Monthly	0.25°×0.25°

2.1.1 GRACE and GRACE Reconstructions

The GRACE (and GRACE Follow-On) missions have provided unprecedented estimates of monthly TWSA worldwide from April 2002 up to the present, however, with the 33 months missing because of the instrumental issues and mission interruption (Tapley et al., 2004). We use the GRACE mascon solution from the Center for Space Research at the University of Texas at Austin (UTCSR) to serve as the benchmarking product from the period 2002-2014 (Watkins et al., 2015). Compared to conventional GRACE products (e.g., spherical harmonic solutions), mascon solutions do not need spatial (e.g., smoothing) or spectral (e.g., de-striping) filtering or other empirical scaling and therefore have higher signal-to-noise ratio, higher spatial resolutions, and eventually reduced errors (Save et al., 2016; Watkins et al., 2015). However, the GRACE observational products were not adequate to assess the long-term trends of TWSA due to relatively short temporal coverage (~20 years). Therefore, we obtain the GRACE reconstruction provided by Li et al. (2021b) for evaluation of the DDWW paradigm, which is generated using state-of-the-art machine learning and statistical methods and is also trained by the consistent GRACE mascon product from the UTCSR institution. The GRACE reconstruction applies four meteorological variables (i.e., precipitation, 2 m air temperature, sea surface temperature, and multiple climate indices) and three hydrological variables (i.e., soil moisture, runoff, and evaporation) to simulate the temporally decomposed GRACE signals (i.e., the seasonal, inter-annual, and residual components) (Li et al., 2021a). We would like to mention that the trend components in GRACE reconstructions are directly added by the linear GRACE trends, which are mainly caused by glacier melt and anthropogenic factors (e.g., dam constructions and water abstractions). These factors are difficult to predict using the climatic and hydrologic inputs and may change over time, causing the possible bias in the long-term trend estimates from GRACE reconstructions. The accuracy and applicability of the GRACE reconstruction have been fully evaluated over global land in several previous studies (Xu et al., 2021; Yi et al., 2021).

2.1.2 Global Hydrological Models

We use two global hydrological models, including the Variable Infiltration Capacity macroscale model (VIC-v4.1.2) and the WaterGAP hydrological model (WGHM-v2.2d), to estimate TWS and P-E-R for independent evaluation of the DDWW paradigm. The physically-based, semi-distributed, and grid-based VIC model is managed by the NASA Global Land Data Assimilation System Version 2.0 (GLDAS-v2.0) (Liang et al., 1994; Syed et al., 2008). Forced by the Global Data Assimilation System atmospheric analysis fields (Derber et al., 1991) and the Air Force Weather Agency's AGRicultural METeorological modeling system radiation fields, the VIC model can effectively capture the terrestrial water cycle by simulating the water stored in the canopies, snow, and soil moisture within three soil layers up to a depth of 200 cm. The VIC model has been widely used to analyze terrestrial water storage changes at regional and global scales (Hao and Singh, 2015; Hao et al., 2018). The WGHM is a grid-based global hydrological model quantifying the human water use and continental water fluxes for all land areas excluding Antarctica (Müller Schmied et al., 2021). Unlike most global hydrological models, the WGHM forced by the ERA40 and ERA-Interim reanalysis can simulate groundwater storage by coupling with global water use models like the Groundwater-Surface Water Use, suggesting a comparably better representation of TWS (Döll et al., 2014). Several frequently-used model outputs such as TWS, discharge, and water use have been evaluated against global observations (Wan et al., 2021). E and R from the VIC and WGHM models are also extracted for the calculation of the variable "P-ET-R" by combining the P from their meteorological inputs of GLDAS2.0.

2.1.3 Land Surface Models

We use two land surface models consisting of the Noah (v3.6) and Catchment (CLSM-vF2.5) models to calculate TWS and P-E-R globally for parallel assessment of the DDWW paradigm. Similar to the VIC model, both Noah and CLSM models are managed by GLDAS (v-2.0) from the NASA GSFC institute. GLDAS is a composite of global hydrological and land surface models that simulate the optimal fields of the land by using state-of-the-art data assimilation and land surface simulation techniques (Rodell et al., 2004). GLDAS has been widely used to compare with GRACE TWSA in data-sparse regions such as Africa and Qinghai-Tibetan Plateau (Ogou et al., 2021; Xing et al., 2021). The Noah-modelled TWS is considered as the sum of canopy water storage, snow water equivalent, and soil moisture of four layers with a total depth of 200 cm. Different from that, the CLSM simulates shallow groundwater and the vertical levels of soil moisture are not explicitly divided within the depth of 100 cm. Similarly, we used the E and R modelled by the CLSM and Noah models to calculate the index P-E-R. We note that the three GLDAS models (i.e., VIC, CLSM, and Noah) share the same P estimations due to the consistent meteorological inputs, which might reduce the bias in the estimates of the metric P-E-R.

2.1.4 Global Climate Models

We use a suite of eight global climate models belonging to the ensemble "r1i1p1f1" of CMIP6 to evaluate the DDWW paradigm under climate change. The CMIP6 serves as a category of experiments of GCMs coupled to the dynamic ocean, simple land surface, and thermodynamic sea ice (Eyring et al., 2016). We choose these eight models out of the 34 CMIP6 models because they are the only models for which TWSA outputs are available in both the historical and future periods under multiple emission scenarios (see Table 1). The CMIP6 TWSA represents the sum of total soil moisture and snow equivalent water, which has been comprehensively validated with the GRACE data, though with embedded uncertainties, over global major river basins (Freedman et al., 2014; Wu et al., 2021). The CMIP6 comparisons have become a diagnostic tool to better understand climate change in past, present, and future periods (Eyring et al. 2016), which includes a total of five Shared Socio-economic Pathways (SSPs) representing global economic and demographic changes under different greenhouse gas emissions. We select three out of five SSP scenarios, including SSP126, SSP245, and SSP585, representing the green roads, middle of the road, and the highway road, respectively (Iqbal et al., 2021). Since the GCMs have different TWSA definitions from the "actual" TWSA observed by GRACE (Table S1), we employ a trend-preserving method to perform bias correction combined with historical GRACE data. The trend-preserving method initially developed by Hempel et al. (2013) modifies the monthly means of the simulated data to match the observed data using a constant offset between simulations and observations and has been widely used in the Intersectoral Model Intercomparison Project (ISIMIP2b). The detailed procedure of the bias correction for CMIP6 TWSA has been described in detail in a recent study (Xiong et al., 2022a). To show the difference before and after the bias correction, we select two typical regions (i.e., Amazon and Mekong River basins) with abundant surface and groundwater resources (Pham et al., 2019). Of the two selected basins, the Mekong River basin experiences severe human interventions such as groundwater pumping, dam constructions, and urbanization, while the Amazon River basin is considered as one of the largest natural river basins with low impacts of human activities (Xiong et al., 2022b). It is discovered that the GCM simulations without bias correction show obvious underestimations over two regions with large uncertainty, which have, however, significantly reduced after bias correction along with a lower spread range (Figure S16). The amplitudes of the GCM series are adjusted to nearly the same as GRACE data, with the long-term trends unaffected. It is noteworthy that the trend-preserving method would not affect the long-term trends of the GCM TWSA, and, therefore, not influence our current DDWW evaluation results. In addition to the TWSA, we also derive the predictions of P, E, and R for the construction of the P-E-R to compare with TWSA similar to those from GHMs and LSMs.

2.2 Detection of Wetting and Drying

TWSA, consisting of the water volume stored in the land surface and subsurface, is applied to define the "wetting" and "drying" conditions of the landmass in this study. The non-dimensional TWS drought severity index (TWS-DSI) is established at both $1° \times 1°$ grid cell and regional/global scales, which is normalised by the regional hydroclimatological variability because a given magnitude of TWS deficit could indicate different dryness/wetness conditions in different climate regions. TWS-DSI has clear classification categories based on U.S. Drought Monitor (USDM) and is suitable for comparing dryness/wetness status for different locations and periods (Table S2). It has been widely used in hydrology and climate fields due to its simple structure and effective ability to capture drying and wetting conditions (Pokhrel et al., 2021). The monthly TWS-DSI is calculated for all ensemble members and their mean from CMIP6 as follows (Zhao et al., 2017):

$$TWS\text{-}DSI_{i,j} = \frac{TWS_{i,j} - \mu_j}{\sigma_j} \tag{1}$$

where $TWS_{i,j}$ is the TWS value in year i and month j; μ_j and σ_j denote the mean and standard deviation of the annual TWS in month j, respectively. We convert the monthly TWS-DSI into annual means to calculate the long-term trends using (Figure S1) the linear regression method. We examine the first-order autocorrelation of each TWSA dataset using the Durbin-Watson test (Durbin and Watson, 1950, 1951). We find a total of 20% (GRACE reconstruction), 43% (WGHM), 41% (VIC), 23% (CLSM), 29% (Noah), and 20% (GCM) of the grid cells not presenting autocorrelation during 1985-2014, respectively. For the future period, the percentage is 25%, 26%, and 22% under the SSP126, SSP245, and SSP585 scenarios, respectively. In this case, the significance of the long-term trends is evaluated using the modified Mann-Kendall trend test at a 5% level to avoid autocorrelation (Hamed and Rao, 1998). Similarly, we also estimate the long-term trends of the index P-E-R for comparison with TWS-DSI using the same methods. The area having a significant trend of increasing/decreasing TWS-DSI or P-E-R is considered to be undergoing wetting/drying; otherwise, it is defined as a region with a non-significant trend.

To evaluate the DDWW paradigm over global land, the effective Aridity index (AI) is used to classify a grid cell as an arid, humid, and transitional region following Yang et al. (2019) because TWS-DSI/TWSA approximates zero for the long-term mean. The AI is calculated as the ratio of annual precipitation to potential evapotranspiration provided by the CRU TS-v4.06 during the same period as TWS-DSI (i.e., 1985-2014). The global distribution of multi-year average AI and the classifications during the period 1985-2014 is presented in Figure S3, which is also highly consistent with the widely used Köppen-Geiger climate classification maps (Beck et al. 2018) (Figure S2). It can be seen that most of the arid regions (AI<0.5) are located in southwestern America, north and south Africa, central Asia, Arabian regions, and Australia, accounting for 39.3% of the land. The percentage of humid areas (AI>0.65) that are mainly located in east America, the Amazon region, central Africa, south China, west Europe, and Russia reaches 52.8% of the land. An approximate 7.9% of the land area is defined as the transitional region, referring to an intermediate between arid and humid climates. The transitional region generally lies in the shared boundaries of the humid and arid regions (e.g., western America, northern Canada, central Asia, western Africa, East Russia, and Australia). The DDWW paradigm is evaluated at a 5% significance level in this study, combined with the standard AI-derived climate classifications. We calculate the global mean trends of TWS-DSI using a spatially weighted method to account for the changing area of grid cells with latitudes.

3 Results and Discussion

3.1 Global Trends of Dryness and Wetness

We firstly assess the reliability of the GRACE reconstruction, GHMs, and LSMs by comparing them with the GRACE observations. Figure S4 presents the global distribution of the normalized root mean square error (NRMSE) between the mean GRACE TWSA and the ensemble means of CMIP6 datasets after bias correction during the period April 2002-December 2014, with the NRMSE calculated as the ratio of RMSE to the differences between the maximum and minimum GRACE TWSA. The GRACE reconstruction shows the best performance over five TWSA datasets, with the NRMSE generally lower than 0.2 (for 97.4% of the global land area), of which 48.0% of the land area shows NRMSE below 0.1. In particular, NRMSE ranging from 0.1 to 0.3 occurs in the west and central Asia, North China, South Australia, eastern Russia, north and south Africa, and central northern and southern America (Fig. S4). Two GHMs (i.e., WGHM and VIC) and two LSMs (CLSM and Noah) present a similar spatial pattern of NRMSE to the GRACE reconstruction but with a relatively higher bias, among which the VIC model outperforms the other three models. The CLSM model shows comparatively poor performance, which is also confirmed by the probability density distributions of NRMSE compared with GRACE (Figure S4). The outperformance of the GRACE reconstruction over other data may be because they are directly calibrated with the GRACE measurements during 2002-2017, while their performances need more validation beyond the GRACE era (i.e., prior to April 2002 and during July 2017-May 2018). A temporal comparison of global average TWSA derived from different datasets and CMIP6 and GRACE during 2002-2014 is shown in Figure S5. The GRACE TWSA ranges from roughly −20 to 20 mm and shows obvious seasonal characteristics with relatively higher uncertainty in the dry season than that in the wet season. A similar temporal pattern is captured by various models, with the change spread covering the variations of GRACE data. The NRMSE between multiple datasets and GRACE data ranges from 0.08 (GRACE reconstruction) and 0.16 (Noah), coinciding with the strong correlation within different datasets (Figures S4 and S6). Moreover, the fluctuation range of the CMIP6 is generally larger than different historical models/products, highlighting the considerable uncertainty sourced from different forcing variables and model parameterizations. Then, we examine the difference between GCMs-simulated TWSA before and after the trend-preserving bias correction using GRACE. It is discovered their correlation coefficients improve by comparing with GRACE, while slightly decreasing within the eight GCMs, which can be attributed to the introduced uncertainty when performing the bias correction (Figure S7). In addition, the spatial distributions clearly show that the ensemble mean of eight GCMs outperforms each member globally, particularly in Australia, southern Africa, and North America (Figures S8 and S9). The outperformance becomes more obvious after bias correction. An overall decrease in NRMSE is also observed according to the probability density functions after performing bias correction, which is also detected from the Taylor diagram results (see Figure S10). We also provide the evaluation of the bias-corrected TWSA changes (i.e., TWSC) using the water balance estimates (i.e., P-E-R=TWSC) during 1985-2014 (Figures S11 and S12). The observation-based water balance estimates correlate well with GRACE TWSA and GCM-modelled P-E-R with a correlation coefficient of 0.62 and 0.93, respectively. The GCM-simulated changes in TWSA also present a strong correlation with the observational products before and after bias correction. The spatial distribution of correlation coefficients between TWSC from observations and GCMs with and without bias correction shows the performances in regions with good accuracy, like Alaska, western parts of the Tibetan Plateau, and northern Russia, decrease after bias correction, which might be caused by the simplified treatment of permafrost in GCMs due to the prevailing uncertainties in, e.g., changes in thermophysical properties of the soil during freezing and thawing cycles (Burke et al., 2020). On the contrary, the areas with relatively poorer accuracy before bias correction, such as North Africa and northern South America, slightly improve after bias correction. Notwithstanding the observed differences in some regions, our trend-preserving method used for bias correction would not influence the long-term trend estimations of both TWSA and TWS-DSI and therefore does not impact our evaluation of the DDWW paradigm (Hempel et al., 2013). Although the bias correction has been performed on the CMIP6 TWSA, some biases inherent to the uncertainty in parameters, hydrometeorological forcing, and internal variability of GCMs still exist, which may influence the assessment of the DDWW paradigm in the future period (2071-2100) climate change.

We access the long-term trends TWS-DSI during the historical period 1985-2014 (based on a GRACE reconstruction, two GHMs (WGHM and VIC), two LSMs (CLSM and Noah), and the ensemble mean of eight GCMs) and the future period 2071-2100 (based on the ensemble mean of eight GCMs) under SPSP126, SSP245, and SSP585 scenarios to provide insights into the terrestrial water storage changes for the DDWW paradigm (Figures 1 and S14). The GRACE reconstruction, having the best accuracy among all other model-based TWSA, is selected for detailed analysis, which also shows the highest proportion of areas with significant trends. During the historical period, a clear spatial homogeneity (clustered patterns) of TWS-DSI trends is observed globally and the average TWS-DSI has a significant decreasing slope of -0.11/yr ($p<0.05$) (Figure

1), similar to the results from SPI, SPEI, and AI (Wang et al., 2018; Yang et al., 2019) together with the results from other models (WGHM: −0.07/yr, VIC: −0.05/yr, CLSM: −0.06/yr, Noah: −0.04/yr, the ensemble mean of GCMs: −0.05/yr). Spatially, severe drying ($p<0.05$) exists in the Gulf of Alaska coast, the Canadian archipelago, Chile, and the QTP with significant slopes of TWS-DSI ranging from −0.09/yr to −0.12/yr (Figure 1), which is caused by the rapid melt of ice-sheet, glacier ablation, and increase in the active permafrost layer under a warming climate (Luthcke et al., 2013; Velicogna et al., 2014). Triggered by severe historical droughts and extensive water use from groundwater and surface water over decades, the drying trends in North Canada, southern California, and Texas can be clearly discovered, with decreasing trend of TWS-DSI ranging from −0.06/yr to −0.12/yr ($p<0.05$) (Bouchard et al., 2013; Haacker et al., 2016), so as in the eastern Brazil (Getirana, 2016). Moreover, overwhelming groundwater depletion due to unsustainable human water use such as irrigation is responsible for the increasing dryness at significant slopes ranging from −0.09/yr to −0.12/yr in southeast and north regions of Africa, eastern and centre of Europe, central Asia, North China, and northern India (Rodell et al., 2009; Feng et al., 2013; Ramillien et al., 2014; Peña-Angulo et al., 2020). The decreasing TWS-DSI is also reported over European Russia because of the decline in the storage of surface and ground waters (Vadim et al., 2018). Additionally, the significant decreases in TWS-DSI ranging from −0.09/yr to −0.12/yr ($p<0.05$) over the Caspian and Aral seas are seen to arise from the reductions of inflow discharge and precipitation as well as evapotranspiration increase (Zmijewski and Becker, 2014). Naturally, a moderate drying trend in southwestern Africa and central Mediterranean Europe caused by precipitation decrease is detected by the reduction of TWS-DSI (−0.06/yr to −0.12/yr) (Peña-Angulo et al., 2020). On the contrary, increasing precipitation dominates the wetting trend in mid-latitude regions, including southern Russia and Canadian, west Africa, southeastern and southwestern Europe, southeast Asia and northwestern China, with significant slopes roughly ranging from 0.06/yr to 0.12/yr (Figure 1) (Siebert et al., 2010; Ndehedehe et al., 2017; Peña-Angulo et al., 2020). On the contrary, some regions, such as the Amazon River basin, south Africa and eastern Australia, presenting wetting trends, are considered to experience a climatic shift from dry to the wet period (Chen et al., 2010; Gaughan and Waylen, 2012). When looking at the test results of the GHMs and LSMs, we notice the regional differences with generally consistent spatial patterns with the GRACE reconstruction. For example, the WGHM model shows depletion trends in TWS-DSI for the southwest of the South American continent. The three GLDAS models (i.e., VIC, CLSM, and Noah) do not capture the increasing trends in South China (i.e., Yangtze and Pearl River basins), of which the VIC model surprisingly shows the increasing trends over the Arab region. We additionally compare the trend estimations of the GCMs' ensemble mean during the period 1985-2014 (Figures 2 and S14). Despite the overall similarity to the above-mentioned datasets, the existing regional differences in western South Africa (drying), South China Sea Islands (drying), and West Asia (wetting) compared with multiple models provide additional insights, indicating the great potential of the CMIP6 ensemble in TWSA projections.

Further, we perform an independent assessment based on the metric P-E-R for comparison with the TWS-DSI results to reveal the inherent mechanisms of the changes (Figures 2 and S15). The observational product of the variable P-E-R presents a similar pattern to the test results using TWS-DSI, however, with non-significant trends over most regions. This can be explained by the fact that the magnitude of the changes in the water storage, i.e., TWSC, in a region are minimal compared to the actual TWSA trends (Lv et al., 2021). In particular, the decreasing P-E-R (=TWSC) in southwest South America, North and South Africa, West Australia, North China, European Russia, and central Asia is observed with trends <–2 mm/yr, while increasing trends in northern Canada, central America, central Africa, East Australia, South India, and South and East Russia are found with rates >2 mm/yr. The local differences over the Arab region, south China, and the Caspian Sea might be caused by the propagated uncertainty in multiple observational datasets, especially for the arid regions (e.g., North Africa and West America), where accurately estimating E is very challenging (Goyal, 2004). For South China, consisting of the Yangtze and Pearl River basins, the difference might arise from the extensive reservoir filling, such as the Three Gorges Dam (Zhong et al., 2009), highlighting the significant role of human activities in the regional variations of TWS. Similarities are also seen over the land around the Caspian Sea, which is largely affected by the direct diversions and extractions of water from the rivers that sustain it (e.g., Volga River) instead of the conventionally dominant precipitation/evapotranspiration patterns over the sea surface (Rodell et al., 2018). It is worth mentioning again that the P-E-R equals the changes in TWSA (TWSC) rather than TWSA in terms of the water balance equation. Therefore, unlike TWSA, there are no significant trends in P-E-R over most regions of the world, which is also mentioned by several previous studies (Lv et al., 2019; Lv et al., 2021). Inter-comparisons with the GHMs and LSMs further confirm our observation-based evaluations, with relatively fewer magnitudes and significance derived from the substantial uncertainties in simulated E and R. In this case, we find an abnormally wetting trend in southwestern America, which might be caused by the severe groundwater pumping and water diversion implicitly considered in the metric P-E-R (Perrone and Jasechko, 2017). Satisfactory consistencies of GHMs and LSMs are also discovered by comparing each subset of P-E-R to the corresponding test results using TWS-DSI. The historical simulations of P-E-R from the ensemble mean of eight GCMs also compare reasonably well with different subsets, though showing the spatial differences over certain regions (e.g., central Europe and south Africa).

Furthermore, we investigate the long-term trends in P, E, and R, respectively, to explain the mechanisms for the changes in land mass wetness/dryness (Figures S16-S18). Different products and models show consistent spatial patterns for P, in which significant ($p<0.05$) increasing trends are detected in eastern North America (5-10 mm/yr), central Amazon (10-20 mm/yr), North Central and Southern Africa (0-5 mm/yr), northern Mediterranean basin (5-10 mm/yr), northwestern China (0-5 mm/yr), East Russia (0-5 mm/yr), North Europe (0-5 mm/yr), and North Australia (0-10 mm/yr). The highest trends are found in the South China Sea Islands (>20 mm/yr). However, decreasing trends over some areas, including North Canada (–5-0 mm/yr), Southwest parts of the United States (–10-–5 mm/yr), central South America (–15-0 mm/yr), Arab regions (–5-0 mm/yr), and Northeast India (<–20 mm/yr) also exist. In terms of E, multiple datasets illustrate generally similar trend distributions with the regional variability in specific areas (e.g., central Africa and Amazon River basin). Significant increases in E are observed over south and north Asia, North Australia, central and northern Europe, eastern North America, and South and central North Africa are seen by all the datasets, with the trends mainly ranging from 0 to 6 mm/yr. This increase might be caused by the warming climate and precipitation changes (Wang et al., 2022). However, we also notice the decreasing trends in western United States (–4-0 mm/yr), central South America (–8-–4 mm/yr), and Arab regions (–2-0 mm/yr), probably related to the heavy land-cover changes (Ruscica et al., 2022). Moreover, we discover overall similarities among trend estimates in R from different datasets, which are mainly dominated by the precipitation changes regionally with relatively lower amplitudes (roughly between –12—12 mm/yr) except for arid central Asia and East Europe. In addition. we want to mention that despite the general agreement with different observational products and models, the GCMs-based historical trends estimates may have significant uncertainties over some regions, including South Africa, West America, Amazon, and central Asia (Figures S16-S18), and hence caution should be taken when interpreting the regional wetting/drying trends in the future scenarios over these regions.

When looking into the respective contributions of P, E, and R to the changes in P-E-R, we find P controls the variations of P-E-R over the majority of the land, including North America, Australia, East Russia, North Europe, and North Africa. Because the trends in P over these regions are apparently larger than those of E and R, resulting in good agreement with P-E-R. Similarly, E governs the changes in P-E-R for southern Africa, Northwest India, South China, most of Europe and central Russia. It is worth noting that P, E, and R jointly cause the changes in P-E-R for South America since P and E/R have opposite trends based on the observational products. The South China Sea Islands, including Indonesia and Malaysia, present consistent increasing trends in P, E, and R, thus, the approximately identical contribution of these variables can be attributed. However, it should be noted that the variability of either of these three water balance components (or their combination) may not always translate to the changes in TWSA because human interventions such as reservoir impoundment, water diversion, and groundwater pumping may substantially alter the natural water cycle, as we have discussed previously by taking the Yangtze River basin as an example (e.g., filling of the reservoirs). Although these changes can also be included in the climatic and hydrologic observations in an indirect/implicit way (e.g., increase of E from water impoundment or increase in soil moisture from infiltration), these signals are very difficult to be captured given the considerable uncertainty in different datasets, causing the nonclosure of the water balance at global scale (Lehmann et al., 2022). In this case, the assessment of the dryness/wetness from the TWSA perspective becomes more needful and convincing.

3.2 Future projections using ensemble CMIP6 outputs

We project the multi-model ensemble mean trends under different climate change scenarios (SSP126, SSP245, and SSP585) during the future period 2071-2100 using both TWS-DSI and P-E-R (Figures 1, 2, S14, and S15). Favorably good agreement between TWS-DSI and P-E-R is detected, with the latter presenting a less significant trend, similar to the observations made in previous studies (Lv et al., 2019; Lv et al., 2021). The general consistency might be associated with the incomplete considerations of human interventions in GCMs. However, we also discover the differences in TWS-DSI and P-E-R over the high-latitude regions such as northern North America and Russia, which shows the wetting trend in P-E-R due to precipitation increase while drying in TWS-DSI probably because of the snow melt under global warming. GCMs present higher spatial heterogeneity than the historical datasets such as GHMs and LSMs, possibly due to the original coarse spatial resolution of the GCMs and the biases in the models. Specifically, all three scenarios confirm the significant ($p<0.05$) wetting trends in North China, South Mongolia, central Asia, northern border of Canada, and South Europe, with the increase in the intensity and spread along with the enhancement of climate scenarios (Figures 1, 2, S14, and S15). Similarities are found in the drying trends in the majority of Russia, northern North America, and South Africa. The wetting trends are apparently caused by the increase in precipitation (Figure S16) (Milly et al., 2005; Seneviratne et al., 2006). The arid Arab region is also projected to become wetter because of the increase in precipitation and the decrease in evapotranspiration. On the contrary, the drying trends are mainly controlled by the rapidly intensifying evapotranspiration in a warming climate (Figure S17) (Allen et al., 2010; Vicente-Serrano et al., 2010), with the precipitation and runoff slightly increasing (Figures S16 and S18). The obvious drying trend around Canada's subarctic lakes is attributed to the high vulnerability to droughts when snow cover declines under increasing temperature (Bouchard et al., 2013). However, there exist scenario-variable divergences over the continents of South America, Australia, India, and the Mediterranean basin, which are generally caused by the various patterns in precipitation under different scenarios with the increasing evapotranspiration over there. The runoff also follows the patterns of precipitation but with comparably lesser magnitudes.

We conduct a regional study for the QTP as an indicator for global climate change and to demonstrate the temporal changes in the regional dryness/wetness during 1985-2100 (Figures S19-20). A significant decrease in TWSA and the derived TWS-DSI is observed during the reference period 1985-2014 based on different datasets except for the WGHM output. The depletion trend is consistent with previous studies reporting the sublimation/ablation of glaciers and ice caps due to climate warming over decades (Huang et al., 2013, 2021). The drying QTP is also evidenced by the metric P-E-R with a non-significant trend based on various datasets, in which both precipitation and evapotranspiration increase. In addition, the QTP is expected to undergo continuous drying trends based on TWSA and TWS-DSI stemming from a warming climate, which can be more intensive under higher climate scenarios from SSP245 and SSP585 conditions (Figure S19). Similarly, regional precipitation and evapotranspiration also show increasing patterns, with the runoff generally unchanged (except during the end of the 21st century under the SSP585 scenario). However, the variable P-E-R does not present the decreasing trends as TWSA (and TWS-DSI). The differences might be attributable to the biases in the projected evapotranspiration and runoff, which might underestimate some key components such as an increase in sublimation and surface runoff due to warming-induced melt of ice, snow, and glaciers. Despite this, it is worth noting that the modelled TWS-DSI-based evaluation can also overestimate the true trend of the land mass because the important surface water is not physically considered, especially in the context of significantly growing lake volume over the QTP (Zhang et al., 2021).

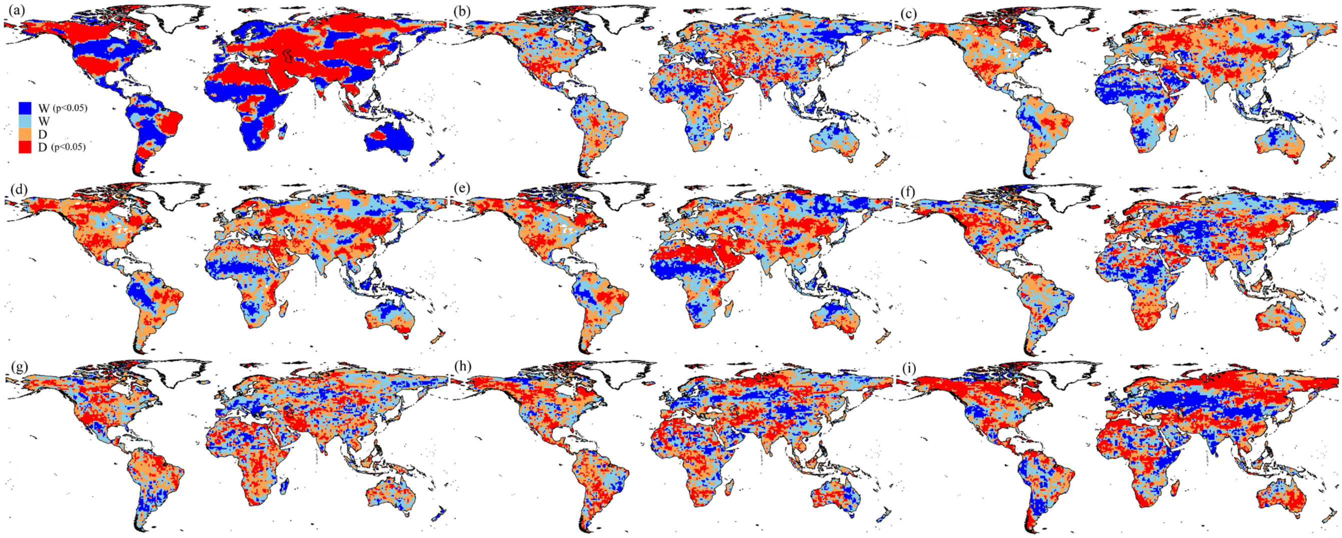

Figure 1. Global distribution of the classification in long-term trends in TWS-DSI during (a-f) the historical (1985-2014) and future (2071-2100) period under (g) SSP126, (h) SSP245, and (i) SSP585 scenarios. Note: The historical results are based on the (a) GRACE reconstruction, (b) WGHM, (c) VIC, (d) CLSM, (e) Noah, and (f) ensemble mean of eight GCMs, respectively. The future results are based on the ensemble of eight GCMs. "D" and "W" indicate regions with drying and wetting trends, respectively.

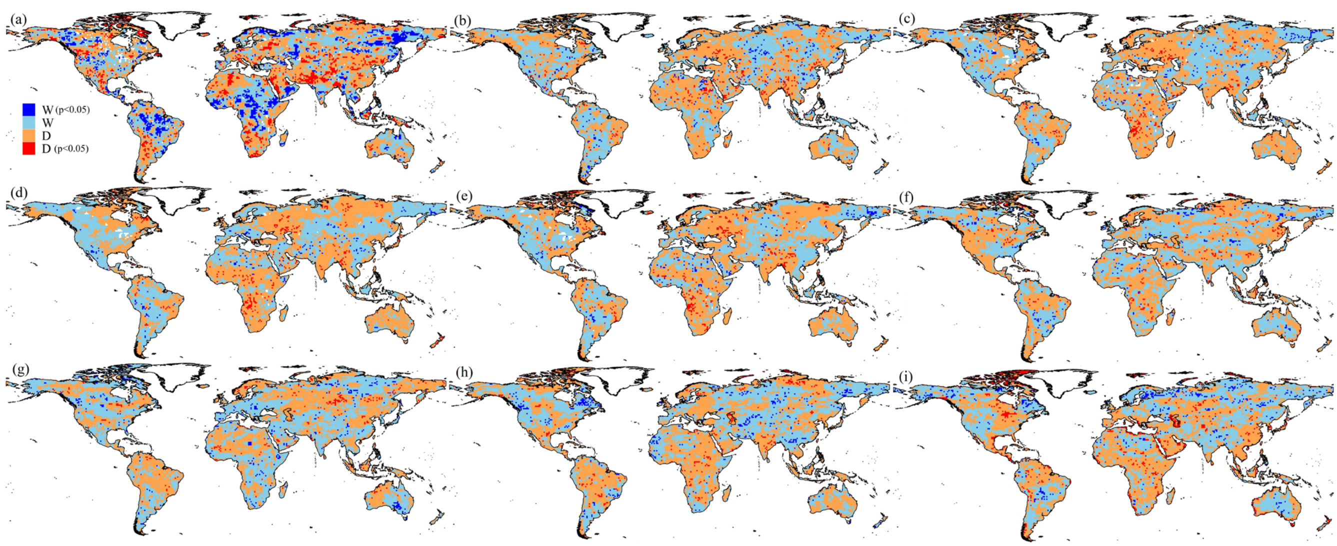

Figure 2. Global distribution of the classification in long-term trends in P-E-R during (a-f) the historical (1985-2014) and future (2071-2100) period under (g) SSP126, (h) SSP245, and (i) SSP585 scenarios. Note: The historical results are based on the (a) observation-based products (i.e., CRU P, GLEAM E, and GRUN R), (b) WGHM, (c) VIC, (d) CLSM, (e) Noah, and (f) ensemble mean of eight GCMs, respectively. The future results are based on the ensemble of eight GCMs. "D" and "W" indicate regions with drying and wetting trends, respectively.

3.3 Assessment of the DDWW Paradigm

Combined with the climate regions classified by AI, we further test the DDWW paradigm at a 5% significance level using both TWS-DSI and P-E-R over global land in the past and future (Figures 3 and 4). We observe apparent consistency in the spatial distribution of the test results based on different indices except for the high-latitude regions under future projections, in line with the long-term trend estimations. While the land area having significant patterns from TWS-DSI is more than that from P-E-R as investigated previously. In addition, different datasets (e.g., GHMs and LSMs) produce reasonably consistent spatial distributions except for the regional variabilities over certain regions such as North Africa. We also note that relatively larger biases could occur in several regions including western United States and central Asia, highlighting the uncertainties in the future projections based on the CMIP6 GCMs. As reported in Table R3, limited proportions (<10%) of area illustrating the "transition gets drier" (TD) and "transition gets wetter" (TW) patterns are estimated in both past and future periods. Much of the land area over the Arab regions, East Asia, and southwest United States show the "dry gets drier" (DD) phenomenon. In contrast to that, a substantial portion of area over the arid regions of the north and south of Africa, Australia, and central Asia shows the "dry gets wetter" (DW) hypothesis. Moreover, the "wet gets wetter" (WW) paradigm is confirmed mainly in East Russia, north Amazon, South China, and East United States, with the "wet gets drier" (WD) pattern happening in central Africa, eastern Amazon, middle Europe. western Canada, and North Asia. The differences between test results from TWS-DSI and P-E-R are mainly in South China and northern lands of the Caspian Sea, which are caused by the divergent meanings in the metrics. For example, a significant increase in E over South China is shown as the drying trends of P-E-R, instead of the wetting trends of TWS-DSI induced by the extensive reservoir impoundment (e.g., Three Gorges Dam). The differences are highlighted by the future projections over high-latitude regions such as northern Russia and North America as well as central Africa, especially under the SSP585 scenario. Despite this, a similar pattern revealed by both variables under the SSP126 scenario shows the continued tendency when compared with the historical results (Figures 3 and 4). However, some regions like South Europe and southeastern South America present strong wetting trends due to an increase in precipitation (Coppola et al., 2021), the opposite changes are discovered over northern South America. Nevertheless, the SSP245 scenario presents a slightly different distribution from historical results, with many regions in the north and center of Asia and central Europe showing DW and WW situations instead of DD and WD. In addition to that, the south and northeast parts of China, together with the majority of Russia show the WD situation, while the DD paradigm is gradually dominating Australia. This difference is further confirmed based on the results under the SSP585 scenario (Figures 3 and 4). These results correspond well with the climatic and hydrologic fluxes such as P, E, and R as well as their residuals (P-E-R), indicating the consistency between the atmospheric and terrestrial conditions under climate change.

Global statistics of the regions with various patterns during the historical (1985-2014) and future periods (2071-2100) are shown in Figure 5. During the period 1985-2014, a percentage of as high as 82.8% of the land area shows significant trends in either wetting or drying ($p<0.05$) based on the GRACE reconstruction. Further, 40.84% of the area shows the DDWW paradigm, in which 20.17% and 20.67% of the area is drying and wetting, respectively. 35.43% of the area, however, shows the opposite pattern of DW (16.13%) and WD (19.30%), respectively. The percentages of the global land supporting/opposing the DDWW paradigm from the GHMs and LSMs are relatively lower than those from the GRACE reconstruction using TWS-DSI, which are reflected by the fewer proportions with significant trends. For example, the percentage of the land area showing the DDWW paradigm ranges from 11.01% (VIC) to 18.95% (Noah), and from 10.21% (WGHM) to 16.4% (VIC) for the opposite pattern. The test results based on P-E-R indicate a similar mismatch of the DDWW paradigm with 12.54% and 6.62% of the land area validating and combating the DDWW paradigm, respectively based on the observational products (Figure S21 and Table S4). Nevertheless, GHMs and LSMs report non-significant trends ($p>0.05$) over more than 90% of land area. In short, the confirmed percentage for the DDWW paradigm (11.01% to 40.84%) for the land mass (represented by TWS-DSI) in our study is higher than that for the land surface (represented by precipitation, evaporation, and aridity) in a previous study (10.8%) (Greve et al., 2014). Feng and Zhang (2015) used soil moisture to conclude a proportion of 15.12% followed the DDWW pattern while a percentage of 7.7% of the land showed an opposite pattern between 1979 and 2013, which is relatively lower than our study. Yang et al. (2019) applied a combined measure employing six different drought indices to evaluate the DDWW paradigm and discovered the percentage following and opposing the DDWW paradigm is 29% and 20%, respectively, during the period 1982-2012, typically consistent with our study. Cheng et al. (2020) utilized the GRACE data during 2002-2017 and reported the area having the DDWW pattern reached 11.2% except for the 4.7% of cold regions over global land, which is comparatively lower than our study. Observed differences among various studies are attributed to the differences in datasets used, metrics employed for assessment and their governing mechanisms, and the study period.

In climate model projections, the proportion of areas supporting the DDWW paradigm is 14.66%, 14.26%, and 17.08% under SSP126, SSP245, and SSP585 scenarios, respectively for TWS-DSI. Alternatively, the fraction of the global land area having the opposite DDWW pattern achieves 13.84%, 18.72%, and 26.64%, respectively. The percentage of areas with significant wetting and drying trends slightly increases over the enhancement of emission scenarios, consistent with the increase of DDWW-validated areas from SSP126 to SSP585 scenarios (Figures 3 and 4). The evaluation results from the perspective of P-E-R are generally lower than 5% because of the non-significant trends in the variable, highlighting the unsupported DDWW paradigm in this regard. However, as we have mentioned previously, the internal variability of climate models might affect the potential agreement with the DDWW pattern (Kumar et al., 2015), which is also reflected by the differences between the GCMs and different models/products during the historical period (Tables S3-S4). Greve and Senevirtne (2015) used climate projections from CMIP5 to establish the measure $P - E$ for the assessment of the DDWW paradigm and discovered the hypothesis was validated over 19.5% of land area between 2080 and 2100 under the RCP8.5 scenario, which is close to our result (17.08%). Moreover, Li et al. (2021a) further applied the P-E index to test the DDWW theory based on GCMs from the third phase of Paleoclimate Modelling Intercom-parison Project (PMIP3) simulations, concluding a similar proportion of 22.8% of the global land to our study that held the DDWW paradigm. This similarity reveals the consistent terrestrial responses to the atmospheric variations under future warming for both metrics.

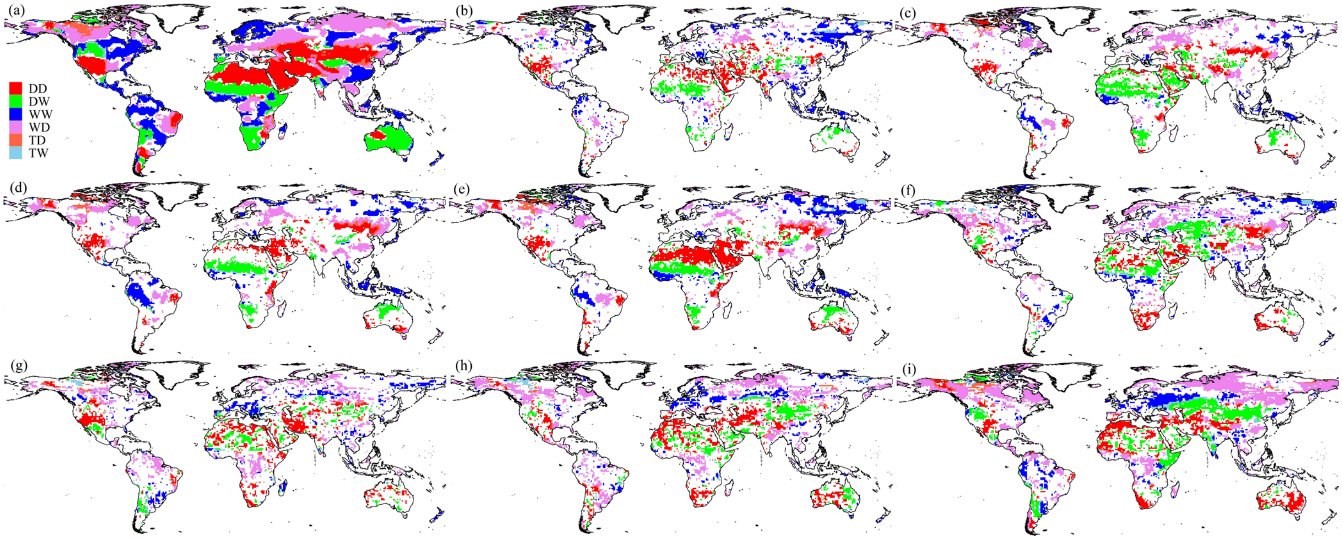

Figure 3. Global assessment of the DDWW paradigm based on TWS-DSI during the (a-f) historical (1985-2014) and (g-i) future (2071-2100) period under (g) SSP126, (h) SSP245, and (i) SSP585 scenarios. Note: The historical results are based on the (a) GRACE reconstruction, (b) WGHM, (c) VIC, (d) CLSM, (e) Noah, and (f) ensemble mean of eight GCMs, respectively. The future results are based on the ensemble of eight GCMs. DD indicates the dry gets drier; DW indicates the dry gets wetter; WW indicates the wet gets wetter; WD indicates the wet gets drier; TD indicates the transition gets drier; TW indicates the transition gets wetter.

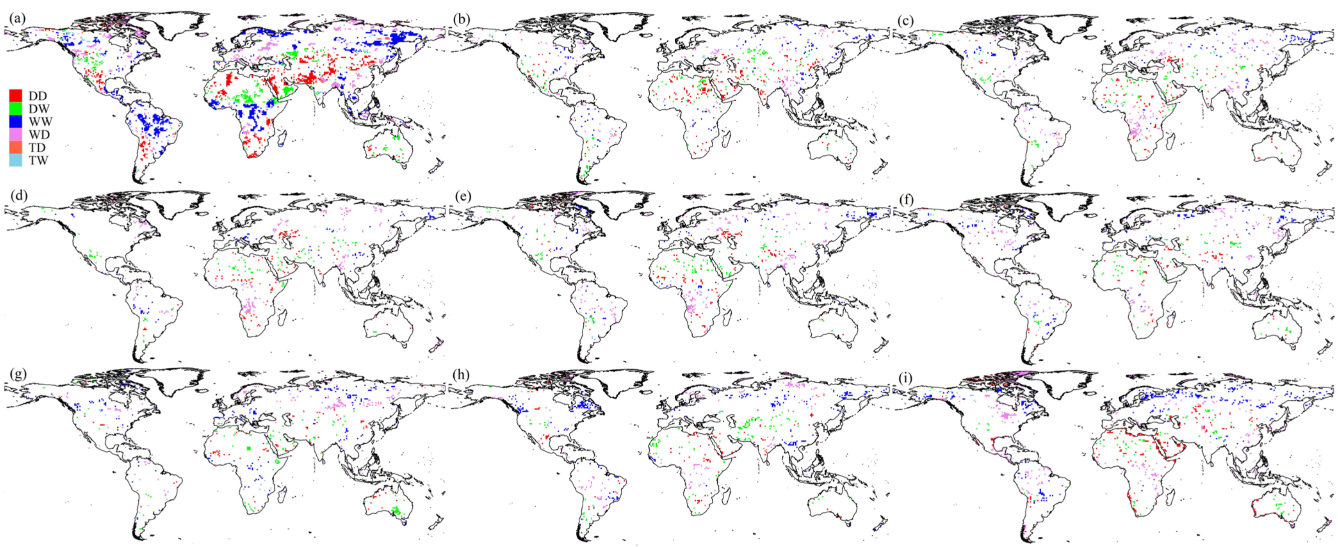

Figure 4. Global assessment of the DDWW paradigm based on P-E-R during the (a-f) historical (1985-2014) and future (2071-2100) period under (g) SSP126, (h) SSP245, and (i) SSP585 scenarios. Note: The historical results are based on the (a) observation-based products (i.e., CRU P, GLEAM E, and GRUN R), (b) WGHM, (c) VIC, (d) CLSM, (e) Noah, and (f) ensemble mean of eight GCMs, respectively. The future results are based on the ensemble of eight GCMs. DD indicates the dry gets drier; DW indicates the dry gets wetter; WW indicates the wet gets wetter; WD indicates the wet gets drier; TD indicates the transition gets drier; TW indicates the transition gets wetter.

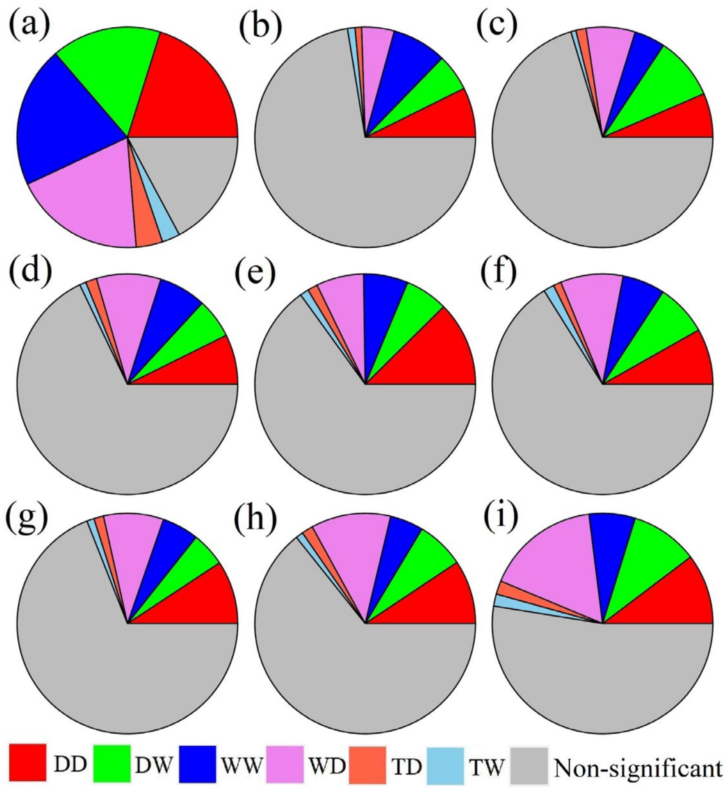

Figure 5. Fraction of the global land area (in percentage) with different patterns during the (a-f) historical (1985-2014) and (g-i) future (2071-2100) period under (g) SSP126, (h) SSP245, and (i) SSP585 scenarios based on TWS-DSI. Note: DD indicates the dry gets drier; DW indicates the dry gets wetter; WW indicates the wet gets wetter; WD indicates the wet gets drier; TD indicates the transition gets drier; TW indicates the transition gets wetter; Non-significant indicates the regions showing non-significant ($p>0.05$) trends in TWS-DSI.

3.4 Uncertainties, Implications, and Way Forward

Each ensemble member of the datasets used in this study has embedded uncertainties inherently originating from one or more forcing variables, simplified assumptions of complex processes in the models and their physical structure, retrieval algorithms, and systematic biases, which might have inevitably propagated to the results presented herein. For example, the original GRACE mascon observations contain the measurement error and signal leakage at the gridded scale, which persists in the reconstruction of TWSA when training via statistical methods (Li et al., 2021a). Unlike observed GRACE and reconstructed GRACE-like data, simulations from the models (GHMs, LSMs, and GCMs) are inherently featured by incomplete TWSA representation (Table S1). They are generally based on simplified hydrological processes, resulting in the lack of certain TWSA components. For example, the widely used Noah and VIC models lack surface water and groundwater storage in TWSA (Scanlon et al., 2018). Similarly, GCMs can only simulate the snow water and soil moisture within a limited depth from 2 to 10 m below the land surface (Xiong et al., 2022a). This inadequate representation of TWSA (and hence TWS-

DSI) in these global models can lead to regional bias in some aquifers with overexploitation of the particular TWSA components (e.g., groundwater depletion in North China Plain) and therefore should be cautioned especially dealing with the seasonal analyses. Overall, the models with completed TWS components are more suitable for assessing the TWSA changes at the global scale for future research, such as the continuously developing hyper-resolution global hydrological models (e.g., WGHM), which can help to avoid the uncertainty associated with the lack of key TWSA elements in most LSMs (e.g., surface water and groundwater) (Pokhrel et al., 2021).

Moreover, the eight CMIP6 GCMs are forced with the future projections of many meteorological variables such as precipitation and air temperature, which have been reported to show variable-specific biases over the global land (Eyring et al., 2016; Kim et al., 2020). Despite employing bias correction with GRACE data, uncertainty from the forcing and models can influence the accuracy of TWSA simulations (Xiong et al., 2022a). Advanced bias-correction methods (e.g., Lange, 2019 and Francois et al., 2020) might play critical roles in reducing such errors in meteorological variables for future hydrologic impact studies, especially when combined with the start-of-the-art GHMs and LSMs as mentioned above. The inclusion of more GCMs can also help to estimate the uncertainties in the meteorological inputs in climate change scenarios. Although it is challenging to explicitly attribute and quantify these uncertainties in the absence of a ‘true’ reference observation dataset, the ensemble averaging method has been used to integrate the multi-source TWSA data. Moreover, since the meaning and hence the results and interpretation of the 'dry' and 'wet' varies across disciplines, land or ocean, target variable(s), and the problem in question (Roth et al., 2021), future studies may focus on various spatial (e.g., local, regional, basin, zonal averages) and temporal (monthly, seasonal, annual) scales using our processed data with additional model outputs (e.g., more number of GCMs).

To investigate the influence of different models on the robustness of the evaluation for the DDWW paradigm, we carry out an independent analysis at the individual member level during the future period 2071-2100 (see Figure S22). We find the differences among different members of the CMIP6 archive. The GFDL-ESM4 and MIROC6 models present overestimations, but the IPSL-CM6A and CanESM5 models underestimate different percentages compared with the ensemble mean. Specifically, the area dominated by the DDWW paradigm changes from 8.16% (ACCESS-ESM1-5) to 19.36% (MIROC6), while that showing the opposite pattern ranges from 7.33% (CanESM5) to 14.57% (MPI-ESM1-2-HR) under the SSP126 scenario. For the SSP245 scenario, the DDWW-validated regions account for 6.98% (CanESM5) to 18.54% (GFDL-ESM4); the opposite pattern occurs over a range from 8.71% (CanESM5) to 12.64% (MPI-ESM1-2-HR) of land. The proportion supporting the DDWW paradigm varies from 9.71% (CanESM5) to 20.08% (GFDL-ESM4), while that presenting the opposite pattern ranges from 8.19% (MPI-ESM1-2-LR) to 18.68% (ACCESS-CM2) under the SSP585 scenario. Overall, the comparatively large difference among various models might source from unforced internal climate variability of distinctive CMIP6 members and different emission scenarios (Kumar et al., 2015).

Our choice of the significance level (i.e., 0.05) may also affect the rationale of the DDWW examination results. Therefore, different significance levels are alternatively tested (see Figures S23-S24 and Tables S5-S6). At a significance level of 0.01, a decrease in 3.21% (37.63%) of the land area agreeing well with the DDWW theory is detected, with a reduction of 2.65%

(32.78%) in area illustrating the opposite pattern during the period 1985-2014 for the GRACE reconstruction. Similar decreases in the proportion of the DDWW-dominated area ranging from 5.19% (SSP245) to 7.2% (CLSM) are also discovered in the GHMs, LSMs, and GCMs. As for the 0.1 significance level, the DDWW-validated regions account for 42.49% (+1.65%) of the total area, with 36.89% (+1.46%) of land agreeing with the opposite hypothesis compared to those at the 0.05 level. In the future period, a similar pattern is discovered that both DDWW-confirmed and DDWW-opposed regions are increasing on account of the enhancement of projected strength of radiative forcing, with the reduction of the area showing non-significant trends in wetting and drying. However, the magnitudes of results at the 0.01 significance level are generally lower than that at the 0.1 significance level due to the different thresholds of the detected trends in drying and wetting. Considering the similar tendency with marginal effects of the varying choices of the p-value (e.g., 4.86% change in DDWW area from 0.01 to 0.1 level for the GRACE reconstruction during 1985-2014), our adopted significance level (i.e., 0.05) can reasonably and robustly explain the global trends of dryness/wetness. Given the inherent magnitude bias from various GCMs projections, the ensemble averaging method has the potential to provide alternative estimates over data-sparse areas globally like Africa and central Asia.

Despite the multisource uncertainties, our study provides important implications for the long-term trends in dryness/wetness of the global land mass in the past and future from the perspective of TWSA. Compared with other widely used indexes that are purely derived from the hydrometeorological variables (e.g., SPI, SPEI, and PDSI) or incorporate a single component of the TWSA (e.g., SSI, SGI, and SRI), our developed TWS-DSI is able to describe the overall status of the land system, which is jointly influenced by different components including soil moisture, river runoff, and groundwater that play different roles in the hydrological cycle (Tapley et al., 2019). Although other indices may undoubtedly perform at par for the specific variable in question, they tend to present equivocal inferences for the total water storage. It can be easily understood by the example of soil moisture or evapotranspiration-based indices in a highly irrigated area such as the Ganges river basin. TWS is unremittingly declining due to the overexploitation of groundwater for agriculture in this region (Rodell et al., 2009), while E or soil moisture may have positive trends, thus attenuating the actual TWS situation. Moreover, the adopted TWS-DSI is suitable and feasible for comparing dryness/wetness status for different locations and periods (Zhao et al., 2017). Furthermore, the projected changes in global TWSA and associated TWS-DSI improve our understanding of the large-scale hydrological response to climate change, particularly in regions with strong human interventions, such as the south and east of Asia.

4 Conclusion

In this study, the historical TWS-DSI monthly time series over global land during 1985-2014 is calculated from an ensemble of two GHMs (VIC and WGHM), two LSMs (Noah and CLSM), and one GRACE reconstruction. In addition, future projections of TWS-DSI from 2071 to 2100 under SSP126, SSP245, and SSP585 scenarios are derived from the average of eight selected CMIP6 GCMs after bias-correction using GRACE observations. Subsequently, we detect the long-term trends in dryness/wetness in both the past and future periods based on TWS-DSI. Further, the DDWW paradigm has been evaluated with a significance level of 0.05 from the perspective of terrestrial water storage change. We also establish the metric P-E-R based on multiple observational products and from the same models as the TWS-DSI for comparison. The uncertainty sourced from different choices of models, methods, and confidence levels has been discussed systematically. The new findings are summarised as follows.

(1) During the historical period, the percentages of global land area presenting significant ($p<0.05$) drying and wetting trends range from 13.06% (WGHM)-43.35% (GRACE reconstruction) and 13.7% (CLSM)-39.43% (GRACE reconstruction), respectively. The wetting trends are mainly in North Australia, North and South Africa, South and Northwest China, western South America, central United States, and East Russia. While the drying trends are found in Arab region, West Brazil, Northeast Asia, and southern and northern American continent. During the future period under climate change, the proportion of drying areas (always ~10% higher than wetting) with a significant slope increases from SSP126 (19.52%) to SSP585 (29.04%) scenario. A similar change is detected in the percentage with significant wetting trends, which reaches 11.48%, 13.01%, and 18.42% under SSP126, SSP245, and SSP585 scenarios, respectively.

(2) A total of 11.01% (VIC) to 40.84% (GRACE reconstruction) of the global land area shows the DDWW paradigm valid, in which the drying and wetting area account for 6.47% (VIC)-20.17% (GRACE reconstruction) and 4.54% (VIC)-20.67% (GRACE reconstruction), respectively during the period 1985-2014. However, the area showing the opposite patterns, like “dry gets wetter” (DW) or “wet gets drier” (WD), account for the 10.21% (WGHM)-35.43% (GRACE reconstruction) of the global land, respectively. The proportion of areas supporting (opposing) the DDWW paradigm is 14.66% (16.76%), 14.26% (18.72%), and 17.08% (26.64%) under SSP126, SSP245, and SSP585 scenarios, respectively.

(3) Parallel estimates of the water balance variables and their comparison with the TWSA-based analysis, on the one hand, shed light on the governing mechanisms and translation of hydrometeorological fluxes to the land water storage, on the other hand, outline additional insights into the varying and sometimes even contrasting behavior of the various metrics.

(4) Sensitivity analysis on different choices of significance levels from 0.01 to 0.1 for the long-term trends indicates similar patterns, in which the maximum decrease (increase) in the DDWW-validated regions reaches –7.4% (4.47% historically under the 0.01 (0.1) level, respectively. Such consistency is also evidenced by the projected TWS-DSI in the future under various scenarios. Moreover, independent experiments based on the individual TWSA datasets suggest that the divergent data sources might lead to model-variable biases for both the DDWW-agreed and DDWW-opposed patterns. The use of distinctive GCMs also suggests slightly overrated (e.g., GFDL-ESM4) and underrated (e.g., CanESM5) percentages of such patterns in the future under multiple emission scenarios.

New insights from the TWSA perspective highlight that the widely-used DDWW paradigm is still challenged in both historical and future periods under climate change. The differences between test results based on P-E-R imply the robustness of our developed TWS-DSI in capturing the total land water variations induced by climate changes and human activities, suggesting potentially new knowledge in the land hydrology field. The regional aggregation of our study in the Qinghai-Tibetan Plateau can provide important inferences for decision-makers and stakeholders for the sustainable management and efficient utilization of water resources under global change.

Data Availability

The data used in this study are open access and available at: GRACE solution(http://www2.csr.utexas.edu/grace/), GRACE reconstruction (https://doi.org/10.1029/2021GL093492), GHMs (WGHM, https://gmd.copernicus.org/articles/14/1037/2021/; VIC, https://ldas.gsfc.nasa.gov/gldas), LSMs (Noah, https://ldas.gsfc.nasa.gov/gldas;CLSM, https://ldas.gsfc.nasa.gov/gldas), GCMs (https://esgf-node.llnl.gov/projects/cmip6/), Climatic and hydrologic datasets (Precipitation and potential evapotranspiration, https://crudata.uea.ac.uk/cru/data/hrg/cru_ts_4.06/; Runoff, https://doi.org/10.1029/2020WR028787; Evapotranspiration; https://www.gleam.eu/).

Supplement

The supplement related to this article is available online

Author contributions

Jinghua Xiong conceived and designed the experiments. Jinghua Xiong performed the experiments. Jinghua Xiong and Abhishek analyzed the data. Jinghua Xiong, Shenglian Guo, Abhishek, Jie Chen, and Jiabo Yin wrote and edited the paper.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgments

This study was financially supported by the National Key Research and Development Program of China (2021YFC3200303), the National Natural Science Foundation of China [U20A20317]. The numerical calculations in this paper have been done on the supercomputing system in the Supercomputing Center of Wuhan University.

References

Abhishek, Kinouchi, T., and Sayama, T.: A comprehensive assessment of water storage dynamics and hydroclimatic extremes in the Chao Phraya River Basin during 2002–2020, J. Hydrol., 603, https://doi.org/10.1016/j.jhydrol.2021.126868, 2021.

AghaKouchak, A., 2014. A baseline probabilistic drought forecasting framework using standardized soil moisture index: application to the 2012 United States drought. Hydrol. Earth Syst. Sci. 18, 2485–2492. https://doi.org/10.5194/hess-18-2485-2014

Allan, R.P., Soden, B.J., John, V.O., Ingram, W., Good, P., 2010. Current changes in tropical precipitation. Environ. Res. Lett. 5, 025205. https://doi.org/10.1088/1748-9326/5/2/025205

Allen, C.D., Macalady, A.K., Chenchouni, H., Bachelet, D., McDowell, N., Vennetier, M., Kitzberger, T., Rigling, A., Breshears, D.D., Hogg, E.H., Gonzalez, P., Fensham, R., Zhang, Z., Castro, J., Demidova, N., Lim, J.-H., Allard, G., Running, S.W., Semerci, A., Cobb, N., 2010. A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. For. Ecol. Manag. 259, 660–684. https://doi.org/10.1016/j.foreco.2009.09.001

An, L., Wang, J., Huang, J., Pokhrel, Y., Hugonnet, R., Wada, Y., et al. (2021). Divergent Causes of Terrestrial Water Storage Decline Between Drylands and Humid Regions Globally. Geophysical Research Letters, 48(23). https://doi.org/10.1029/2021GL095035

Barnard, D.M., Germino, M.J., Bradford, J.B., Connor, R.C., Andrews, C.M., Shriver, R.K. 2021. Are drought indices and climate data good indicators of ecologically relevant soil moisture dynamics in drylands? Ecol. indic. 133, 108379. https://doi.org/10.1016/j.ecolind.2021.108379

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, https://doi.org/10.1038/sdata.2018.214

Bouchard, F., Turner, K.W., MacDonald, L.A., Deakin, C., White, H., Farquharson, N., Medeiros, A.S., Wolfe, B.B., Hall, R.I., Pienitz, R., Edwards, T.W.D., 2013. Vulnerability of shallow subarctic lakes to evaporate and desiccate when snowmelt runoff is low. Geophys. Res. Lett. 40, 6112–6117. https://doi.org/10.1002/2013GL058635

Burke, E.J., Zhang, Y., Krinner, G. 2020. Evaluating permafrost physics in the coupled model intercomparison project 6 (CMIP6) models and their sensitivity to climate change. Cryosphere., 14 (9) , pp. 3155-3174

Byrne, M.P., O'Gorman, P.A., 2015. The Response of Precipitation Minus Evapotranspiration to Climate Warming: Why the "Wet-Get-Wetter, Dry-Get-Drier'' Scaling Does Not Hold over Land. J. Clim. 28, 8078–8092. https://doi.org/10.1175/JCLI-D-15-0369.1

Chen, J.L., Tapley, B., Rodell, M., Seo, K.W., Wilson, C., Scanlon, B.R., Pokhrel, Y., 2020. Basin-Scale River Runoff Estimation From GRACE Gravity Satellites, Climate Models, and In Situ Observations: A Case Study in the Amazon Basin. Water Resour. Res. 56, e2020WR028032. https://doi.org/10.1029/2020WR028032

Chang, L.-L., Yuan, R., Gupta, H. V., Winter, C. L., Niu, G.-Y. 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 Resources Research, 56, e2020WR027102.

Chen, J.L., Wilson, C.R., Tapley, B.D., 2010. The 2009 exceptional Amazon flood and interannual terrestrial water storage change observed by GRACE. Water Resour. Res. 46, W12526. https://doi.org/10.1029/2010WR009383

Chou, C., Neelin, J.D., Chen, C.-A., Tu, J.-Y., 2009. Evaluating the "Rich-Get-Richer'' Mechanism in Tropical Precipitation Change under Global Warming. J. Clim. 22, 1982–2005. https://doi.org/10.1175/2008JCLI2471.1

Coppola, E., Nogherotto, R., Ciarlo', J. M., Giorgi, F., van Meijgaard, E., Kadygrov, N., Iles, C., Corre, L., Sandstad, M., Somot, S., Nabat, P., Vautard, R., Levavasseur, G., Schwingshackl, C., Sillmann, J., Kjellström, E., Nikulin, G., Aalbers, E., Lenderink, G., Christensen, O. B., Boberg, F., Sørland, S. L., Demory, M.-E., Bülow, K., Teichmann, C., Warrach-Sagi, K., Wulfmeyer, V. 2021. Assessment of the European Climate Projections as Simulated by the Large EURO-CORDEX Regional and Global Climate Model Ensemble, J. Geophys. Res.-Atmos., 126, e2019JD032356, https://doi.org/10.1029/2019JD032356.

Dai, A., 2011. Drought under global warming: a review. Wiley Interdiscip. Rev.-Clim. Change 2, 45–65. https://doi.org/10.1002/wcc.81

Derber, J., Parrish, D., Lord, S., 1991. The New Global Operational Analysis System at the National-Meteorological-Center. Weather Forecast. 6, 538–547. https://doi.org/10.1175/1520-0434(1991)006<0538:TNGOAS>2.0.CO;2

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

Donat, M.G., Lowry, A.L., Alexander, L.V., O'Gorman, P.A., Maher, N., 2016. More extreme precipitation in the world's dry and wet regions. Nat. Clim. Change 6, 508-+. https://doi.org/10.1038/NCLIMATE2941

Durack, P.J., Wijffels, S.E., Matear, R.J., 2012. Ocean Salinities Reveal Strong Global Water Cycle Intensification During 1950 to 2000. Science 336, 455–458. https://doi.org/10.1126/science.1212222

Durbin, J., and Watson, G. S., "Testing for Serial Correlation in Least Squares Regression, I." Biometrika 37 (1950): 409-428.

Durbin, J., and Watson, G. S., "Testing for Serial Correlation in Least Squares Regression, II." Biometrika 38 (1951): 159-179.

Eyring, V., Bony, S., Meehl, G.A., Senior, C.A., Stevens, B., Stouffer, R.J., Taylor, K.E., 2016. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geosci. Model Dev. 9, 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Feng, H., Zhang, M., 2015. Global land moisture trends: drier in dry and wetter in wet over land. Sci. Rep. 5, 18018. https://doi.org/10.1038/srep18018

Feng, W., Zhong, M., Lemoine, J.-M., Biancale, R., Hsu, H.-T., Xia, J., 2013. Evaluation of groundwater depletion in North China using the Gravity Recovery and Climate Experiment (GRACE) data and ground-based measurements: Groundwater Depletion In North China. Water Resour. Res. 49, 2110–2118. https://doi.org/10.1002/wrcr.20192

François, B., Vrac, M., Cannon, A. J., Robin, Y., and Allard, D.: Multivariate bias corrections of climate simulations: which benefits for which losses?, Earth Syst. Dynam., 11, 537–562, https://doi.org/10.5194/esd-11-537-2020, 2020.

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

Gampe, D., Zscheischler, J., Reichstein, M., O'Sullivan, M., Smith, W.K., Sitch, S., Buermann, W., 2021. Increasing impact of warm droughts on northern ecosystem productivity over recent decades. Nat. Clim. Change 11, 772-+. https://doi.org/10.1038/s41558-021-01112-8

Gaughan, A.E., Waylen, P.R., 2012. Spatial and temporal precipitation variability in the Okavango-Kwando-Zambezi catchment, southern Africa. J. Arid Environ. 82, 19–30. https://doi.org/10.1016/j.jaridenv.2012.02.007

Getirana, A., 2016. Extreme Water Deficit in Brazil Detected from Space. J. Hydrometeorol. 17, 591–599. https://doi.org/10.1175/JHM-D-15-0096.1

Ghiggi, G., Humphrey, V., Seneviratne, S. I., Gudmundsson, L. 2021. G-RUN ENSEMBLE: A Multi-Forcing Observation-Based Global Runoff Reanalysis, Water Resour. Res., 57, e2020WR028787, https://doi.org/10.1029/2020WR028787, 2021b.

Goyal, R. K. 2004. Sensitivity of evapotranspiration to global warming: A case study of arid zone of Rajasthan (India). Agricultural Water Management, 69(1), 1– 11.

Greve, P., Orlowsky, B., Mueller, B., Sheffield, J., Reichstein, M., Seneviratne, S.I., 2014. Global assessment of trends in wetting and drying over land. Nat. Geosci. 7, 716–721. https://doi.org/10.1038/NGEO2247

Greve, P., Seneviratne, S.I., 2015. Assessment of future changes in water availability and aridity. Geophys. Res. Lett. 42, 5493–5499. https://doi.org/10.1002/2015GL064127

Guo, M., Yue, W., Wang, T., Zheng, N., Wu, L., 2021. Assessing the use of standardized groundwater index for quantifying groundwater drought over the conterminous US. J. Hydrol. 598, 126227. https://doi.org/10.1016/j.jhydrol.2021.126227

Haacker, E.M.K., Kendall, A.D., Hyndman, D.W., 2016. Water Level Declines in the High Plains Aquifer: Predevelopment to Resource Senescence. Groundwater 54, 231–242. https://doi.org/10.1111/gwat.12350

Hamed, K. H., & Rao, A. R. (1998). A modified Mann-Kendall trend test for autocorrelated data. Journal of hydrology, 204(1-4), 182-196.

Hao, Z., Singh, V.P., 2015. Drought characterization from a multivariate perspective: A review. J. Hydrol. 527, 668–678. https://doi.org/10.1016/j.jhydrol.2015.05.031

Hao, Z., Singh, V.P., Xia, Y., 2018. Seasonal Drought Prediction: Advances, Challenges, and Future Prospects. Rev. Geophys. 56, 108–141. https://doi.org/10.1002/2016RG000549

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 (1). pp. 1-8

Held, I.M., Soden, B.J., 2006. Robust responses of the hydrological cycle to global warming. J. Clim. 19, 5686–5699. https://doi.org/10.1175/JCLI3990.1

Hempel, S., Frieler, K., Warszawski, L., Schewe, J., Piontek, F., 2013. A trend-preserving bias correction: the ISI-MIP approach. Earth Syst. Dyn. 4, 219–236. https://doi.org/10.5194/esd-4-219-2013

Hu, B., Wang, L., Li, X., Zhou, J., & Pan, Y. (2021). Divergent Changes in Terrestrial Water Storage Across Global Arid and Humid Basins. Geophysical Research Letters. https://doi.org/10.1029/2020GL091069

Hu, Z., Chen, X., Chen, D., Li, J., Wang, S., Zhou, Q., Yin, G., Guo, N., 2019. “Dry gets drier, wet gets wetter”: A case study over the arid regions of central Asia. Int. J. Climatol. 39, 1072–1091. https://doi.org/10.1002/joc.5863

Huang, J., Ji, M., Xie, Y., Wang, S., He, Y., Ran, J., 2016. Global semi-arid climate change over last 60 years. Clim. Dyn. 46, 1131–1150. https://doi.org/10.1007/s00382-015-2636-8

Huang, L., Li, Z., Tian, B., Chen, Q., Zhou, J. 2013. Monitoring glacier zones and snow/firn line changes in the Qinghai–Tibetan Plateau using C-band SAR imagery. Remote Sensing of Environment 137, 17–30. https://doi.org/10.1016/j.rse.2013.05.016

Huang, L., Li, Z., Zhou, J. M., Zhang, P. 2021. An automatic method for clean glacier and nonseasonal snow area change estimation in High Mountain Asia from 1990 to 2018, Remote Sens. Environ., 258, 112376, https://doi.org/10.1016/j.rse.2021.112376.

Huntington, T.G., 2006. Evidence for intensification of the global water cycle: Review and synthesis. J. Hydrol. 319, 83–95. https://doi.org/10.1016/j.jhydrol.2005.07.003

Immerzeel, W.W., van Beek, L.P.H., Bierkens, M.F.P. 2010, Climate change will affect the Asian water towers. Science, 328, 1382–1385.

Iqbal, Z., Shahid, S., Ahmed, K., Ismail, T., Ziarh, G.F., Chung, E.-S., Wang, X., 2021. Evaluation of CMIP6 GCM rainfall in mainland Southeast Asia. Atmospheric Res. 254, 105525. https://doi.org/10.1016/j.atmosres.2021.105525

Kim, Y.H., Min, S.K., Zhang, X., Sillmann, J., Sandstad, M. 2020. Evaluation of the CMIP6 multi-model ensemble for climate extreme indices. Weather Clim. Extremes, 29, p. 100269, https://doi.org/10.1016/j.wace:2020.100269

Kumar, S., Allan, R.P., Zwiers, F., Lawrence, D.M., Dirmeyer, P.A., 2015. Revisiting trends in wetness and dryness in the presence of internal climate variability and water limitations over land. Geophys. Res. Lett. 42, 10867–10875. https://doi.org/10.1002/2015GL066858

Lange, S.: Trend-preserving bias adjustment and statistical downscaling with ISIMIP3BASD (v1.0), Geosci. Model Dev., 12, 3055–3070, https://doi.org/10.5194/gmd-12-3055-2019, 2019.

Lehmann, F., Vishwakarma, B. D., Bamber, J. 2022. How well are we able to close the water budget at the global scale? Hydrology and Earth System Sciences, 26(1), 35– 54. https://doi.org/10.5194/hess-26-35-2022

Lelieveld, J., Hadjinicolaou, P., Kostopoulou, E., Chenoweth, J., El Maayar, M., Giannakopoulos, C., Hannides, C., Lange, M.A., Tanarhte, M., Tyrlis, E., Xoplaki, E., 2012. Climate change and impacts in the Eastern Mediterranean and the Middle East. Clim. Change 114, 667–687. https://doi.org/10.1007/s10584-012-0418-4

Li, F., Kusche, J., Chao, N., Wang, Z., Loecher, A., 2021a. Long-Term (1979-Present) Total Water Storage Anomalies Over the Global Land Derived by Reconstructing GRACE Data. Geophys. Res. Lett. 48, e2021GL093492. https://doi.org/10.1029/2021GL093492

Li, Y., Zhang, Y., Ye, W., Zhang, X., 2021b. Global Wet/Dry Patterns and Mechanisms Since the Last Glacial Maximum: A Key to Future Projection. Earths Future 9, e2020EF001907. https://doi.org/10.1029/2020EF001907

Liang, X., Lettenmaier, D., Wood, E., Burges, S., 1994. A Simple Hydrologically Based Model of Land-Surface Water and Energy Fluxes for General-Circulation Models. J. Geophys. Res.-Atmospheres 99, 14415–14428. https://doi.org/10.1029/94JD00483

Liu, X., Yin, Z., Shao, X., Qin, N. 2006. Temporal trends and variability of daily maximum and minimum, extreme temperature events, and growing season length over the eastern and central Tibetan Plateau during 1961-2003. J. Geophys. Res., 111 (D19), 10.1029/2005JD006915

Long, D., Shen, Y., Sun, A., Hong, Y., Longuevergne, L., Yang, Y., Li, B., Chen, L., 2014. Drought and flood monitoring for a large karst plateau in Southwest China using extended GRACE data. Remote Sens. Environ. 155, 145–160. https://doi.org/10.1016/j.rse.2014.08.006

Luthcke, S.B., Sabaka, T.J., Loomis, B.D., Arendt, A.A., McCarthy, J.J., Camp, J., 2013. Antarctica, Greenland and Gulf of Alaska land-ice evolution from an iterated GRACE global mascon solution. J. Glaciol. 59, 613–631. https://doi.org/10.3189/2013JoG12J147

Lv, M., Ma, Z., Chen, L., Peng, S. 2019. Evapotranspiration reconstruction based on land surface models and observed water budget components while considering irrigation. Journal of Hydrometeorology, 20, 2163– 2183. https://doi.org/10.1175/jhm-d-19-0090.1

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

Martens, B., Miralles, D.G., Lievens, H., van der Schalie, R., de Jeu, R.A.M., Fernández-Prieto, D., Beck, H.E., Dorigo, W.A., and Verhoest, N.E.C.: GLEAM v3: satellite-based land evaporation and root-zone soil moisture, Geoscientific Model Development, 10, 1903–1925, 2017.

Meng, F., Su, F., Li, Y., Tong. K. 2019. Changes in Terrestrial Water Storage During 2003–2014 and Possible Causes in Tibetan Plateau. J. Geophys. Res.: Atmos., 124 (6), pp. 2909-2931, 10.1029/2018JD029552

Milly, P.C.D., Dunne, K.A., Vecchia, A.V., 2005. Global pattern of trends in streamflow and water availability in a changing climate. Nature 438, 347–350. https://doi.org/10.1038/nature04312

Moreno-Jimenez, E., Plaza, C., Saiz, H., Manzano, R., Flagmeier, M., Maestre, F.T., 2019. Aridity and reduced soil micronutrient availability in global drylands. Nat. Sustain. 2, 371–377. https://doi.org/10.1038/s41893-019-0262-x

Müller Schmied, H., Cáceres, D., Eisner, S., Flörke, M., Herbert, C., Niemann, C., Peiris, T.A., Popat, E., Portmann, F.T., Reinecke, R., Schumacher, M., Shadkam, S., Telteu, C.-E., Trautmann, T., Döll, P., 2021. The global water resources and use model WaterGAP v2.2d: model description and evaluation. Geosci. Model Dev. 14, 1037–1079. https://doi.org/10.5194/gmd-14-1037-2021

Ndehedehe, C.E., Awange, J.L., Kuhn, M., Agutu, N.O., Fukuda, Y., 2017. Climate teleconnections influence on West Africa's terrestrial water storage. Hydrol. Process. 31, 3206–3224. https://doi.org/10.1002/hyp.11237

Ogou, F.K., Ojeh, V.N., Naabil, E., Mbah, C.I., 2021. Hydro-climatic and Water Availability Changes and its Relationship with NDVI in Northern Sub-Saharan Africa. Earth Syst. Environ. https://doi.org/10.1007/s41748-021-00260-3

Peña-Angulo, D., Vicente-Serrano, S. M., Domínguez-Castro, F., Murphy, C., Reig, F., Tramblay, Y., Trigo, R. M., Luna, M. Y., Turco, M., Noguera, I., Aznárez-Balta, M., García-Herrera, R., Tomas-Burguera, M., and El Kenawy, A.: Long-term precipitation in Southwestern Europe reveals no clear trend attributable to anthropogenic forcing, Environ. Res. Lett., 2020. 15, 094070, https://doi.org/10.1088/1748-9326/ab9c4f.

Perera, A. T. D., Nik, V. M., Chen, D., Scartezzini, J. L., Hong, T. 2020. Quantifying the impacts of climate change and extreme climate events on energy systems; Nat. Energy 5(2) 150–159, https://doi.org/10.1038/s41560-020-0558-0.

Perrone, D. Jasechko, S. 2017. Dry groundwater wells in the western United States. Environ. Res. Lett., 12 (10), p. 104002, 10.1088/1748-9326/aa8ac0

Pham-Duc, B., Papa, F., Prigent, C., Aires, F., Biancamaria, S., and Frappart, F. 2019. Variations of surface and subsurface water storage in the Lower Mekong Basin (Vietnam and Cambodia) from multisatellite observations. Water, 11(1). https://doi.org/10.3390/w11010075

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. Nat. Clim. Change 11, 226–233. https://doi.org/10.1038/s41558-020-00972-w

Polson, D., Hegerl, G.C., 2017. Strengthening contrast between precipitation in tropical wet and dry regions. Geophys. Res. Lett. 44, 365–373. https://doi.org/10.1002/2016GL071194

Rajah, K., O'Leary, T., Turner, A., Petrakis, G., Leonard, M., Westra, S., 2014. Changes to the temporal distribution of daily precipitation. Geophys. Res. Lett. 41, 8887–8894. https://doi.org/10.1002/2014GL062156

Ramillien, G., Frappart, F., Seoane, L., 2014. Application of the Regional Water Mass Variations from GRACE Satellite Gravimetry to Large-Scale Water Management in Africa. Remote Sens. 6, 7379–7405. https://doi.org/10.3390/rs6087379

Rodell, M., Houser, P.R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J.K., Walker, J.P., Lohmann, D., Toll, D., 2004. The global land data assimilation system. Bull. Am. Meteorol. Soc. 85, 381-+. https://doi.org/10.1175/BAMS-85-3-381

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

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(7707): 651–659. 10.1038/s41586-018-0123-1

Roderick, M. L., Sun, F., Lim, W. H., Farquhar, G. D. 2014. A general framework for understanding the response of the water cycle to global warming over land and ocean. Hydrology and Earth System Sciences, 18(5), 1575-1589.

Roth, N., Jaramillo, F., Wang-Erlandsson, L., Zamora, D., Palomino-Ángel, S., Cousins, S. A. 2021. A call for consistency with the terms 'wetter'and 'drier'in climate change studies. Environmental Evidence, 10(1), 1-7.

Ruscica, R.C., Sörensson, A. A., Diaz, L. B., Vera, C., Castro, A., Papastefanou, P., et al. 2021. Evapotranspiration trends and variability in southeastern South America: The roles of land-cover change and precipitation variability. International Journal of Climatology. Accepted AuthorManuscript. https://doi.org/10.1002/joc.7350

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., 2018. Global models underestimate large decadal declining and rising water storage trends relative to GRACE satellite data. Proc. Natl. Acad. Sci. U. S. A. 201704665.

Seneviratne, S.I., Luethi, D., Litschi, M., Schaer, C., 2006. Land-atmosphere coupling and climate change in Europe. Nature 443, 205–209. https://doi.org/10.1038/nature05095

Shugar, D.H., Burr, A., Haritashya, U.K., Kargel, J.S., Watson, C.S., Kennedy, M.C., Bevington, A.R., Betts, R.A., Harrison, S., Strattman, K., 2020. Rapid worldwide growth of glacial lakes since 1990. Nat. Clim. Change 10, 939-+. https://doi.org/10.1038/s41558-020-0855-4

Siebert, S., Burke, J., Faures, J.M., Frenken, K., Hoogeveen, J., Doell, P., Portmann, F.T., 2010. Groundwater use for irrigation - a global inventory. Hydrol. Earth Syst. Sci. 14, 1863–1880. https://doi.org/10.5194/hess-14-1863-2010

Slette, I.J., Smith, M.D., Knapp, A.K., Vicente-Serrano, S.M., Camarero, J.J., Beguería, S., 2020. Standardized metrics are key for assessing drought severity. Glob. Change Biol. 26, e1–e3.

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

Tapley B.D., Bettadpur S, Ries J.C., Thompson P.F., Watkins M.M. 2004. GRACE measurements of mass variability in the Earth system. Science 305(5683):503–505. https://doi.org/10.1126/science.1099192

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. Change 9, 358–369. https://doi.org/10.1038/s41558-019-0456-2

Trenberth, K.E., Dai, A., van der Schrier, G., Jones, P.D., Barichivich, J., Briffa, K.R., Sheffield, J., 2014. Global warming and changes in drought. Nat. Clim. Change 4, 17–22. https://doi.org/10.1038/NCLIMATE2067

Vadim Yu. Grigoriev, Natalia L. Frolova (2018) Terrestrial water storage change of European Russia and its impact on water balance. Geography, Environment, Sustainability, 11, 1, p. 38-50. https://doi.org/10.24057/2071-9388-2018-11-1-38-50

Velicogna, I., Sutterley, T.C., van den Broeke, M.R., 2014. Regional acceleration in ice mass loss from Greenland and Antarctica using GRACE time-variable gravity data. Geophys. Res. Lett. 41, 8130–8137. https://doi.org/10.1002/2014GL061052

Vicente-Serrano, S.M., Beguería, S., López-Moreno, J.I., 2010. A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index. J. Clim. 23, 1696–1718. https://doi.org/10.1175/2009JCLI2909.1

Wan, W., Zhao, J., Popat, E., Herbert, C., Döll, P., 2021. Analyzing the Impact of Streamflow Drought on Hydroelectricity Production: A Global-Scale Study. Water Resour. Res. 57, e2020WR028087. https://doi.org/10.1029/2020WR028087

Wan, W., Xiao, P., Feng, X., Li, H., Ma, R., Duan, H., Zhao, L. 2014. Monitoring lake changes of Qinghai-Tibetan Plateau over the past 30 years using satellite remote sensing data. Chinese Science Bulletin, 59(8), 701– 714. https://doi.org/10.1007/s11434-014-0128-6

Wang, R., Li, L., Gentine, P., Zhang, Y., Chen, J., Chen, X., Chen, L., Ning, L., Yuan, L., Lu, G. 2021. Recent increase in the observation-derived land evapotranspiration due to global warming Environ. Res. Lett. 17 024020

Wang, Z., Li, J., Lai, C., Wang, R.Y., Chen, X., Lian, Y., 2018. Drying tendency dominating the global grain production area. Glob. Food Secur.-Agric. Policy Econ. Environ. 16, 138–149. https://doi.org/10.1016/j.gfs.2018.02.001

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

Wu, J., Miao, C., Tang, X., Duan, Q., He, X., 2018. A nonparametric standardized runoff index for characterizing hydrological drought on the Loess Plateau, China. Glob. Planet. Change 161, 53–65. https://doi.org/10.1016/j.gloplacha.2017.12.006

Wu, R.-J., Lo, M.-H., Scanlon, B.R., 2021. The Annual Cycle of Terrestrial Water Storage Anomalies in CMIP6 Models Evaluated against GRACE Data. J. Clim. 34, 8205–8217. https://doi.org/10.1175/JCLI-D-21-0021.1

Xing, Z., Fan, L., Zhao, L., De Lannoy, G., Frappart, F., Peng, J., Li, X., Zeng, J., Al-Yaari, A., Yang, K., Zhao, T., Shi, Jiancheng, Wang, M., Liu, X., Hu, G., Xiao, Y., Du, E., Li, R., Qiao, Y., Shi, Jianzong, Wen, J., Ma, M., Wigneron, J.-P., 2021. A first assessment of satellite and reanalysis estimates of surface and root-zone soil moisture over the permafrost region of Qinghai-Tibet Plateau. Remote Sens. Environ. 265, 112666. https://doi.org/10.1016/j.rse.2021.112666

Xiong, J., Guo, S., Yin, J., Ning, Z., Zeng, Z., Wang, R. 2022a. Projected changes in terrestrial water storage and associated flood potential across the Yangtze River basin. Sci. Total Environ. 817, 152998. https://doi.org/10.1016/j.scitotenv.2022.152998

Xiong, J., Yin, J., Guo, S., He, S., Chen, J, Abhishek. 2022b. Annual runoff coefficient variation in a changing environment: a global perspective. 6, 064006. 10.1088/1748-9326/ac62ad.

Xu, Z., Cheng, L., Liu, P., Makarieva, O., Chen, M., 2021. Detecting and quantifying the impact of long-term terrestrial water storage changes on the runoff ratio in the head regions of the two largest rivers in China. J. Hydrol. 601, 126668. https://doi.org/10.1016/j.jhydrol.2021.126668

Yang, T., Ding, J., Liu, D., Wang, X., Wang, T., 2019. Combined Use of Multiple Drought Indices for Global Assessment of Dry Gets Drier and Wet Gets Wetter Paradigm. J. Clim. 32, 737–748. https://doi.org/10.1175/JCLI-D-18-0261.1

Yi, W., Feng, Y., Liang, S., Kuang, X., Yan, D., Wan, L., 2021. Increasing annual streamflow and groundwater storage in response to climate warming in the Yangtze River source region. Environ. Res. Lett. 16, 084011. https://doi.org/10.1088/1748-9326/ac0f27

Zeng, H., Wu, B., Zhang, N., Tian, F., Phiri, E., Musakwa, W., Zhang, M., Zhu, L., Mashonjowa, E., 2019. Spatiotemporal Analysis of Precipitation in the Sparsely Gauged Zambezi River Basin Using Remote Sensing and Google Earth Engine. Remote Sens. 11(24), 2977. https://doi.org/10.3390/rs11242977

Zhang, C., Tang, Q., Chen, D., Li, L., Liu, X., Cui, H. 2017. Tracing changes in atmospheric moisture supply to the drying Southwest China. Atmos. Chem. Phys., 17 (17), pp. 10383-10393, 10.5194/acp-17-10383-2017

Zhang, G., Ran, Y., Wan, W., Luo, W., Chen, W., Xu, F. 2021. 100 years of lake evolution over the Qinghai-Tibet pLateau Earth Syst. Sci. Data Discuss. , pp. 1-26, 10.5194/essd-2021-130

Zhao, M., Geruo, A., Velicogna, I., Kimball, J.S., 2017. Satellite Observations of Regional Drought Severity in the Continental United States Using GRACE-Based Terrestrial Water Storage Changes. J. Clim. 30, 6297–6308. https://doi.org/10.1175/JCLI-D-16-0458.1

Zhao, M., Geruo, A., Zhang, J., Velicogna, I., Liang, C., Li, Z., 2021. Ecological restoration impact on total terrestrial water storage. Nat. Sustain. 4, 56-U85. https://doi.org/10.1038/s41893-020-00600-7

Zhong, M., Duan, J., Xu, H., Peng, P., Yan, H., Zhu, Y. , 2009. Trend of China land water storage redistribution at medi-and large-spatial scales in recent five years by satellite gravity observations, Chinese Sci. Bull., 54, 816–821, https://doi.org/10.1007/s11434-008-0556-2.

Zmijewski, K., Becker, R. 2014. Estimating the effects of anthropogenic modification on water balance in the Aral Sea watershed using GRACE: 2003–12. Earth Interact. 18, 1–16.