Emerging global freshwater challenges unveiled through observation-constrained projections

Fei Huo[1], Yanping Li[1], Zhenhua Li[1]

[1]Department of Physics and Astronomy, Western University, London, N6G 2V4, Canada.

Correspondence to: Yanping Li (yli4972@uwo.ca)

Abstract. Future hydrological projections exhibit significant discrepancies among models, undermining confidence in the predicted magnitude and timing of hydrological extremes. Here we show that observation-constrained changes in global mean terrestrial water storage (TWS), excluding Greenland and Antarctica, could be approximately 83 mm lower than raw projections from the Inter-Sectoral Impact Model Intercomparison Project phase 3b (ISIMIP3b) by the end of this century under both the low (SSP1-2.6) and high (SSP3-7.0) future forcing scenarios. Notably, the 95th percentile upper bounds are substantially reduced from 2 mm to −96 mm under the low-emissions scenario and from 8 mm to −105 mm under the high-emissions scenario, revealing a notable overestimation of global freshwater availability in the raw model projections. Global models are intricate process representations, making it challenging to isolate causes of their differences with observations. However, by leveraging the emergent constraint (EC) methodology and inter-model spread to empirically adjust biases against observations, we derive more tightly constrained estimates of future TWS changes than those obtained from conventional, unconstrained approaches. The EC-corrected estimates are substantially lower than the raw ISIMIP3b projections, implying that current water resource planning may underestimate the severity of future water shortages, particularly if global water demand remains stable or continues to rise. Our findings pinpoint the urgent need to reduce model uncertainties and enhance the reliability of future hydrological projections to better inform water resource management and climate adaptation strategies.

1 Introduction

Terrestrial water storage (TWS) encompasses all water stored on and beneath the land surface, representing the net balance of precipitation, evapotranspiration, and runoff (Getirana et al., 2017; Rodell & Famiglietti, 2001). As a critical element of the hydrological cycle, TWS plays a key role in regulating freshwater availability (Rodell & Famiglietti, 2001) and Earth's energy budget (Getirana et al., 2017), supporting freshwater ecosystems (Tapley et al., 2019; Wu et al., 2024), influencing biogeochemical cycles (Rodell et al., 2018), driving socioeconomic development (Scanlon et al., 2023; Vörösmarty et al., 2000), and mitigating sea level rise by enhancing continental water storage (Tapley et al., 2019).

A warming climate impacts TWS by accelerating the hydrological cycle through enhanced evapotranspiration and by modifying global precipitation patterns (A. Dai et al., 2018). These changes can exacerbate freshwater scarcity in many regions

while increasing flood risk in others, highlighting the crucial need for accurate future projections. However, the discrepancies among models in simulating historical hydrological budgets lead to substantial differences in their projections of extreme hydrological events (Herrera-Estrada et al., 2017; Trenberth et al., 2003; Vogel et al., 2018), which in turn reduce confidence in the predicted magnitude and timing of these extremes. These discrepancies can be attributed to various factors, including uncertainties in climate forcing (Scanlon et al., 2018), the absence of key components such as surface water storage, groundwater storage, and human interventions in most land surface models (LSMs), as well as limited storage capacities within both LSMs and global hydrological models (GHMs).

Modeling uncertainties pose as a major challenge to producing reliable projections of future freshwater availability. To address these uncertainties, the emergent constraint (EC) approach has been widely applied to identify statistical relationships across global climate models between observable historical climate metrics and projected future responses (Brient, 2020; Hall et al., 2019). These relationships are not evident in individual models but emerge only when large ensembles are analyzed. Specifically, the EC technique aims to reduce the often uncomfortably large spread in future projections within multi-model ensembles, thereby providing more tightly constrained estimates of variables of interest (Hall et al., 2019). Such improved and physically informed constraints are crucial for climate mitigation and adaptation policymaking. While a large number of ECs have been proposed for global mean temperature changes and other hydro-climatic variables (Bowman et al., 2018; Brient, 2020; Hall et al., 2019; Petrova et al., 2024; Shiogama et al., 2022), the potential constraints on global mean changes in terrestrial water storage remain largely unexplored, probably due to the complex and heterogeneous components of land water storage, their differing representations across numerical models, and the historical lack of continental-scale TWS observations. The launch of the Gravity Recovery and Climate Experiment (GRACE) and GRACE-Follow On (GRACE-FO) missions has, however, provided spaceborne measurements of global freshwater changes (Tapley et al., 2019; Velicogna et al., 2020). Here, we combine a proposed EC with historical GRACE observations and the latest datasets from the Inter-Sectoral Impact Model Intercomparison Project phase 3b (ISIMIP3b; https://protocol.isimip.org/#/ISIMIP3a). By incorporating as many models and ensemble members as possible to account for varying model representations, we apply the EC framework to constrain projections of future TWS changes.

2 Data and methods

2.1 Observations

Global TWS anomaly products (excluding Greenland and Antarctica) for the period 2004–2023 were derived from GRACE and GRACE-FO satellites. The period 2004–2023 is shorter than the conventionally defined climatological baseline, but it provides a reasonable basis for estimating long-term trends. In addition, given that our primary focus is on long-term changes in TWS rather than interannual variability, this study period is considered adequate, though longer records would further

strengthen the analysis. Because research has demonstrated that GRACE data processing in terms of mass concentration (mascon) solutions results in higher correlations with in situ data compared to spherical harmonic solutions (Watkins et al., 2015), we utilized all three available GRACE mascon solution datasets (i.e., JPL RL06.3M v04, CSR RL0603M, and GSFC mascon RL06 v1.0), produced by the Jet Propulsion Laboratory (Watkins et al., 2015), the Center for Space Research (Save et al., 2016), and the Goddard Space Flight Center (Luthcke et al., 2013), respectively, to estimate observational uncertainty. Missing months in the GRACE time series were filled using linear interpolation. This approach is appropriate because the analysis focuses on long-term changes and linear trends in TWS, rather than on resolving seasonal variability. Furthermore, no scaling factors were applied, as previous studies have indicated that standard scale factors may not improve the accuracy of TWS trend estimates (Humphrey et al., 2023). The mean of three mascon products was used for EC calibration, with the associated standard deviation representing observational uncertainty.

2.2 Global models and climate forcing

Five LSMs and GHMs (see details in Table S1) from the ISIMIP3b were employed to assess historical and future relationships in TWS anomalies (see details below). These projections were based on climate forcing from the Coupled Model Intercomparison Project Phase 6 (CMIP6) (Eyring et al., 2016). Climate forcing data were sourced from five general circulation models (GCMs)—GFDL-ESM4, IPSL-CM6A-LR, MPI-ESM1-2-HR, MRI-ESM2-0, and UKESM1-0-LL—under three scenarios: historical climate (HIST, 1850–2014), a low greenhouse gas (GHG) emissions scenario (shared socioeconomic pathway (SSP)1-2.6), and a high GHG emissions scenario (SSP3-7.0). These scenarios were selected to maximize the number of available models and ensemble members included in the analysis, thereby enhancing both model diversity and within-model ensemble completeness. The historical climatology (2004–2023) was constructed by combining the end of the HIST run with the beginning of the SSP1-2.6 simulation, following the approach adopted in previous TWS studies using the ISIMIP datasets (e.g., Pokhrel et al., 2021). SSP2-4.5, a "middle of the road" scenario, would be the most appropriate for extending the historical period due to its alignment with historical socioeconomic trajectories (Fricko et al., 2017). Nevertheless, only a limited number of models provide the SSP2 forcing in the ISIMIP datasets. Consequently, trade-offs were required to maintain sufficient model diversity. While the use of SSP1-2.6 to extend the historical climatology may introduce discontinuities in forcing conditions, this approach enables a more comprehensive ensemble analysis. Each individual model realization was treated independently during the EC processing. To further ensure robustness, the proposed ECs were repeated using an ensemble of simulations from eight ISIMIP2b models (Table S2). These simulations were based on climate forcing from four CMIP5 GCMs under three scenarios: historical climate (HIST, 1861–2005), the Representative Concentration Pathway (RCP)2.6 scenario, and the RCP6.0 scenario. All outputs were provided at a monthly temporal resolution and on a 0.5° × 0.5° global grid. Monthly data were then regridded to a common 1° × 1° global grid using bilinear interpolation. The ensemble members (i.e., outputs from each GHM or LSM driven by different climate forcings) were compared with GRACE data. To ensure consistency, we computed TWS anomalies at each grid point relative to a baseline period of 2004–2009. This aligns all datasets to a common reference, conducting direct comparison of observed and simulated TWS anomalies as deviations from the same

climatological mean. Consequently, historical and future climatologies can be derived from these TWS anomalies (see details in the next subsection). While the baseline period is relatively short (affecting only the calculation of anomalies), our analysis focuses on long-term TWS changes, which are expected to be minimally influenced by the length of the chosen baseline.

2.3 Emergent constraint approach and calibration

To implement the EC framework, we begin by identifying statistically significant relationships between projected future annual global mean changes in TWS (e.g., at mid-century or by the end of the century), the predictand **y**, and historical climatologies (2004–2023) of annual global mean TWS anomalies, the predictor **x**, across a suite of global models. To ensure the robustness of our EC results, we additionally used linear trends of historical TWS anomalies as alternative predictors. The long-term trend in historical annual TWS anomalies at each grid point was estimated using ordinary least squares regression. A two-sided linear least-squares regression model was then employed to depict the EC relationship between **x** and **y** at the global scale. A total of 25 and 31 individual realizations were included in the EC relationships using ISIMIP3b and ISIMIP2b datasets, respectively (Fig. S1). The nonparametric Spearman rank-order correlation coefficient was also calculated, with corresponding p-values used to assess statistical significance. Residual plots were employed to identify outliers that consistently deviate from the regression-defined EC relationship across scenarios and to evaluate the linearity assumption of the regression.

Once **x** was estimated from historical GRACE observations, the resulting regression was applied to narrow the inter-model spread of **y**. In particular, by substituting **x** with the mean of historical observations, we derived the mean of EC-corrected changes from the regression model. No bias correction was applied before the regression. We further calibrated the projected future changes by applying the EC relationship to each grid cell, resulting in the spatial distribution of EC-corrected projections along with the corresponding biases. It should be noted that if observational uncertainties and regression residuals are large, these errors may propagate through the EC calibration, resulting in less well-constrained projections.

To estimate the uncertainty of the calibrated future changes, a stationary bootstrap method (Brient, 2020; Politis & White, 2004) was applied. Bootstrapping was performed 500 times, generating 500 regression line samples. This approach robustly quantifies the sampling uncertainty without requiring assumptions about the underlying probability distributions. Following the method of Brient and Schneider (Brient & Schneider, 2016), confidence intervals for the calibrated future changes were estimated by projecting the observed value x_o from each GRACE dataset (using all three mascon solutions) onto the generated 500 regression line samples.

Spearman's rank correlation coefficients were calculated for TWS-related, including precipitation, evapotranspiration, and total runoff, to elucidate the potential physical mechanisms underlying the EC relationship. Partial Spearman's correlation

analysis was also employed between TWS and the other variables, controlling for precipitation effects. These variables were derived from simulations under the historical and SSP3-7.0 scenarios. To understand the spatial distribution of differences in future changes, we categorized the six driest and six wettest ensemble members (Table S1) using the 25th and 75th percentiles of the ensemble distribution of the historical climatology of annual global mean TWS anomalies, after excluding outliers identified from the residual analysis (Fig. S9). To assess the sensitivity of this classification, we also applied the 35th and 65th percentile thresholds and found that the results are robust to the choice of threshold under the SSP3-7.0 scenario (Fig. S11). For simplicity, we refer to models with historically low and high simulated TWS anomalies as "dry" and "wet" models, respectively. The observed discrepancies are likely due to differences in how models represent storage compartments and account for human influences. The driest models produced present-day global mean TWS anomalies that closely align with values derived from mascon solutions, while the wettest models appeared to overestimate the historical climatology (Fig. S1).

3 Results

3.1 Proposed emergent constraint derived from historical climatology

Our approach follows established EC protocols by relating an observable historical metric (that is, area-weighted global mean of annual climatologies of TWS anomalies) to a targeted future change (global mean of end-of-century annual TWS change) across an ensemble of models. This strategy is consistent with prevailing EC studies (e.g., Cai et al., 2025; P. Dai et al., 2024; Kim et al., 2023; Petrova et al., 2024), which often relate present-day climatological states (or trends when they offer stronger physical connections) to future changes rather than link historical and future trends directly. To align GRACE observations with model outputs, we first compute TWS anomalies at each grid point relative to a baseline period of 2004–2009. Anchoring all datasets to this common reference ensures that both historical and future TWS anomalies (whether from observations or models) are directly comparable as deviations from the same climatological mean. Consequently, because both historical and future TWS anomalies share this baseline, their difference naturally represents changes in TWS (hereafter referred to as "TWS change") rather than changes in anomalies. In fact, our approach to calculating TWS change has shown results comparable to the conventional method (cf. Fig. 2c,d in this article with Fig. 1b,d in Pokhrel et al., 2021). Finally, we derive climatologies of TWS anomalies by averaging monthly anomalies for each calendar year and then computing the multi-year mean over the period of interest.

We analyze historical and future relationships in TWS using multi-model hydrological simulations, including 25 ensemble members from the ISIMIP3b (Warszawski et al., 2014) and 31 ensemble members from the ISIMIP2b (Frieler et al., 2017), considered separately (Methods and Tables S1 and S2). Future projections are based on climate forcing from the CMIP6 (Eyring et al., 2016) and CMIP5 (Taylor et al., 2012), respectively. To maximize the inclusion of available models and compare with previous studies (e.g., Pokhrel et al., 2021), we select a low-emissions scenario (SSP1-2.6/RCP2.6) and a high-emissions

scenario (SSP3-7.0/RCP6.0). Significant positive correlations ($R > 0.99$ for both ISIMIP2b and ISIMIP3b models) are found between historical and late century annual area-weighted global mean TWS anomalies across realizations, irrespective of the emissions scenario (Fig. S1). Statistical significance is adjusted for the number of realizations (25 for ISIMIP3b and 31 for ISIMIP2b). These findings indicate that models simulating higher TWS anomalies in the historical climate tend to predict higher TWS anomalies in the future, in agreement with the "wet-gets-wetter" atmospheric response to warming identified globally (Held & Soden, 2006) and over land (Petrova et al., 2024) in the CMIP models. Various hydrological datasets also corroborate this "wet-gets-wetter" signal over water-sufficient lands (Greve et al., 2014; Kumar et al., 2015; Markonis et al., 2019). GHMs and LSMs are developed based on distinct numerical frameworks and parameterizations, leading to structural diversity in how they represent key storage components. This structural diversity such as lack of surface water and groundwater storage compartments and human intervention in most LSMs; see Tables S1 and S2) is responsible for inter-model differences (Scanlon et al., 2018), which may lead to the "wet-gets-wetter" pattern. For instance, under the SSP1-2.6 scenario, MIROC-INTEG-LAND simulates relatively low climatologies of TWS anomalies in both historical and future periods, whereas JULES-W2 produces much higher values under the same conditions (Fig. S1). Our methodology does not assume identical storage physics across models; instead, it leverages the multi-model spread to empirically constrain biases against GRACE observations.

Notably, however, the strength and spatial extent of the "wet-gets-wetter" response are sensitive to differences in study periods, datasets, and the variables used to characterize hydrological conditions. For example, even when historical and future periods are held fixed, analyses based on satellite observations, GHM/LSM outputs, and GCM simulations report that the "wet-gets-wetter" response spans ~11% to ~41% of global land area, as quantified using a nondimensional TWS drought severity index (Xiong et al., 2022).

As shown in Fig. 1, the magnitude of future TWS increases is positively correlated with the historical climatology of TWS anomalies. The Spearman's rank correlation coefficients between the historical climatology and mid-century TWS changes are 0.79 and 0.70 ($p < 0.01$) for the ISIMIP2b and ISIMIP3b ensembles, respectively, under the low-emissions scenario. Under the high-emissions scenario, statistically significant correlations of 0.61 and 0.71 ($p < 0.01$) are observed for the ISIMIP2b and ISIMIP3b ensembles, respectively. For late century TWS changes, significant correlations with the historical climatology persist across both ISIMIP2b and ISIMIP3b ensembles. Residual plots are used to assess the linearity assumption underlying the EC framework (Fig. S9). Overall, the linear EC assumption is broadly reasonable across the four SSP panels. However, late-century projections show residuals that are not fully captured by a single linear EC regression, indicating increased nonlinearity and warranting caution in the interpretation of late-century results. It is worth noting that several modeling factors

including initial conditions, structural differences (especially the diversity of storage compartments), human intervention, and calibration can contribute to the large inter-model spread and thus the resulting discrepancies with GRACE observations.

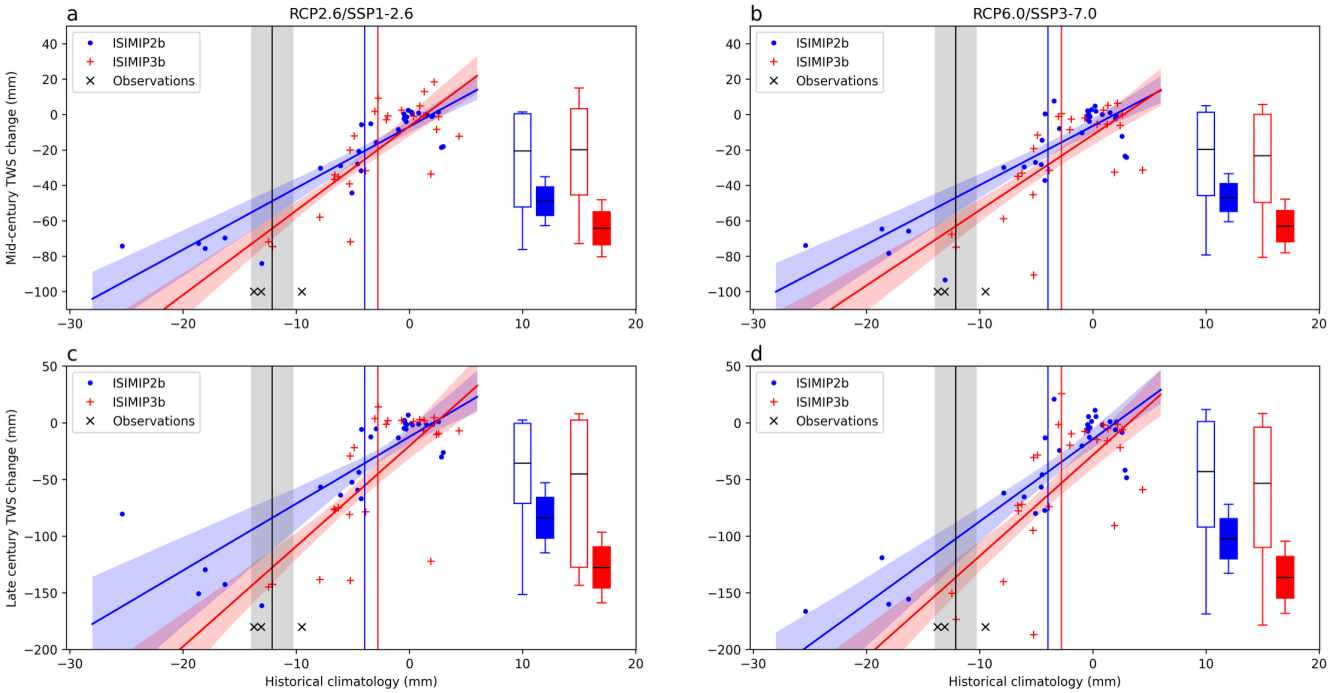

Figure 1: Inter-model relationships between historical (2004–2023) climatologies and future changes at mid-century (2040–2059; upper panels) and late century (2080–2099; lower panels) from ISIMIP2b (blue) and ISIMIP3b (red) models under the RCP2.6/SSP1-2.6 (a,c) and RCP6.0/SSP3-7.0 (b,d) scenarios. Dots and crosses represent global (excluding Greenland and Antarctica) averages from ensemble members. Blue and red lines represent linear regression fits, with 90% confidence intervals estimated through bootstrapping. Blue and red vertical lines mark the ensemble mean. Black vertical lines indicate the average of GRACE observations (black cross), and grey shading represents the standard deviation. Box plots indicate the mean (black line), 66% (box), and 90% (whisker) confidence intervals of future TWS changes before (empty box) and after (filled box) applying observational constraints.

The robust relationship provides a basis for constraining future TWS changes. After applying the EC calibration (Methods), mid-century global mean TWS changes are reduced by 44 mm and 40 mm compared to the raw projections from the ISIMIP3b ensembles under the low- and high-end forcing scenarios, respectively (Fig. 1, upper panels). For late century projections, EC-corrected changes could be ~83 mm lower than the raw estimates from the ISIMIP3b ensembles irrespective of the forcing scenario (Fig. 1, lower panels), highlighting potentially lower global freshwater availability than initially indicated by the ISIMIP3b models. Furthermore, the EC correction constrains the uncertainty in late century TWS changes by 63% for the SSP1-2.6 scenario and 69% for the SSP3-7.0 scenario, based on the 5–95% ranges of future projections, with 90% confidence

intervals estimated via bootstrapping. Specifically, the upper bound (95th percentile) is reduced from 2 mm to −96 mm under the low-end forcing scenario and from 8 mm to −105 mm under the high-end forcing scenario, indicating an initial overestimation of global freshwater availability in the raw ISIMIP3b ensemble projections. Global models are sophisticated process representations, making it challenging to isolate causes of their differences with GRACE (Haddeland et al., 2011). However, by leveraging the EC methodology and inter-model spread in water storage partitioning constrained by GRACE observations, we produce more tightly constrained estimates of future TWS changes than those obtained from conventional approaches, while recognizing the inherent uncertainties associated with observational constraints. Note that the reduction in ensemble uncertainty shown in our study is unexpectedly greater than that reported in previous EC-related research (e.g., Brient, 2020; Brient & Schneider, 2016; Petrova et al., 2024), highlighting a potential limitation of the EC approach when applied to freshwater availability. This discrepancy arises because GRACE products used for EC correction are all derived from a single source—the GRACE satellites—resulting in an underestimated uncertainty for TWS observations compared to other multi-source observational variables, such as air temperature and precipitation. EC-corrected changes based on the ISIMIP2b models show similar results to those from the ISIMIP3b models, albeit with higher ensemble averages, which can be attributed to the shallower slopes observed in their linear regression relationships. As ISIMIP2b and ISIMIP3b provide different sets of GHMs/LSMs and ensemble members, driven by different GCMs, these differences may also contribute to discrepancies in the EC-corrected results between ISIMIP2b and ISIMIP3b. Although here we illustrate results using three mascon-based GRACE datasets, extending the analysis to include four spherical-harmonic solutions yields EC-corrected projections that remain robust but produces slightly higher mean TWS changes at the end of this century (Fig. S7). We also analyze the past-future EC relationship using linear trends of historical TWS anomalies as potential predictors (Fig. S2, lower panels). The corresponding results closely align with those obtained using historical climatology of TWS anomalies as predictors, showing identical ensemble averages and consistent 5–95% ranges. Therefore, subsequent analyses are based on the past-future EC relationship with historical climatology of TWS anomalies as predictors, due to the current models' inability to accurately simulate the large regional TWS trends observed by GRACE (Scanlon et al., 2018). This limitation may reduce the reliability of the spatial patterns in future TWS change projections presented below.

3.2 Spatial patterns in future TWS change projections

By the end of the 21st century, TWS is projected to decrease considerably across several regions under the SSP1-2.6 scenario, including the southern United States, Mexico, northwestern South America, both coasts of the strait of Gibraltar, the majority of Central, West, and South Asia, as well as North China (Fig. 2a). Under the SSP3-7.0 scenario, although the spatial pattern of TWS changes is similar, the signal becomes much stronger, especially in the low latitudes (Fig. 2b). The future changes projected by the ISIMIP2b ensembles are consistent with those from the ISIMIP3b ensembles across most land areas, with pattern correlations of 0.44 and 0.55 under the low- and high-end forcing scenarios, respectively. Although this study uses an unweighted ensemble mean and a slightly different period to represent the late century, the spatial distributions closely align

with those reported by Pokhrel et al. (2021). The main difference lies in the latitudinal mean, which in our analysis reveals a much steeper decline in TWS over the northern midlatitudes, irrespective of the emissions scenario (Fig. 2c,d).

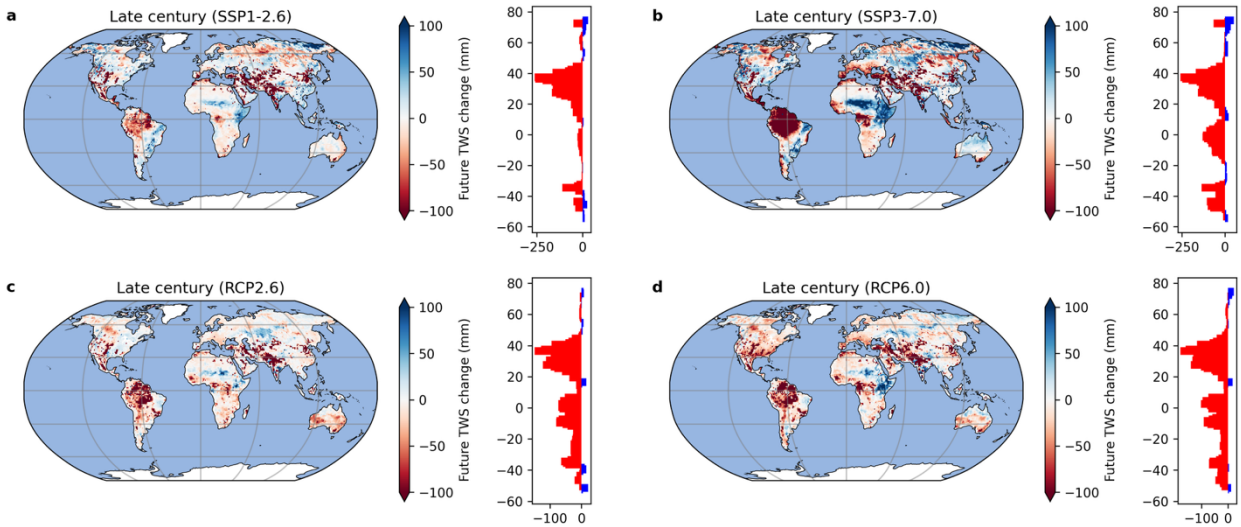

Figure 2: Late century (2080–2099) multi-model ensemble averages of TWS changes under the SSP1-2.6/RCP2.6 (a,c) and SSP3-7.0/RCP6.0 (b,d) scenarios, shown relative to historical (2004–2023) climatologies. The accompanying histograms on the right indicate zonally averaged TWS changes for each scenario.

To provide accurate regional-scale projections of future TWS changes, we obtained EC-corrected maps (Fig. 3a,c) by calibrating the ISIMIP3b outputs against observed historical climatologies at each grid cell (Methods). The proposed EC correlations are robust ($R > 0$ and $p < 0.05$) across ~26% of global land areas (excluding Greenland and Antarctica) under both scenarios, based on the ISIMIP3b ensembles. In contrast to the raw ISIMIP3b outputs (Fig. 2a,b), the EC-corrected maps show similar patterns, such as pronounced TWS declines in the northern midlatitudes, though with greater magnitudes under both the low- and high-emissions scenarios, as highlighted in the histograms of the latitudinal mean (Fig. 3a,c). The similarity in spatial patterns between the raw model outputs and the EC-corrected projections, along with the bias maps (Fig. 3b,d), corroborates our global analysis: observation-constrained TWS changes (ensemble mean) shift away from zero after the EC correction (Fig. 1). Under both emissions scenarios, the ISIMIP3b outputs notably overestimate TWS across mid- and high latitudes in the Northern Hemisphere, including regions such as the Northwest Territories in Canada, the Southwestern United States, much of the Middle East, the Danube River Basin, Siberia, and North China. Particularly, the zonal-mean TWS is overestimated by over 300 mm near 30° N and 60° N under the SSP1-2.6 scenario, with an even larger overestimation exceeding 400 mm at the same latitudes under the high-emissions scenario. This overestimation is consistent with previously reported underestimations of decreasing TWS trends in models over Northern Hemisphere river basins compared to satellite

observations (Scanlon et al., 2018), which may consequently lead to overestimated projections of future freshwater availability in these regions. Conversely, considerable underestimations of TWS occur in the Amazon and the Murray Basin in southeastern Australia, aligning with models' tendency to underestimate increasing TWS trends in humid regions or nonirrigated basins (Scanlon et al., 2018).

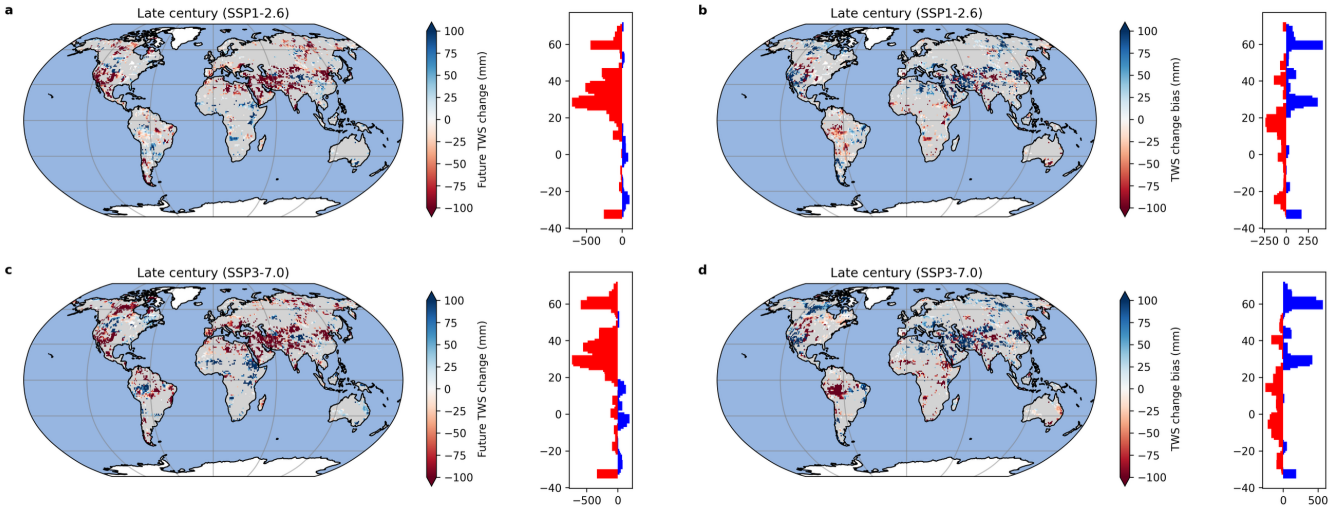

Figure 3: a,c, Late century (2080–2099) EC-corrected TWS changes under the SSP1-2.6 and SSP3-7.0 scenarios, shown relative to historical (2004–2023) climatologies. b,d, Biases in projected late century TWS changes (raw model outputs minus EC-corrected values). Only grid cells with statistically significant positive EC correlations ($R > 0$ and $p < 0.05$) are shown; non-significant regions are shaded in grey. The histograms on the right represent zonally averaged values, with data shown only for latitudes between 70°N and 35°S due to sparse coverage outside this range.

3.3 Underlying physical processes of the emergent relationship

It is crucial to elucidate the physical mechanisms linking historical and future variability in the EC relationship (Caldwell et al., 2014; Hall et al., 2019; Schlund et al., 2020). Under the SSP3-7.0 scenario, significant positive inter-model correlations ($p < 0.05$) are found between local precipitation and TWS changes over most regions globally (Fig. 4a). Similarly, positive correlations are evident between local precipitation and other TWS-related variables, such as evapotranspiration and total runoff (Fig. 4b,c). These results are consistent with established physical understandings, affirming that local precipitation changes strongly correlate with TWS changes. More importantly, they suggest that models projecting higher precipitation changes tend to predict larger TWS increases. This finding highlights the transfer of the "wet-gets-wetter" atmospheric response to warming into future hydrological projections through precipitation forcing. Actually, climate models predicting higher warming trends often anticipate greater precipitation increases as a result of thermodynamics (Emori & Brown, 2005;

Shiogama et al., 2022). At local scales, atmospheric warming-induced increases in water vapor enhance precipitation in wet regions and reduce it in dry regions (Chou et al., 2013; Held & Soden, 2006).

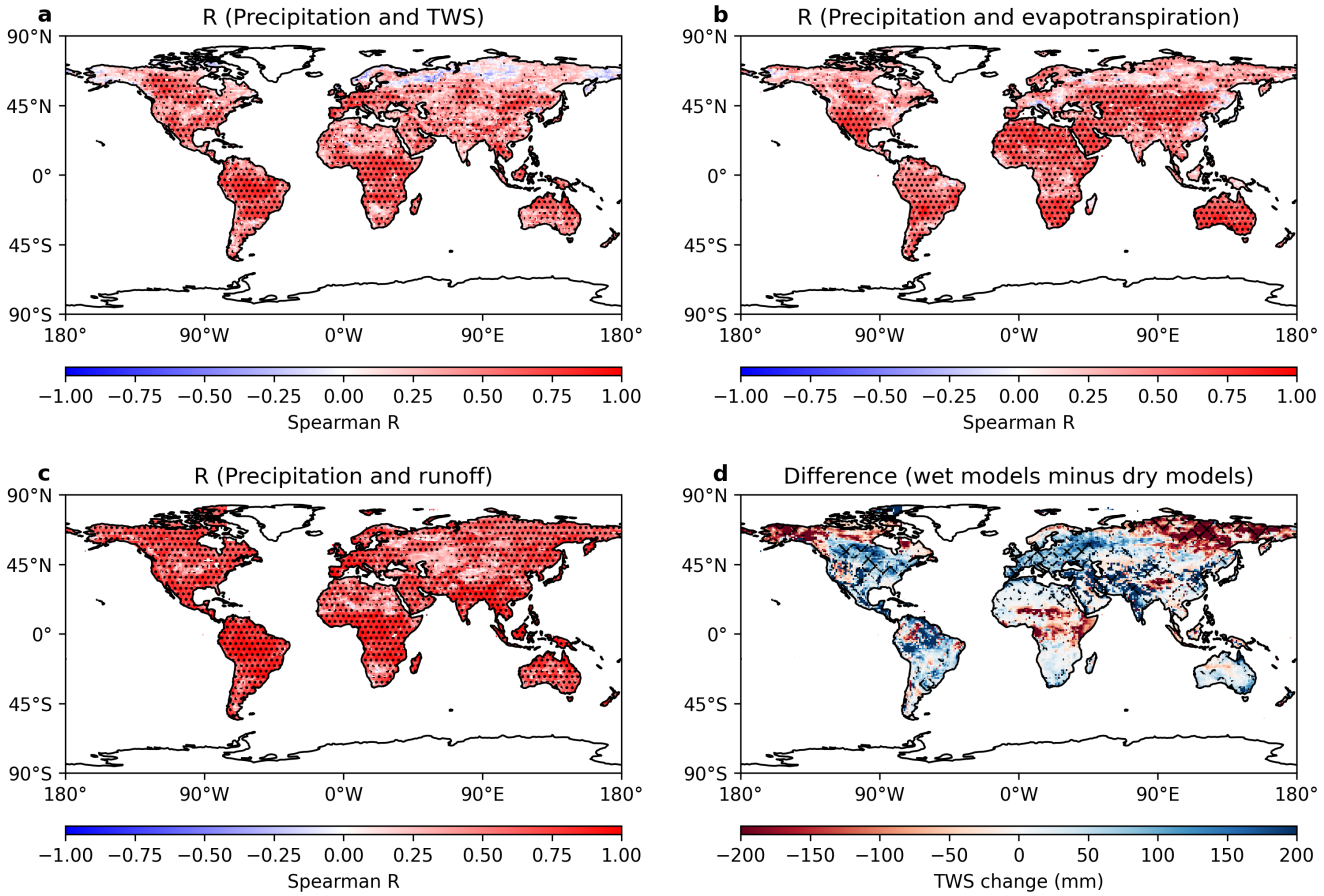

Figure 4: Inter-model Spearman's rank correlations between late century (2080–2099) precipitation changes and changes in TWS (a), evapotranspiration (b), and total runoff (c) under the SSP3-7.0 scenario. Black stippling marks regions of statistical significance ($p < 0.05$). d, Differences in late-century TWS changes (the wettest minus the driest models). Black hatches indicate statistically significant differences at the 5% level, as determined by Welch's *t*-test. A permutation test with 100 random permutations was conducted to estimate the *p*-values.

Positive inter-model correlations between changes in TWS and changes in other variables, such as evapotranspiration and runoff, are found across most global land areas (Fig. S8a,b), likely reflecting the dominant influence of precipitation. After controlling for precipitation, however, partial Spearman's correlations reveal a noticeably different pattern, in particular for evapotranspiration–TWS relationships (Fig. S8c), which become negative over much of the land areas. This result indicates

that enhanced evapotranspiration accelerates the depletion of future TWS, a process consistently captured by GHMs and LSMs. Conversely, the partial correlations between runoff and TWS closely resemble the original correlations even after controlling for precipitation (Fig. S8d). The widespread positive partial correlations underscore the role of high wetness and water storage in regulating runoff generation, including groundwater-driven baseflow and saturation-excess runoff (Dunne & Black, 1971; Li et al., 2025; Penna et al., 2011; Spence et al., 2010), whereby regions with greater water storage tend to produce more runoff because their buffering capacity is limited.

To further explore inter-model variations, we examine differences in TWS-related variables between the wettest and driest ensemble members based on their historical climatologies of TWS anomalies (Fig. 1, horizontal axis values; Methods). Significant differences in future TWS changes emerge, particularly in northern midlatitudes such as North America and Europe, where the driest ensemble members predict substantially more severe TWS reductions compared to the wettest ones (Fig. 4d). Similar patterns of inter-model correlations and subdued precipitation changes are also found under the SSP1-2.6 scenario (Fig. S5).

Under the ISIMIP3b framework, the wet–dry classification aligns primarily with hydrological model structure rather than the driving climate forcing: MIROC-INTEG-LAND and WaterGAP2-2e consistently emerge as dry models, while the remaining models are classified as wet (Table S1). This structural dependence is illustrated by cases where the same climate forcing produces contrasting responses across GHMs/LSMs. For instance, MPI-ESM1-2-HR yields a wet response when driving CWatM and JULES-W2, but a dry response when coupled with MIROC-INTEG-LAND and WaterGAP2-2e. One possible explanation is that MIROC-INTEG-LAND, as a land surface model, represents subsurface hydrology with more limited deep groundwater storage compared to dedicated hydrological models (Yokohata et al., 2020). WaterGAP, while including multiple storage compartments (e.g., soil, groundwater, lakes, and wetlands), employs a relatively simplified vertical structure, lacking explicit representation of deep groundwater layering, and therefore exhibits less storage complexity than fully physically-based hydrological models (Müller Schmied et al., 2021, 2024). It is worth noting that the classification of MIROC-INTEG-LAND and WaterGAP2-2e as dry models does not imply poor performance in their historical simulations. In fact, the driest models produce present-day global mean TWS anomalies that closely match those derived from GRACE mascon solutions, whereas the wettest models tend to overestimate the historical climatology (Fig. S1).

4 Discussion

While the proposed EC relationship in Fig. 1 appears robust at the global scale, showing a high and statistically significant correlation, substantial uncertainty remains at regional scales. In particular, significant EC correlations are found over only

~26% of global land areas (Fig. 3). Regions where the proposed EC relationship does not hold include large portions of high-latitude regions such as northwestern Canada and Siberia, as well as central Asia, the Middle East, central Africa, the Amazon basin, and central Australia. These regions largely overlap with data-sparse or climatically extreme areas. There are two possible explanations for this pattern. First, as discussed earlier, the "wet-gets-wetter" responses are sensitive to the choice of study periods and regions. As a result, the signal may be attenuated in the regions with intricate hydrological processes, particularly where nonlinear land surface feedbacks, such as groundwater depletion or vegetation responses, play a role, often in remote regions subject to climatic extremes. Second, GHMs and LSMs lack sufficient observational data to adequately constrain simulations in remote regions. This limitation may weaken or obscure the "wet-gets-wetter" signal in GHMs and LSMs, even when such a signal is present in the driving GCMs.

The EC relationship does not hold for several individual realizations, which appear as outliers in the residual plots (Fig. S9). These outliers consistently deviate from the regression-defined EC relationship across scenarios (Figs. S9 and S10). Notably, realizations driven by GFDL-ESM4 and IPSL-CM6A-LR (numbered 1, 4, 6, 7, and 8 in Fig. S9) show pronounced deviations. Such systematic departures suggest that inconsistencies in these realizations may originate from their driving GCMs, likely reflecting model-specific biases in representing the "wet-gets-wetter" pattern. Furthermore, the reliability of proposed ECs could be compromised due to the lack of independence among climate models (Brient, 2020; Caldwell et al., 2014). Climate models are frequently derived from one another (Knutti et al., 2013). This challenge becomes more pronounced in future hydrological projections. LSMs and GHMs often share structural similarities in simulating water storage compartments and human water use sectors (Telteu et al., 2021). For example, all ISIMIP2b models included in this study employ the same DDM30 drainage direction map to characterize surface water routing (Döll & Lehner, 2002). In addition, three models (CWatM, MATSIRO, and MPI-HM) use an identical formulation to calculate potential evapotranspiration; while four models (CWatM, JPJmL, PCR-GLOBWB, and WaterGAP2) rely on the same set of variables to calculate changes in canopy water storage (Telteu et al., 2021). These findings underscore the critical need to ensure diversity in global models as well as in their climate forcings, particularly in water-resources-focused projects like the ISIMIP. Model selection should prioritize inclusion of models with distinct formulations of key hydrological processes and encourage integration of independently developed models such as dynamic global vegetation models alongside GHMs and LSMs. Incorporating a broad range of GCM forcings to drive all GHMs and LSMs would allow a more robust evaluation of how sensitive projections are to external climate drivers.

Several limitations should be acknowledged in this study. First, the GRACE products used in the analysis originate from a single observational system and therefore are likely affected by the same systematic biases. Second, the applicability of the EC technique is inherently limited by the knowledge space represented by the ensemble of climate models (Hall et al., 2019). If key physical processes are oversimplified or absent in the models, the EC method cannot identify or constrain those

processes. Thus, the inter-model spread captured by the EC relationship may be unrealistic or unjustified due to the lack of a sound physical basis. In addition, the EC approach relies on statistical assumptions, including the stationarity of present-day relationships under future climate conditions, which may not hold and can lead to overconfident constraints (Brient, 2020; Hall et al., 2019).

In conclusion, our observation-constrained results highlight that warming-induced reductions in TWS translate into diminished freshwater availability on both global and regional scales. This is especially evident in mid- and high-latitude regions of the Northern Hemisphere, where low historical climatologies are prevalent. Our constrained estimates of global mean TWS change are roughly 83 mm lower than the raw ISIMIP3b projections, far exceeding the observed historical range of global mean TWS anomalies relative to the 2003–2020 mean (−15 to 16 mm) (Rodell et al., 2024). This large discrepancy reveals a significant overestimation of future water availability in both GHMs and LSMs. Consequently, current water resource planning may underestimate potential shortages, especially as uneven water gaps are projected to widen under warming scenarios (Rosa & Sangiorgio, 2025). Accordingly, there is an urgent need to reduce model uncertainties through robust observational constraints and enhance diversity among GHMs and LSMs, thereby improving the reliability of future hydrological projections for informed water resource management and climate adaptation.

Acknowledgments

YL acknowledges support from a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant.

Data availability

All ISIMIP output and climate forcing data are available at https://data.isimip.org. GRACE products can be obtained from https://grace.jpl.nasa.gov/data/get-data/jpl_global_mascons/, https://www2.csr.utexas.edu/grace/RL05_mascons.html, https://earth.gsfc.nasa.gov/geo/data/grace-mascons, https://podaac.jpl.nasa.gov/dataset/GRACEFO_L2_JPL_MONTHLY_0063, https://podaac.jpl.nasa.gov/dataset/GRACEFO_L2_GFZ_MONTHLY_0063, and https://gracefo.jpl.nasa.gov, and https://cost-g.org.

Code availability

Python and the Climate Data Operators (CDO) scripts were used to prepare model output, climate forcing, and GRACE data. The code can be accessed at Zenodo (https://doi.org/10.5281/zenodo.15838299).

Author contributions

FH performed the research, analyzed the data, and wrote the manuscript draft; FH, YL, and ZL reviewed and edited the manuscript.

Conflict of Interest Disclosure

The authors declare there are no conflicts of interest for this manuscript.

References

Bowman, K. W., Cressie, N., Qu, X., & Hall, A. (2018). A Hierarchical Statistical Framework for Emergent Constraints: Application to Snow-Albedo Feedback. *Geophysical Research Letters*, *45*(23), 13,050-13,059. https://doi.org/10.1029/2018GL080082

Brient, F. (2020). Reducing Uncertainties in Climate Projections with Emergent Constraints: Concepts, Examples and Prospects. *Advances in Atmospheric Sciences*, *37*(1), 1–15. https://doi.org/10.1007/s00376-019-9140-8

Brient, F., & Schneider, T. (2016). Constraints on Climate Sensitivity from Space-Based Measurements of Low-Cloud Reflection. *Journal of Climate*, *29*(16), 5821–5835. https://doi.org/10.1175/JCLI-D-15-0897.s1

Cai, Z., You, Q., Screen, J. A., Chen, H. W., Zhang, R., Zuo, Z., Chen, D., Cohen, J., Kang, S., & Zhang, R. (2025). Lessened projections of Arctic warming and wetting after correcting for model errors in global warming and sea ice cover. *Science Advances*, *11*(10), 6413. https://doi.org/10.1126/sciadv.adr6413

Caldwell, P. M., Bretherton, C. S., Zelinka, M. D., Klein, S. A., Santer, B. D., & Sanderson, B. M. (2014). Statistical significance of climate sensitivity predictors obtained by data mining. *Geophysical Research Letters*, *41*(5), 1803–1808. https://doi.org/10.1002/2014GL059205

Chou, C., Chiang, J. C. H., Lan, C.-W., Chung, C.-H., Liao, Y.-C., & Lee, C.-J. (2013). Increase in the range between wet and dry season precipitation. *Nature Geoscience*, *6*(4), 263–267. https://doi.org/10.1038/ngeo1744

Dai, A., Zhao, T., & Chen, J. (2018). Climate Change and Drought: a Precipitation and Evaporation Perspective. *Current Climate Change Reports*, *4*(3), 301–312. https://doi.org/10.1007/s40641-018-0101-6

Dai, P., Nie, J., Yu, Y., & Wu, R. (2024). Constraints on regional projections of mean and extreme precipitation under warming. *Proceedings of the National Academy of Sciences*, *121*(11). https://doi.org/10.1073/pnas.2312400121

Döll, P., & Lehner, B. (2002). Validation of a new global 30-min drainage direction map. *Journal of Hydrology*, *258*(1–4). https://doi.org/10.1016/S0022-1694(01)00565-0

Dunne, T., & Black, R. D. (1971). Runoff Processes during Snowmelt. *Water Resources Research*, *7*(5). https://doi.org/10.1029/WR007i005p01160

Emori, S., & Brown, S. J. (2005). Dynamic and thermodynamic changes in mean and extreme precipitation under changed climate. *Geophysical Research Letters*, *32*(17). https://doi.org/10.1029/2005GL023272

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. *Geoscientific Model Development*, *9*(5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Fricko, O., Havlik, P., Rogelj, J., Klimont, Z., Gusti, M., Johnson, N., Kolp, P., Strubegger, M., Valin, H., Amann, M., Ermolieva, T., Forsell, N., Herrero, M., Heyes, C., Kindermann, G., Krey, V., McCollum, D. L., Obersteiner, M., Pachauri, S., … Riahi, K. (2017). The marker quantification of the Shared Socioeconomic Pathway 2: A middle-of-the-road scenario for the 21st century. *Global Environmental Change*, *42*, 251–267. https://doi.org/10.1016/j.gloenvcha.2016.06.004

Frieler, K., Lange, S., Piontek, F., Reyer, C. P. O., Schewe, J., Warszawski, L., Zhao, F., Chini, L., Denvil, S., Emanuel, K., Geiger, T., Halladay, K., Hurtt, G., Mengel, M., Murakami, D., Ostberg, S., Popp, A., Riva, R., Stevanovic, M., … Yamagata, Y. (2017). Assessing the impacts of 1.5 °C global warming – simulation protocol of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP2b). *Geoscientific Model Development*, *10*(12), 4321–4345. https://doi.org/10.5194/gmd-10-4321-2017

Getirana, A., Kumar, S., Girotto, M., & Rodell, M. (2017). Rivers and Floodplains as Key Components of Global Terrestrial Water Storage Variability. *Geophysical Research Letters*, *44*(20), 10,359-10,368. https://doi.org/10.1002/2017GL074684

Greve, P., Orlowsky, B., Mueller, B., Sheffield, J., Reichstein, M., & Seneviratne, S. I. (2014). Global assessment of trends in wetting and drying over land. *Nature Geoscience*, *7*(10), 716–721. https://doi.org/10.1038/ngeo2247

Haddeland, I., Clark, D. B., Franssen, W., Ludwig, F., Voß, F., Arnell, N. W., Bertrand, N., Best, M., Folwell, S., Gerten, D., Gomes, S., Gosling, S. N., Hagemann, S., Hanasaki, N., Harding, R., Heinke, J., Kabat, P., Koirala, S., Oki, T., … Yeh, P. (2011). Multimodel Estimate of the Global Terrestrial Water Balance: Setup and First Results. *Journal of Hydrometeorology*, *12*(5), 869–884. https://doi.org/10.1175/2011JHM1324.1

Hall, A., Cox, P., Huntingford, C., & Klein, S. (2019). Progressing emergent constraints on future climate change. *Nature Climate Change*, *9*(4), 269–278. https://doi.org/10.1038/s41558-019-0436-6

Held, I. M., & Soden, B. J. (2006). Robust Responses of the Hydrological Cycle to Global Warming. *Journal of Climate*, *19*(21), 5686–5699. https://doi.org/10.1175/JCLI3990.1

Herrera-Estrada, J. E., Satoh, Y., & Sheffield, J. (2017). Spatiotemporal dynamics of global drought. *Geophysical Research Letters*, *44*(5), 2254–2263. https://doi.org/10.1002/2016GL071768

Humphrey, V., Rodell, M., & Eicker, A. (2023). Using Satellite-Based Terrestrial Water Storage Data: A Review. In *Surveys in Geophysics* (Vol. 44, Issue 5, pp. 1489–1517). Springer Science and Business Media B.V. https://doi.org/10.1007/s10712-022-09754-9

Kim, Y.-H., Min, S.-K., Gillett, N. P., Notz, D., & Malinina, E. (2023). Observationally-constrained projections of an ice-free Arctic even under a low emission scenario. *Nature Communications*, *14*(1), 3139. https://doi.org/10.1038/s41467-023-38511-8

Knutti, R., Masson, D., & Gettelman, A. (2013). Climate model genealogy: Generation CMIP5 and how we got there. *Geophysical Research Letters*, *40*(6), 1194–1199. https://doi.org/10.1002/grl.50256

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. *Geophysical Research Letters*, *42*(24). https://doi.org/10.1002/2015GL066858

Li, X., Long, D., Scanlon, B. R., & Slater, L. J. (2025). Retrievals and simulations of terrestrial water storage changes and runoff over the Tibetan Plateau: Challenges and opportunities. *Fundamental Research*. https://doi.org/10.1016/j.fmre.2025.11.012

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. *Journal of Glaciology*, *59*(216), 613–631. https://doi.org/10.3189/2013JoG12J147

Markonis, Y., Papalexiou, S. M., Martinkova, M., & Hanel, M. (2019). Assessment of Water Cycle Intensification Over Land using a Multisource Global Gridded Precipitation DataSet. *Journal of Geophysical Research: Atmospheres*, *124*(21), 11175–11187. https://doi.org/10.1029/2019JD030855

Penna, D., Tromp-Van Meerveld, H. J., Gobbi, A., Borga, M., & Dalla Fontana, G. (2011). The influence of soil moisture on threshold runoff generation processes in an alpine headwater catchment. *Hydrology and Earth System Sciences*, *15*(3), 689–702. https://doi.org/10.5194/hess-15-689-2011

Petrova, I. Y., Miralles, D. G., Brient, F., Donat, M. G., Min, S.-K., Kim, Y.-H., & Bador, M. (2024). Observation-constrained projections reveal longer-than-expected dry spells. *Nature*, *633*(8030), 594–600. https://doi.org/10.1038/s41586-024-07887-y

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., … Wada, Y. (2021). Global terrestrial water storage and drought severity under climate change. *Nature Climate Change*, *11*(3), 226–233. https://doi.org/10.1038/s41558-020-00972-w

Politis, D. N., & White, H. (2004). Automatic Block-Length Selection for the Dependent Bootstrap. *Econometric Reviews*, *23*(1), 53–70. https://doi.org/10.1081/ETC-120028836

Rodell, M., Barnoud, A., Robertson, F. R., Allan, R. P., Bellas-Manley, A., Bosilovich, M. G., Chambers, D., Landerer, F., Loomis, B., Nerem, R. S., O'Neill, M. M., Wiese, D., & Seneviratne, S. I. (2024). An Abrupt Decline in Global Terrestrial Water Storage and Its Relationship with Sea Level Change. *Surveys in Geophysics*, *45*(6), 1875–1902. https://doi.org/10.1007/s10712-024-09860-w

Rodell, M., & Famiglietti, J. S. (2001). An analysis of terrestrial water storage variations in Illinois with implications for the Gravity Recovery and Climate Experiment (GRACE). *Water Resources Research*, *37*(5), 1327–1339. https://doi.org/10.1029/2000WR900306

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. https://doi.org/10.1038/s41586-018-0123-1

Rosa, L., & Sangiorgio, M. (2025). Global water gaps under future warming levels. *Nature Communications*, *16*(1), 1192. https://doi.org/10.1038/s41467-025-56517-2

Save, H., Bettadpur, S., & Tapley, B. D. (2016). High-resolution CSR GRACE RL05 mascons. *Journal of Geophysical Research: Solid Earth*, *121*(10), 7547–7569. https://doi.org/10.1002/2016JB013007

Scanlon, B. R., Fakhreddine, S., Rateb, A., de Graaf, I., Famiglietti, J., Gleeson, T., Grafton, R. Q., Jobbagy, E., Kebede, S., Kolusu, S. R., Konikow, L. F., Long, D., Mekonnen, M., Schmied, H. M., Mukherjee, A., MacDonald, A., Reedy, R. C., Shamsudduha, M., Simmons, C. T., … Zheng, C. (2023). Global water resources and the role of groundwater in a resilient water future. *Nature Reviews Earth & Environment*, *4*(2), 87–101. https://doi.org/10.1038/s43017-022-00378-6

Scanlon, B. R., Zhang, Z., Save, H., Sun, A. Y., Müller Schmied, H., van Beek, L. P. H., Wiese, D. N., Wada, Y., Long, D., Reedy, R. C., Longuevergne, L., Döll, P., & Bierkens, M. F. P. (2018). Global models underestimate large decadal declining and rising water storage trends relative to GRACE satellite data. *Proceedings of the National Academy of Sciences*, *115*(6), E1080–E1089. https://doi.org/10.1073/pnas.1704665115

Schlund, M., Lauer, A., Gentine, P., Sherwood, S. C., & Eyring, V. (2020). Emergent constraints on equilibrium climate sensitivity in CMIP5: do they hold for CMIP6? *Earth System Dynamics*, *11*(4), 1233–1258. https://doi.org/10.5194/esd-11-1233-2020

Shiogama, H., Watanabe, M., Kim, H., & Hirota, N. (2022). Emergent constraints on future precipitation changes. *Nature*, *602*(7898), 612–616. https://doi.org/10.1038/s41586-021-04310-8

Spence, C., Guan, X. J., Phillips, R., Hedstrom, N., Granger, R., & Reid, B. (2010). Storage dynamics and streamflow in a catchment with a variable contributing area. *Hydrological Processes*, *24*(16), 2209–2221. https://doi.org/10.1002/hyp.7492

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., … Velicogna, I. (2019). Contributions of GRACE to understanding climate change. *Nature Climate Change*, *9*(5), 358–369. https://doi.org/10.1038/s41558-019-0456-2

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An Overview of CMIP5 and the Experiment Design. *Bulletin of the American Meteorological Society*, *93*(4), 485–498. https://doi.org/10.1175/BAMS-D-11-00094.1

Telteu, C.-E., Müller Schmied, H., Thiery, W., Leng, G., Burek, P., Liu, X., Boulange, J. E. S., Andersen, L. S., Grillakis, M., Gosling, S. N., Satoh, Y., Rakovec, O., Stacke, T., Chang, J., Wanders, N., Shah, H. L., Trautmann, T., Mao, G., Hanasaki, N., … Herz, F. (2021). Understanding each other's models: an introduction and a standard representation of

16 global water models to support intercomparison, improvement, and communication. *Geoscientific Model Development*, *14*(6), 3843–3878. https://doi.org/10.5194/gmd-14-3843-2021

Trenberth, K. E., Dai, A., Rasmussen, R. M., & Parsons, D. B. (2003). The Changing Character of Precipitation. *Bulletin of the American Meteorological Society*, *84*(9), 1205–1218. https://doi.org/10.1175/BAMS-84-9-1205

Velicogna, I., Mohajerani, Y., A, G., Landerer, F., Mouginot, J., Noel, B., Rignot, E., Sutterley, T., van den Broeke, M., van Wessem, M., & Wiese, D. (2020). Continuity of Ice Sheet Mass Loss in Greenland and Antarctica From the GRACE and GRACE Follow-On Missions. *Geophysical Research Letters*, *47*(8). https://doi.org/10.1029/2020GL087291

Vogel, M. M., Zscheischler, J., & Seneviratne, S. I. (2018). Varying soil moisture–atmosphere feedbacks explain divergent temperature extremes and precipitation projections in central Europe. *Earth System Dynamics*, *9*(3), 1107–1125. https://doi.org/10.5194/esd-9-1107-2018

Vörösmarty, C. J., Green, P., Salisbury, J., & Lammers, R. B. (2000). Global Water Resources: Vulnerability from Climate Change and Population Growth. *Science*, *289*(5477), 284–288. https://doi.org/10.1126/science.289.5477.284

Warszawski, L., Frieler, K., Huber, V., Piontek, F., Serdeczny, O., & Schewe, J. (2014). The Inter-Sectoral Impact Model Intercomparison Project (ISI–MIP): Project framework. *Proceedings of the National Academy of Sciences*, *111*(9), 3228–3232. https://doi.org/10.1073/pnas.1312330110

Watkins, M. M., Wiese, D. N., Yuan, D., Boening, C., & Landerer, F. W. (2015). Improved methods for observing Earth's time variable mass distribution with GRACE using spherical cap mascons. *Journal of Geophysical Research: Solid Earth*, *120*(4), 2648–2671. https://doi.org/10.1002/2014JB011547

Wu, N. C., Bovo, R. P., Enriquez-Urzelai, U., Clusella-Trullas, S., Kearney, M. R., Navas, C. A., & Kong, J. D. (2024). Global exposure risk of frogs to increasing environmental dryness. *Nature Climate Change*, *14*(12), 1314–1322. https://doi.org/10.1038/s41558-024-02167-z

Xiong, J., Guo, S., Abhishek, Chen, J., & Yin, J. (2022). Global evaluation of the 'dry gets drier, and wet gets wetter' paradigm from a terrestrial water storage change perspective. *Hydrology and Earth System Sciences*, *26*(24), 6457–6476. https://doi.org/10.5194/hess-26-6457-2022