A probabilistic approach to characterizing drought using satellite gravimetry

Peyman Saemian[1], Mohammad Javad Tourian[1], Omid Elmi[1], Nico Sneeuw[1], and Amir AghaKouchak[2]

[1]University of Stuttgart
[2]University of California, Irvine

April 16, 2024

Abstract

In the recent past, the Gravity Recovery and Climate Experiment (GRACE) satellite mission and its successor GRACE Follow-On (GRACE-FO), have become invaluable tools for characterizing drought through measurements of Total Water Storage Anomaly (TWSA). However, the existing approaches have often overlooked the uncertainties in TWSA that stem from GRACE orbit configuration, background models, and intrinsic data errors. Here we introduce a fresh view on this problem which incorporates the uncertainties in the data: the Probabilistic Storage-based Drought Index (PSDI). Our method leverages Monte Carlo simulations to yield realistic realizations for the stochastic process of the TWSA time series. These realizations depict a range of plausible drought scenarios that later on are used to characterize drought. This approach provides probability for each drought category instead of selecting a single final category at each epoch. We have compared PSDI with the deterministic approach (SDI) over major global basins. Our results show that the deterministic approach often leans towards an overestimation of storage-based drought severity. Furthermore, we scrutinize the performance of PSDI across diverse hydrologic events, spanning continents from the United States to Europe, the Middle East, Southern Africa, South America, and Australia. In each case, PSDI emerges as a reliable indicator for characterizing drought conditions, providing a more comprehensive perspective than traditional deterministic indices. In contrast to the common deterministic view, our probabilistic approach provides a more realistic characterization of the TWS drought, making it more suited for adaptive strategies and realistic risk management.

A probabilistic approach to characterizing drought using satellite gravimetry

Peyman Saemian[1], Mohammad J. Tourian[1], Omid Elmi[1]
Nico Sneeuw[1], Amir AghaKouchak[2,3]

[1]Institute of Geodesy, University of Stuttgart, Stuttgart, Germany
[2]Department of Civil and Environmental Engineering, University of California, Irvine, CA, USA.
[3]United Nations University Institute for Water, Environment and Health, Hamilton, Ontario, Canada

Key Points:

- A probabilistic framework is introduced to characterize drought using GRACE and GRACE Follow-On observations.
- Our study highlights a tendency of deterministic approaches to consistently overestimate storage-based drought severity.
- The probabilistic approach captures global droughts while delivering more realistic results suited for risk management.

Corresponding author: Peyman Saemian, `peyman.saemian@gis.uni-stuttgart.de`

Abstract

In the recent past, the Gravity Recovery and Climate Experiment (GRACE) satellite mission and its successor GRACE Follow-On (GRACE-FO), have become invaluable tools for characterizing drought through measurements of Total Water Storage Anomaly (TWSA). However, the existing approaches have often overlooked the uncertainties in TWSA that stem from GRACE orbit configuration, background models, and intrinsic data errors. Here we introduce a fresh view on this problem which incorporates the uncertainties in the data: the Probabilistic Storage-based Drought Index (PSDI). Our method leverages Monte Carlo simulations to yield realistic realizations for the stochastic process of the TWSA time series. These realizations depict a range of plausible drought scenarios that later on are used to characterize drought. This approach provides probability for each drought category instead of selecting a single final category at each epoch. We have compared PSDI with the deterministic approach (SDI) over major global basins. Our results show that the deterministic approach often leans towards an overestimation of storage-based drought severity. Furthermore, we scrutinize the performance of PSDI across diverse hydrologic events, spanning continents from the United States to Europe, the Middle East, Southern Africa, South America, and Australia. In each case, PSDI emerges as a reliable indicator for characterizing drought conditions, providing a more comprehensive perspective than traditional deterministic indices. In contrast to the common deterministic view, our probabilistic approach provides a more realistic characterization of the TWS drought, making it more suited for adaptive strategies and realistic risk management.

Plain Language Summary

Total Water Storage (TWS) is defined as the sum of water stored as surface water (e.g., lakes and rivers), groundwater, soil moisture, snow, ice, and vegetation biomass. Since its launch in 2002, the Gravity Recovery and Climate Experiment (GRACE) satellite mission has provided unique TWS change measurements with manifold applications in hydrology,

including characterizing drought events. Scientists have been using satellites like GRACE and its successor, GRACE-FO, to understand drought by measuring the Total Water Storage Anomaly (TWSA). However, previous methods didn't consider uncertainties from satellite orbits, models, and data errors. This study offers a novel probabilistic approach for characterizing drought, Probabilistic Storage-based Drought Index (PSDI), which acknowledges the uncertainties in the GRACE TWS change. We use simulations to create different drought scenarios, offering probabilities for each category instead of one fixed category. Comparing PSDI to traditional methods, we found that traditional methods tend to overestimate drought severity. We tested PSDI across different regions, and it consistently proved to be a reliable way to understand drought conditions, offering a more comprehensive perspective. Our probabilistic approach offers a more realistic view of TWS drought, making it suitable for adaptive strategies and risk management.

1 Introduction

The modern reality of human settlement is the consequence of many historical events, but perhaps none influenced human settlements as much as droughts and famine. DNA analysis indicates that a series of extreme droughts that occurred 75-135 thousand years ago may have been the reason for the first human migration out of Africa (Scholz et al., 2007). Following several consequential droughts over the past century (e.g., the 1921 drought in Europe, the 1930s Dust Bowl drought in the US, 1928-1930 drought in China, 1980s drought and famine in Africa, 2000s Millennium drought in Australia), increasingly more effort has focused on understanding, monitoring and predicting droughts and their impacts (Mishra & Singh, 2010; Heim Jr, 2002; AghaKouchak et al., 2015; Svoboda et al., 2002; Wilhite et al., 2007; Kreibich et al., 2022; AghaKouchak et al., 2021).

Compared to other hazards witnessed over the past four decades, drought impacts are often felt by a much larger number of people worldwide (Wilhite, 2000; FAO, 2021; AghaKouchak et al., 2021). Numerous nations have grappled with significant economic losses resulting from drought events. Notably, according to the NOAA's National Centers for Environmental Information (NCEI) report, the United States has experienced 26 significant droughts in the past century, amounting to a staggering economic loss of at least $ 249 billion, equivalent to nearly $ 10 billion per occurrence. In Europe, the southern and western regions, in particular, face an annual drought-related expenditure estimated at up to € 9 billion, which could surge to over € 65 billion if climate action is not taken (Naumann et al., 2021). Aside from the financial burdens, climate change, and unsustainable water management practices have amplified the frequency and severity of drought occurrences worldwide over the past two decades. This trend is projected to escalate further in the future (see e.g., Hisdal et al., 2001; Coumou & Rahmstorf, 2012; Yu et al., 2014; Donat et al., 2016; Teuling, 2018; Li et al., 2021; C. Zhao et al., 2020).

The negative consequences of drought can be effectively alleviated through the implementation of risk management strategies rather than relying on crisis management (Wilhite, 2000; Zscheischler et al., 2018). Such a proactive response may be achieved by establishing reliable drought monitoring systems, including early warning systems and forecasting capabilities, operating at both national and local levels (Wilhite et al., 2007; AghaKouchak et al., 2023). These systems trigger a series of decisions aimed at helping communities navigate the challenges posed by drought events (Mishra & Singh, 2011; Sun et al., 2017). To enhance drought monitoring efforts and provide valuable guidance to decision-makers, numerous drought indices have been developed (Mishra & Singh, 2010). These indices condense the intricacies of drought into a single numerical value, effectively characterizing its onset, intensity, frequency, and duration (Zargar et al., 2011; Wilhite, 2000; Ahmadalipour et al., 2017). Such indices offer a comprehensive representation of drought by utilizing single or multiple climatic and hydrometeorological variables such as precipitation, streamflow, evapotranspiration, temperature, and snowpack (e.g., Svoboda et al., 2016; Hosseini-Moghari et al., 2020).

A comprehensive understanding of drought dynamics necessitates the observation of Total Water Storage (TWS) including snow, surface water, soil moisture, and groundwater storage (M. Zhao et al., 2017; M. J. Tourian et al., 2023). Traditionally, TWS monitoring has relied on costly and time-consuming site measurements, providing limited regional and local coverage. While hydrological and land surface models partially address this issue, estimating TWS in regions lacking in-situ runoff data for calibrating rainfall-runoff models still yields high uncertainties (Jiang et al., 2014; S. Yi et al., 2023). Since its launch in 2002, the Gravity Recovery And Climate Experiment (GRACE) satellite mission has revolutionized the remote measurement of TWS Anomalies (TWSA) at regional to continental scales (Tapley et al., 2004; M. J. Tourian et al., 2022). The GRACE mission came to an end on 12 October 2017, due to battery failure, after more than 15 years of Earth observation. However, its successor, GRACE Follow-On (GRACE-FO), has continued the GRACE legacy since its launch on 22 May 2018. GRACE(-FO) data have been extensively utilized for manifold applications, including monitoring ice sheets and glaciers (e.g., van den Broeke et al., 2009; Gardner et al., 2013; Shepherd et al., 2018), tracking anthropogenic groundwater depletion (e.g., Rodell et al., 2007, 2009; Famiglietti et al., 2011; Voss et al., 2013; Saemian et al., 2022), forecasting flood events (e.g., Reager & Famiglietti, 2009; Gouweleeuw et al., 2018), and quantifying and comprehending hydrological processes (e.g., Lorenz et al., 2014; Saemian et al., 2020; M. Tourian et al., 2018; Behling et al., 2022), to name but a few.

GRACE-derived estimates of TWS have been employed in developing indices aimed at assessing drought on a regional to global scale. For example, Yirdaw et al. (2008) developed the Total Storage Deficit Index (TSDI), utilizing the Palmer Drought Severity Index (PDSI; Palmer, 1965) and the Soil Moisture Deficit Index (SMDI; Narasimhan & Srinivasan, 2005), to characterize the Canadian Prairie droughts of 2002/2003. Another notable endeavor by Thomas et al. (2014) presented a comprehensive framework for drought characterization based on GRACE-derived TWSA over regions including the Amazon, Zambezi, Texas, and the southeastern United States. Additionally, H. Yi & Wen (2016) devised the GRACE-based Hydrological Drought Index (GHDI) to characterize drought in the continental United

States from 2003 to 2012, building upon the foundation of the PDSI concept. Among recent indicators we can name the Drought Severity Index (DSI) by M. Zhao et al. (2017), the Water Storage Deficit Index (WSDI) by Sinha et al. (2017), and a long-term standardized GRACE reconstructed TWSA index (SGRTI) by Zhong et al. (2023).

The indices mentioned above have the potential for monitoring and assessing the TWS drought at regional to global scales. Nevertheless, they adopt a deterministic approach that disregards the intrinsic uncertainties associated with characterizing drought using GRACE observations. These uncertainties are inherent in the GRACE data due to factors such as its orbit configuration, measurement concept, various post-processing approaches of GRACE data, and different options for de-aliasing products. Besides, the estimation of GRACE uncertainty varies among different GRACE level-2 products, in both magnitude and spatial pattern. Most of the centers offer an uncertainty measure (known as formal errors) in the form of spherical harmonic coefficients. Figure 1 shows the coefficient-wise ratio of average formal errors and empirical errors of the GRACE solution following the approach suggested by Kvas et al. (2019). The ideal ratio is set at one, with values below indicating an underestimation of empirical errors, whereas values exceeding one signify an overestimation. The three official centers (JPL, CSR, and GFZ) together with AIUB, HUST, and SWPU exhibit a similar pattern. The SWPU solution demonstrates a pronounced overestimation of formal errors, particularly for low d/o (under 30). ITSG and Tongji display comparable patterns, although Tongji tends to lean slightly towards a more pessimistic estimation of errors. In contrast, COST-G reflects more realistic formal errors in comparison to empirical errors but appears overly optimistic for lower d/o values. The distinctive pattern observed in the CNES product can be attributed to the regularization applied during the derivation of the gravity field from the Level-1 dataset. The disparities in formal error performance among GRACE Level-2 products underscore the inherent uncertainty in GRACE data and consequently, the necessity for a comprehensive approach that effectively considers GRACE uncertainty while characterizing drought.

To address this gap in conventional methods, we propose a probabilistic approach that considers all possible scenarios and associated impacts. Our approach leverages Monte Carlo simulations to obtain realistic realizations of TWSA compatible with GRACE-TWSA and its corresponding uncertainties. We then characterize drought based on TWSA using our proposed drought index, the Probabilistic Storage-based Drought Index (PSDI). This index indicates not only drought but also its corresponding occurrence probability. We compare our results with those from the conventional deterministic approaches over the major river basins. Moreover, the performance of PSDI in capturing the main hydrological drought extremes is examined within the GRACE era. PSDI facilitates more informed and proactive responses to water resource challenges and serves as a practical tool for decision-makers and water resource managers to assess and manage drought-related risks more realistically.

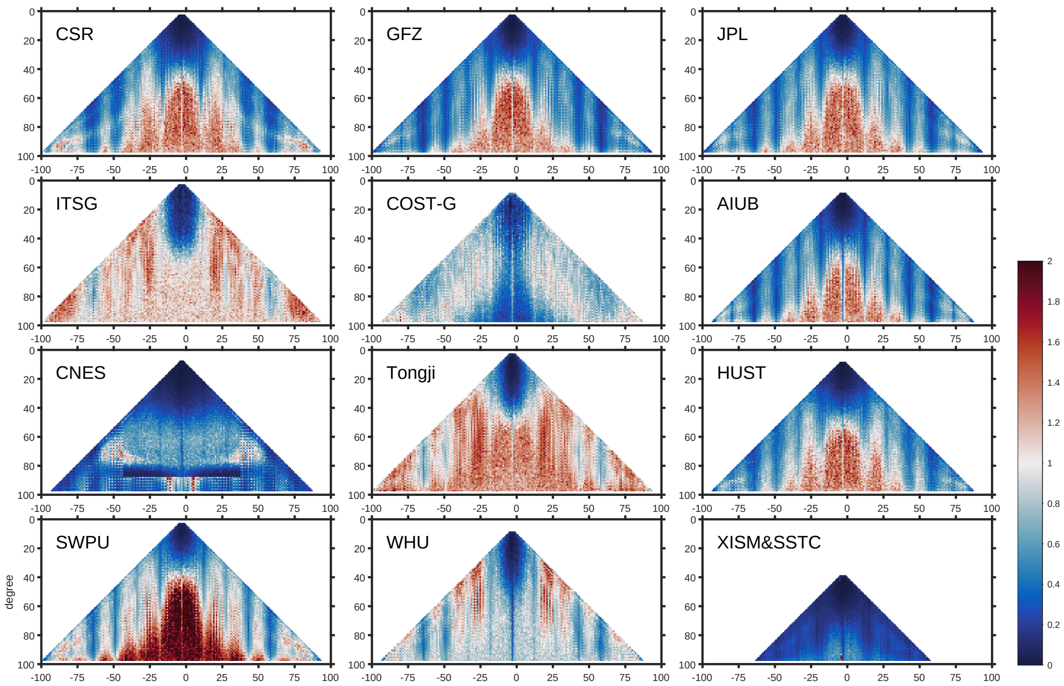

Figure 1. The coefficient-wise ratio of average formal errors and empirical errors of the GRACE solutions following the approach suggested by Kvas et al. (2019). For the average formal errors, the mean of the reported variance of the spherical harmonics coefficients is computed for monthly solutions from January 2005 to December 2010, which is assumed to hold a homogeneous data quality. To estimate the empirical errors, we compute the standard deviation of the coefficients after removing the mean, linear trend, and annual and semi-annual signals. The optimal value of the ratio is one, and values below one indicate an underestimation of the empirical errors, while values bigger than one show overestimation. We have only included solutions that provide formal errors.

2 Data and Method

2.1 GRACE data

The GRACE TWSA can be obtained from the two main approaches, namely Spherical Harmonics (SHs) and mass concentration blocks (mascons). In the former, one needs to apply post-processing steps including noise reduction and signal restoration while the latter is already the Level-3 product (gridded TWSA over the globe). These approaches are briefly described in section 1 of the supplementary file. In line with the common practice within the hydrology community, we have utilized the mascons solutions. The probabilistic approach for characterizing storage-based drought index, however, can readily be applied to any level-3 products that provide estimations for GRACE TWSA and its corresponding uncertainty. Among the mascon products, we have employed the one from Goddard Space Flight Center (GSFC), NASA. The GSFC mascon product has been widely used in the geodesy and Earth science communities to investigate a range of phenomena, including hydrology, glaciology, and solid Earth dynamics, and can be downloaded from `https://earth.gsfc.nasa.gov/geo/data/grace-mascons`. We used the latest version of the dataset available at the time of our analysis, which covers the period from August 2002 to November 2022. The dataset includes monthly gravity field solutions with a grid size of 0.5°.

2.2 Methodology

We propose a probabilistic framework to characterize storage-based drought. The framework is illustrated in Figure 2, Figure 3, and Figure 4 using TWSA over the Death Valley basin in the US as an example. To characterize drought, we must first define a reference, based on which a prolonged relative water deficiency is determined. It is common to consider the long-term monthly average, also known as the *climatology*, as the reference or *normal* condition in a region. Obtaining accurate climatology from short time series can be challenging. Calculating the climatology over at least 30 years, preferably 60 years, is standard practice, as this time frame allows us to average out the effects of short-term variability, resulting in a more robust estimate of the long-term average conditions (e.g., Hulme, 1992; Jones & Hulme, 1996; Svoboda et al., 2012). The GRACE and GRACE-FO missions, with their approximately 20-year duration, fall short of providing sufficient data for calculating long-term climatology. In this study, we utilized a combination of different models to estimate TWSA dating back to 1980. To this end, we incorporated a total of 13 state-of-the-art datasets including Global Hydrological Models (GHMs), Land Surface Models (LSMs), and atmospheric reanalysis models. To combine models, we employed the Multivariate Linear Regression (MLR) method. We compared the results of Multiple Linear Regression (MLR) in reconstructing TWSA with GRACE time series (see Supplementary section 2). The MLR exhibits a strong capability to capture the features and effectively reconstruct TWSA data as far back as 1980. It demonstrated superior performance compared to the ensemble mean of the models, as indicated by a substantial improvement in both the correlation coefficient (on average from 0.87 to 0.97) and the Kling-Gupta Efficiency (KGE) score (on average from 0.27 to 0.95) across major river basins. For more details about the datasets and long-term TWSA, please refer to Supplementary section 2. This extended time frame enables us to capture significant climate events and phenomena that influence long-term climate, such as the El Niño-Southern Oscillation (ENSO) and the North Atlantic Oscillation (NAO) (Sohn et al., 2013; Coelho & Goddard, 2009).

The climatology, along with its corresponding uncertainty (see Figure 2 (a)), is obtained by:

$$\overline{\mathrm{TWSA}}[t_m] = \frac{1}{N}\sum_{y=y_1}^{y_N} \mathrm{TWSA}[t_{y,m}] \tag{1}$$

$$\sigma^2_{\overline{\mathrm{TWSA}}[t_m]} = \frac{1}{N}\sqrt{\sum_{y=y_1}^{y_N} \sigma^2_{\mathrm{TWSA}[t_{y,m}]}} \tag{2}$$

where $\overline{\mathrm{TWSA}}[t_m]$ represents the TWSA climatology for month m, y denotes the year and can vary from y_1 to y_N, m corresponds to the month within a year, taking values $1, 2, 3, \ldots, 12$, and N is the number of years in the long-term dataset. Note that we deliberately retain the trend in the time series. We reason that the trend reflects long-term changes in climate, such as temperature increases or precipitation pattern alterations, and that it affects the frequency and severity of droughts (see Supplementary section 3 for more details). We then subtract the climatology from the GRACE TWSA time series to obtain TWS residual (S):

$$S[t_{y,m}] = \mathrm{TWSA}[t_{y,m}] - \overline{\mathrm{TWSA}}[t_m] \tag{3}$$

$$\sigma_{S[t_{y,m}]} = \sqrt{\sigma^2_{\mathrm{TWSA}[t_{y,m}]} + \sigma^2_{\overline{\mathrm{TWSA}}[t_m]}} \tag{4}$$

where negative values of S represent water storage deficits.

To reduce the effects of short-term fluctuations due to precipitation and other factors, we chose to use a 3-month moving average to smooth the TWS residual (see Figure 2 (b)):

$$S[t] = f * S[t]_{\mathrm{unsmoothed}} \tag{5}$$

where $*$ denotes the convolution operation, and f is the kernel [1/3, 1/3, 1/3] which is convolved with the $S[t]_{\mathrm{unsmoothed}}$ time series.

To address the inherent uncertainty, it becomes essential to employ a stochastic approach that incorporates S along with its associated uncertainty. At each epoch, we postulate a normal distribution with the mean being the obtained TWS residual $S[t]$ and $\sigma = \sigma_{S[t]}$. Sampling from this distribution in each time step allows us to create realizations of the S time series. Several methods exist for sampling from a Gaussian distribution, and one widely used technique is the Box-Muller transform (Box & Muller, 1958). This method guarantees generating a realization of S, which is independent epoch-wise with no artificial correlations. We then use Monte Carlo Simulation (Mooney, 1997; Metropolis & Ulam, 1949) to generate

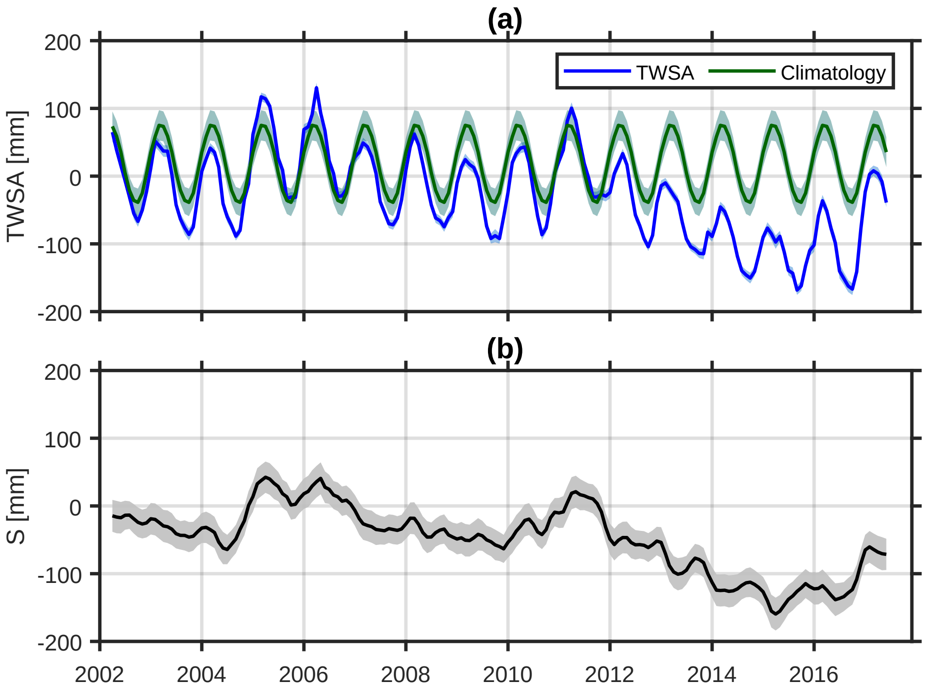

Figure 2. (a) Time series of the long-term TWSA from GRACE and the long-term climatology (1980–2012) from the hindcasted TWSA together with their uncertainties. At each epoch, we assume a Gaussian distribution for the uncertainties and the depicted uncertainty corresponds to the 1-σ level. Here the results are shown for the Death Valley basin in the US.

multiple realizations of S. Figure 3 shows 10 000 realization of TWS residual considering the 3-σ uncertainty. The density of the realizations is highest around the mean signal and decays following a Gaussian distribution. The colored lines overlaid on the time series depict the distribution of outcomes for three specific epochs: July 2004, May 2013, and April 2015.

To characterize drought within each time epoch, one common approach is to use the percentile rank method and the U.S. Drought Monitor (USDM) criteria. To this end, a set of five drought categories is defined (Table 1).

The quantile values for the S time series can be extracted as the inverse of the Cumulative Distribution Function (CDF), also known as the quantile function:

$$Q(p) = F^{-1}(p) \tag{6}$$

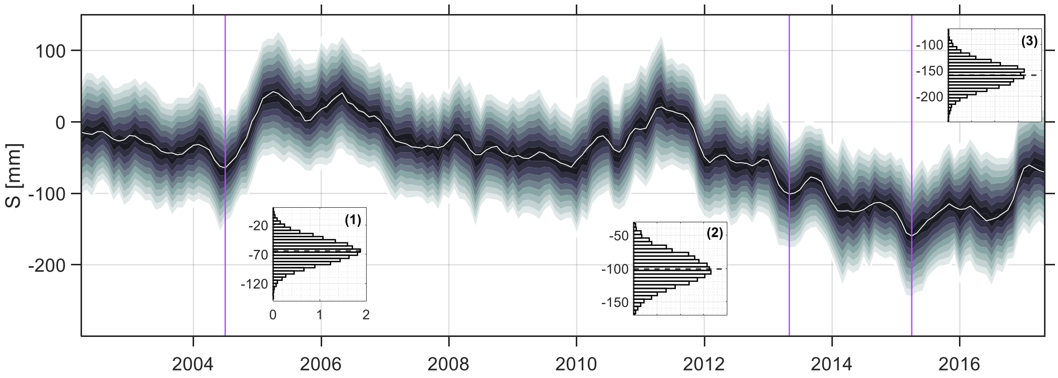

Figure 3. The TWS residual (S) together with its 10 000 realizations, calculated using Monte Carlo simulation. Here the results are shown for the Death Valley basin in the US. The distribution of realizations for three epochs, namely July 2004, May 2013, and April 2015 are marked with colored dots over the time series and are shown in sub-figures.

Table 1. Drought categories and corresponding percentile ranges as defined by the U.S. Drought Monitor (USDM).

Drought Category	Description	Percentile Range
D0	Abnormally dry	20–30%
D1	Moderate drought	10–20%
D2	Severe drought	5–10%
D3	Extreme drought	2–5%
D4	Exceptional drought	Less than 2%

where $Q(p)$ represents the quantile function and $F^{-1}(p)$ denotes the inverse CDF evaluated at probability p. In a conventional deterministic approach, the drought category for each epoch is determined based on its quantile value of S. For example in Figure 4 (a), the dark solid line represents the quantile function of the S time series. Using such a function one can characterize drought for case (1), case (2), and case (3) as D4, D1, and no drought, respectively.

Such an approach overlooks the uncertainty in TWS residual S. However, accounting for uncertainty would entail obtaining the quantile function for all realizations of S. These functions form a cloud of points rather than a single line as it has been illustrated in Figure 4 (a). The quantile functions are shown in grayscale representing the probability $\Pr(p, S)$ for a given percentile p and TWS residual S. Already at this stage, a glance at Figure 4 (a) reveals the complexity introduced by the uncertainty envelope, challenging the conventional approach to assigning a specific class to a particular measurement. It is noteworthy that the uncertainty envelope depicted in Figure 4 (a) exhibits a stationarity character, indicating general uncertainty in the data regardless of the specific time of measurements. This characteristic is reflective of the general uncertainty of S, emphasizing the broader statistical context rather than being tied to specific instances in time.

Now, let's delve into the characterization of drought for one of the measurements illustrated in Figure 3. In this context, alongside the consideration of the stationary uncertainty as reflected in the quantile envelope and represented by $\mathrm{Pr}(p, S)$, it becomes essential to account for the uncertainty associated with the measurement at that specific epoch. This is fundamentally crucial because two GRACE measurements with the same value of S may exhibit varying levels of uncertainty. Therefore, we incorporate the probability density function of the value S_t, denoted by $f(S_t)$, obtained from the mean and uncertainty of that epoch. $f(S_t)$ is shown for the three sample epochs on the top right panel of Figure 4. At each epoch, we multiply this probability density function with the entire distribution $\mathrm{Pr}(p, S)$, as illustrated in Figure 4 (c). Essentially, this multiplication results in a down-weighting of probabilities located in the tails of $f(S_t)$.

Once $\mathrm{Pr}(p, S)\, f(S_t)$ is achieved, to obtain PSDI at each epoch and for each drought category D_i, we can integrate the probabilities both in S and p domains and normalize it with the integral over the entire domain:

$$\mathrm{PSDI}(t, D_i) = \frac{\int_S \int_{D_i} \mathrm{Pr}(p, S) f(S_t)\, dp\, ds}{\int_S \int_0^1 \mathrm{Pr}(p, S) f(S_t)\, dp\, ds} \tag{7}$$

By performing this process for all drought categories and time epochs, we generate a comprehensive probabilistic representation of drought severity over time. For decision-making purposes, the highest-probability category can be judiciously chosen as the definitive drought classification for a particular month. The flowchart of the proposed probabilistic approach is shown in Figure 5.

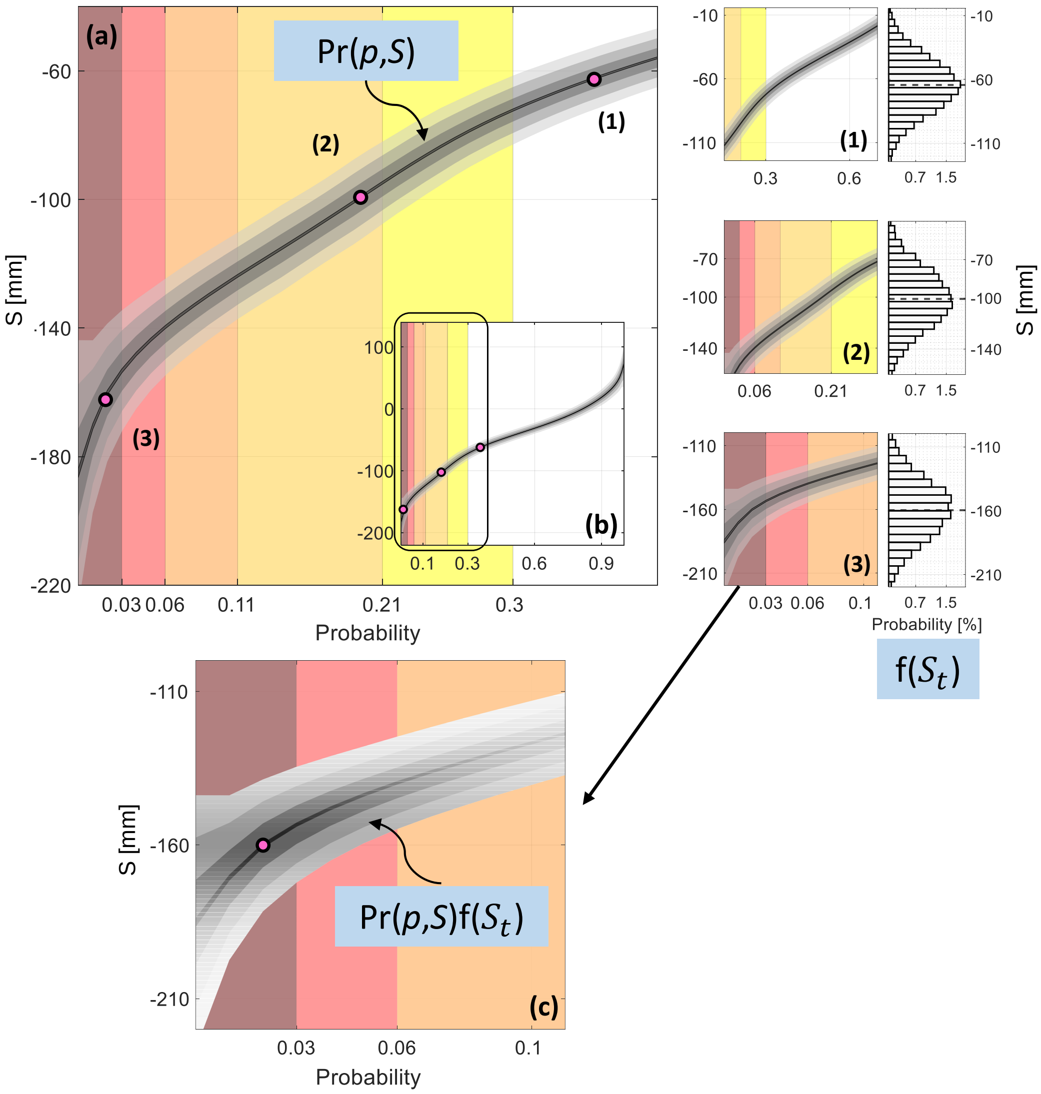

Figure 4. (a) The quantile functions (inverse of the Cumulative Distribution Function (CDF)) of the S realizations are depicted. The varying shades of gray signify the density of data points, with darker shades indicating higher density. The drought categories, ranging from D0 to D4, are delineated within their respective percentile ranges, each denoted by its corresponding color. Colored dots illustrate the positions of the three cases from Figure 3 on the quantile functions plot. These cases are further elaborated in a magnified view, accompanied by the corresponding probability distribution derived from S and its associated uncertainty. (b) Similarly, as in (a), this visualization portrays the quantile functions plot, but encompasses the complete range of quantile values. (c) The density of the counted points after integrating the probability distribution stemming from the S and its corresponding uncertainty. It's important to note that the presented results are centered on the Death Valley basin within the United States.

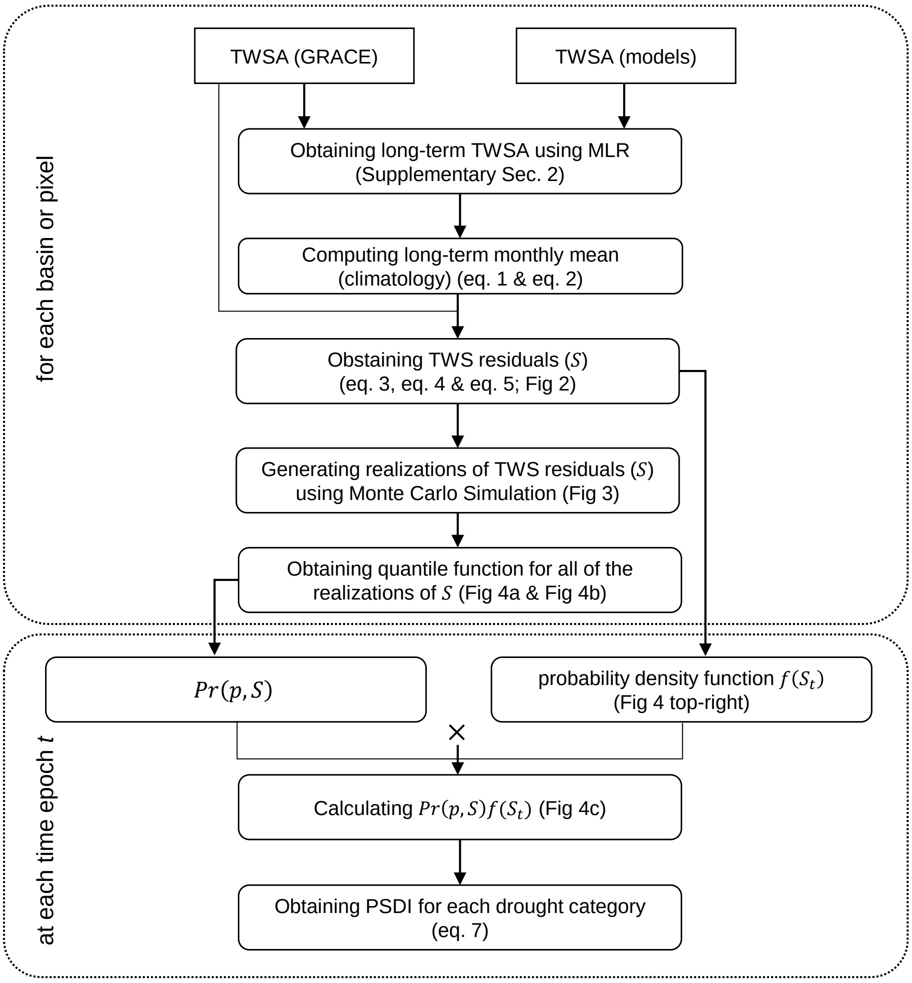

Figure 5. Flowchart of the proposed PSDI framework.

3 Results and Discussion

3.1 PSDI vs SDI

The PSDI approach offers a more nuanced understanding of drought conditions compared to the SDI approach. This is because PSDI captures the uncertainty associated with drought severity, while the SDI approach may oversimplify the classification of drought conditions. Although the SDI categorization is often the most probable category according to the PSDI, the neighboring categories may also have significant probabilities. This tendency becomes more pronounced as the intensity of the drought increases. This can be attributed to the lower slope of the CDF curve over more severe droughts and the wider range of quantile values.

To delve deeper into the analysis, we have quantified the disparities between drought categorizations as defined by SDI and $\mathrm{PSDI}_{\mathrm{max}}$—the category of drought with the highest probability in PSDI—across the world's major river basins, with the exclusion of Greenland and Antarctica. The findings, illustrated in Figure 6 (a), shed light on the prevalence of these discrepancies throughout the study period spanning from 2003 to 2016. The ratio exhibits a range of variations, hovering near zero for basins such as Lake Balkhash in south-eastern Kazakhstan or Po in Italy to a significant value of 30 % over Highland of Ethiopia and Somalia in Africa or Sao Francisco in Brazil. In general, the risk of mischaracterizing storage-based drought through the deterministic approach is notably high (exceeding 10 % in Figure 6 (a)) across Africa (excluding the northern region), Eastern Europe, Mongolia, Russia, and within the river basins of Nelson river, St. Lawrence, and Colorado (Argentina). In instances where discrepancies arise between SDI and $\mathrm{PSDI}_{\mathrm{max}}$, a predominant tendency is for SDI to overestimate the drought category. This is evident when comparing Figure 6 (b) and Figure 6 (c).

To investigate further, Figure 7 provides a visual comparison between two approaches for characterizing drought: probabilistic (PSDI) and deterministic (SDI), over several selected basins. The distribution of the basins is shown in the top panel of the Figure 7. For each basin, the drought categories, ranging from the status of no drought to exceptional drought (D4), are displayed in columns. The probability assigned to each category at every time step is depicted using gray scale. The deterministic perspective is illustrated with red boxes, allowing for a direct comparison of the two approaches.

The Danube and Ganges basins exhibited no disparity between SDI and $\mathrm{PSDI}_{\mathrm{max}}$ from 2015 to 2016. In contrast, the Mississippi basin displayed the most substantial mismatch between SDI and $\mathrm{PSDI}_{\mathrm{max}}$. It's noteworthy that these mismatches were confined to adjacent categories. Specifically, when considering mismatches spanning more than one category, only

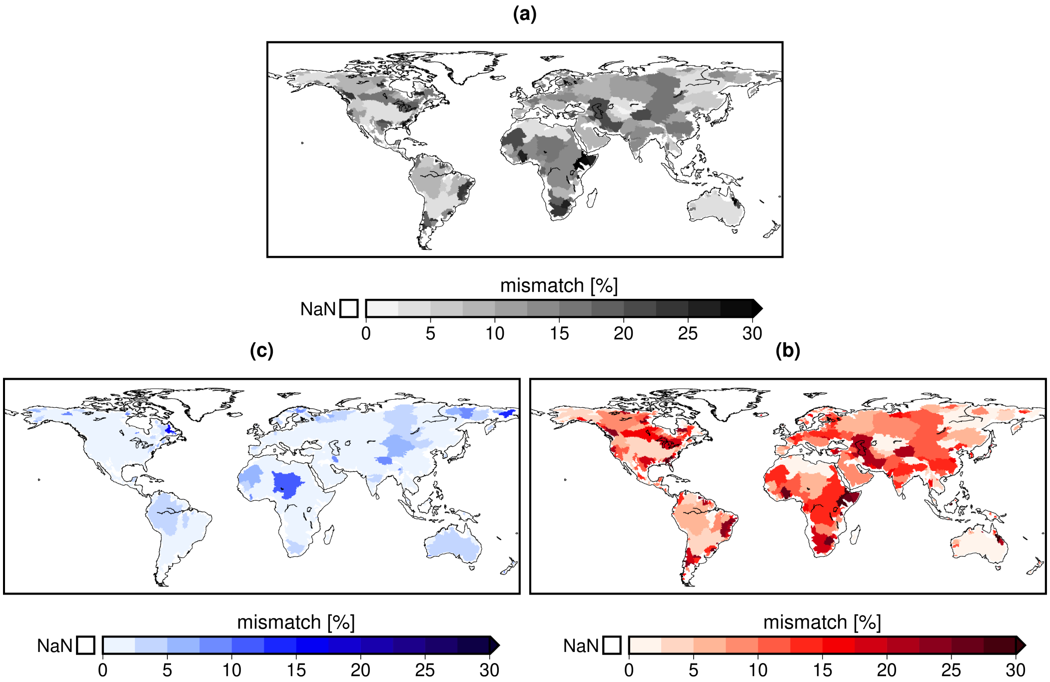

Figure 6. (a) Basin-wise distribution of the discrepancies between SDI and $PSDI_{max}$. The values represent the percentage of epochs where $PSDI_{max}$ differs from SDI by at least two drought categories. (b) the percentage of epochs with a discrepancy of more than one category higher in SDI compared to $PSDI_{max}$. (b) the percentage of epochs with a discrepancy of more than one category lower in SDI compared to $PSDI_{max}$. Greenland and Antarctica are excluded from the maps.

four basins had such occurrences: one month in Amazonas and Nile, two months in Niger, and five months in Murray Darling. Across all basins, when a discrepancy arose between SDI and $PSDI_{max}$, the SDI category consistently indicated a higher severity of drought.

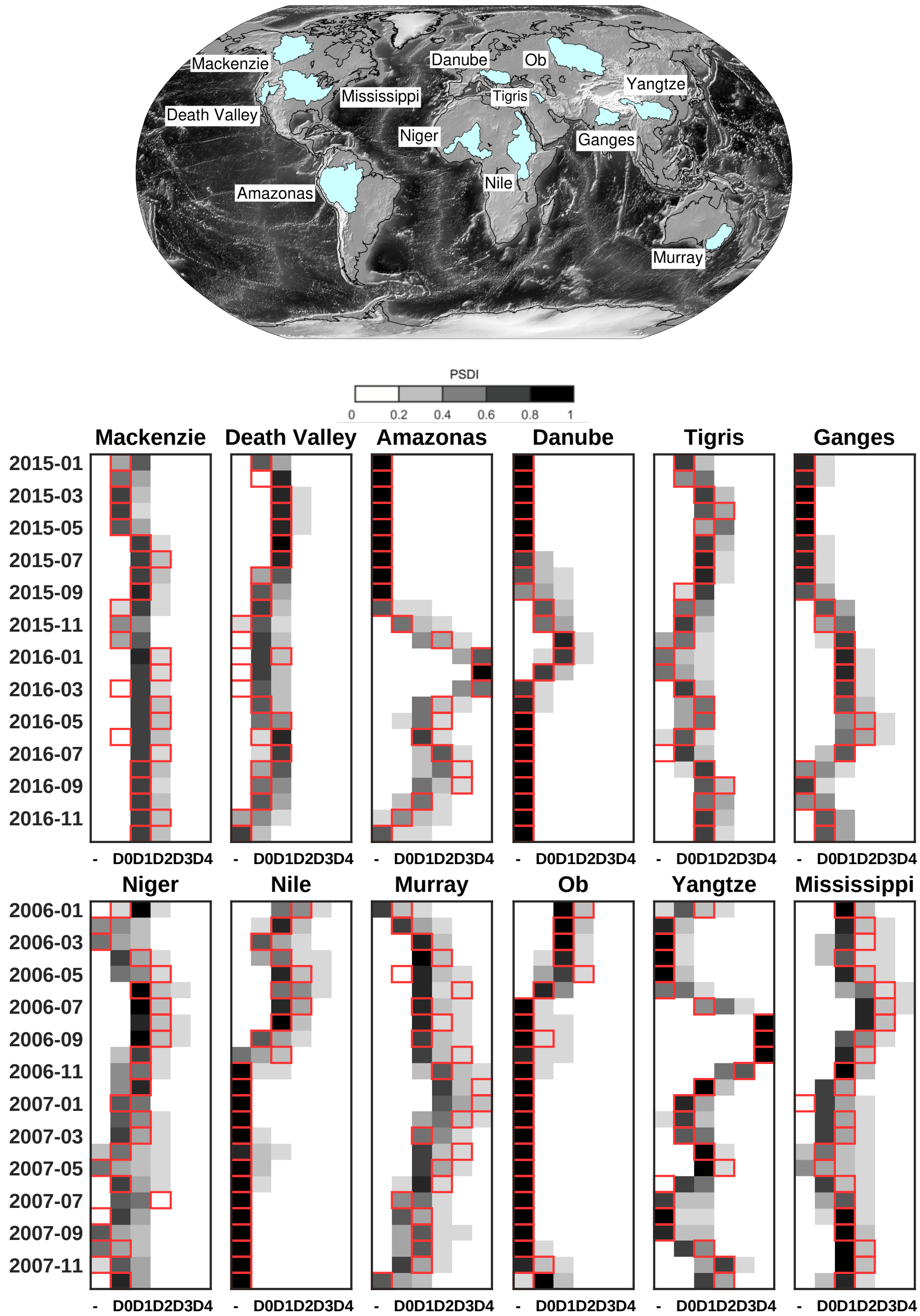

Figure 7. Top: The global distribution of the selected basins. Bottom: The SDI (red boxes) together with PSDI (gray scale probability range) for selected basins. The basins are shown in two groups considering the period with more frequency of drought, the first row between 2015–2016 and the second row between 2006–2007. The "-" represents "no drought" or "normal state" of the water storage.

We have investigated further the sensitivity of different categories of drought to incorporating uncertainties into drought characterization. Figure 8 (a) visualizes the percentage of epochs where $PSDI_{max}$ differs from SDI by at least two drought categories. The results suggest that such discrepancies can diverge significantly in the categorization of drought conditions, especially in the D1, D2, and D3 categories, especially in D2. We have also compared the ratio of the mismatch period with respect to different climate categories (Figure 8 (b)). For more detailed information about the categories and the method of classification, please see section 4 in the supplementary file. Although the mismatch range can vary from arid to humid climate, the average value of the mismatch is the same over different climate regions, with a slightly higher value for the Dry sub-humid regions (Dry sH).

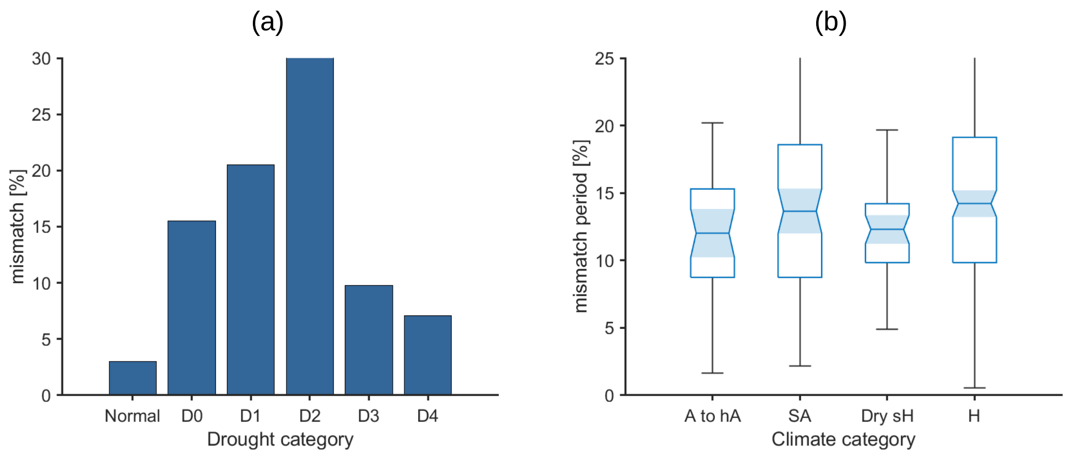

Figure 8. (a) A barplot illustrates the percentage of epochs where $PSDI_{max}$ diverges from SDI by at least two drought categories. (b) Boxplot of the mismatch between the $PSDI_{max}$ and SDI over different climate categories, namely, arid to hyper-arid (A to hA), semi-arid (SA), dry sub-humid (Dry sH), and humid (H). It is noteworthy that to count the number of months, we have considered those with more than one category difference between the $PSDI_{max}$ and SDI.

3.2 Performance of the PSDI during extreme hydrologic events

To assess the PSDI's reliability, we analyzed its performance during several well-documented extreme hydrologic events between 2002 and 2016. The drought events during 2012 included the moderate to exceptional drought over the United state (Boyer et al., 2013; Ault et al., 2013), southern Europe (Oikonomou et al., 2020; Spinoni et al., 2015). The drought affected many Middle East regions between 2007 and 2008 (Barlow et al., 2016). Southern Africa suffered from a severe to exceptional drought between 2005 and early 2006 (Nicholson, 2014), while central Argentina and Paraguay were affected by drought throughout 2009 (Guha-Sapir et al., 2016). Moreover, Australia experienced the worst drought recorded since European settlement in the 2000s, called the *Millennium drought*, with a peak in 2006 that affected many regions of the south to the east, including agricultural lands of the Murray-Darling basin (Van Dijk et al., 2013; Heberger, 2012). Figure 9 illustrates the performance of the PSDI over the events mentioned above. For each region, the category with the maximum probability and the estimated probability is shown for the selected date. Generally, the PSDI shows high performance in characterizing drought in the selected drought events (Figure 9). Comparing the SDI with $PSDI_{max}$ reveals that SDI categorizes higher drought intensities.

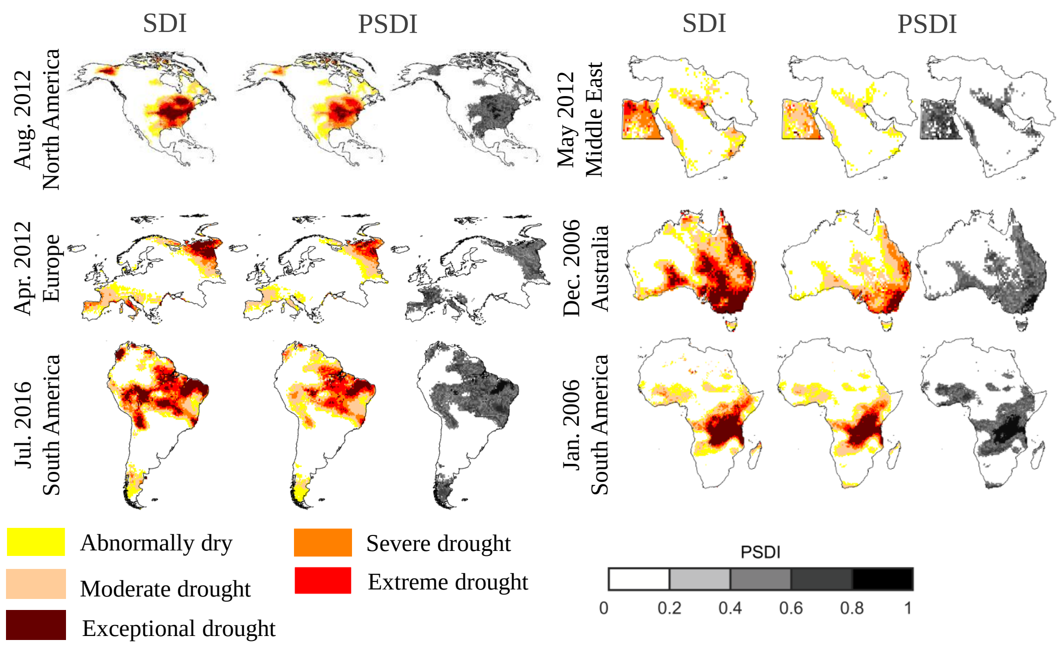

Figure 9. Comparing SDI with PSDI during some reported drought events.

4 Conclusions

For the first time, this study presents a probabilistic approach to characterizing TWS drought using time-variable gravity from satellite gravimetry. Our proposed framework acknowledges and addresses the inherent uncertainties associated with GRACE data. Our approach leverages Monte Carlo simulations to generate realistic realizations, capturing the stochastic nature of the TWSA time series. This ensemble reflects the diverse possible scenarios and their associated uncertainties, paving the way for a more insightful understanding of drought conditions. We have monitored the results of the proposed PSDI over major river basins and compared the result with SDI (deterministic approach). Our spatial analysis underscores the significance of adopting a probabilistic approach. It becomes evident that deterministic methodologies, in certain regions, tend to overestimate the severity of storage-based drought, potentially leading to misleading conclusions. While deterministic indices may tend to oversimplify drought categorization, PSDI accounts for uncertainty, thereby offering a more accurate representation of drought severity, particularly during extreme events.

Furthermore, our study assesses the performance of PSDI during well-documented extreme hydrologic events, spanning from the United States to Europe, the Middle East, Southern Africa, South America, and Australia. In each case, PSDI demonstrates its robustness in characterizing drought conditions. Comparing the SDI with $\mathrm{PSDI}_{\mathrm{max}}$ reveals that the drought can be categorized with more intensity using SDI with respect to the PSDI. We also address the uncertainties associated with different GRACE mascon products, emphasizing the importance of selecting the appropriate data source for reliable drought characterization. Variations in uncertainty estimates among different centers and processing methods highlight the need for caution when utilizing GRACE-derived data for drought analysis. We also shed light on the formal errors associated with GRACE data, highlighting the overestimation and underestimation tendencies of various solutions. This insight serves as a valuable reference for researchers and institutions relying on GRACE data for drought monitoring and assessment.

The findings of this study underscore the importance of a probabilistic approach in characterizing drought over various regions and during several drought events. The new approach provides a more realistic characterization of drought by accounting for the uncertainties in the GRACE(-FO) TWSA data in contrast to the common deterministic approach. By embracing uncertainty and providing a comprehensive ensemble of drought scenarios, PSDI advances the field of drought assessment, offering improved accuracy and insight for decision-makers and researchers alike. In an era marked by changing climate patterns and increasing water stress, our probabilistic approach represents a significant step toward more effective drought management and adaptation strategies.

Author Contribution Statement

Peyman Saemian and Mohammad J. Tourian developed the method, conducted the data analysis, and wrote the paper. Omid Elmi contributed to the analysis and assisted in producing graphics. Amir Aghakouchak and Nico Sneeuw supported the study with discussions on algorithm development. All authors commented on and reviewed the manuscript, and contributed to the final version.

Open Research

In this study, we employed a diverse set of datasets. The GRACE data, the GSFC mascon product, is available at `https://earth.gsfc.nasa.gov/geo/data/grace-mascons`. Nine global water resources datasets, including PCR-GLOBWB, SURFEX-TRIP, HBV-SIMREG, HTESSEL, JULES, LISFLOOD, ORCHIDEE, SWBM, and W3RA, were obtained from the eartH2Observe Water Cycle Integrator (`ftp://wci.earth2observe.eu`). CLM5 products are accessible via Earth System Grid (Oleson et al., 2019). The WaterGAP Global Hydrology Model (WaterGAP v2.2d) data is accessible at `https://doi.pangaea`

.de/10.1594/PANGAEA.918447. Additionally, the fifth generation ECMWF atmospheric reanalysis (ERA5) data can be downloaded from the Copernicus Climate Change Service (C3S) at ECMWF (https://cds.climate.copernicus.eu). For the long-term TWSA from the MLR approach, the data is available in mat format at DaRUS “Data for: A probabilistic approach to characterizing drought using satellite gravimetry”, https://doi.org/10.18419/darus-3832.

Acknowledgments

We extend our gratitude to the numerous organizations and data providers whose contributions were essential to the success of this research. The datasets utilized in this study were sourced from various centers and repositories, as comprehensively listed in Tables 1 and 2 of the Supplementary section 1. We acknowledge the invaluable support and data access made possible by these sources. P. Saemian is funded by the DFG funding in project RASLyBoCa within the Special Priority Program Regional Sea Level Change & Society ‘SeaLevel’, Germany (SPP1889)—Project number 313883668, Grant TO 1042/3-1. O. Elmi is funded by DFG (Deutsche Forschungsgemeinschaft) (project Number: 324641997) within the framework of the Research Unit 2630, GlobalCDA: understanding the global freshwater system by combining geodetic and remote sensing information with modeling using a calibration/data assimilation approach (https://globalcda.de).

References

AghaKouchak, A., Farahmand, A., Melton, F., Teixeira, J., Anderson, M., Wardlow, B., & Hain, C. (2015). Remote sensing of drought: progress, challenges and opportunities. *Reviews of Geophysics*, *53*(2), 452–480. (doi: 10.1002/2014RG000456)

AghaKouchak, A., Huning, L. S., Sadegh, M., Qin, Y., Markonis, Y., Vahedifard, F., ... others (2023). Toward impact-based monitoring of drought and its cascading hazards. *Nature Reviews Earth & Environment*, 1–14.

AghaKouchak, A., Mirchi, A., Madani, K., Di Baldassarre, G., Nazemi, A., Alborzi, A., ...

others (2021). *Anthropogenic Drought: Definition, Challenges, and Opportunities.* Wiley Online Library. doi: 10.1029/2019RG000683

Ahmadalipour, A., Moradkhani, H., Yan, H., & Zarekarizi, M. (2017). Remote sensing of drought: vegetation, soil moisture, and data assimilation. In *Remote sensing of hydrological extremes* (pp. 121–149). Springer.

Ault, T., Henebry, G., De Beurs, K., Schwartz, M., Betancourt, J. L., & Moore, D. (2013). The false spring of 2012, earliest in North American record. *Eos, Transactions American Geophysical Union*, *94*(20), 181–182. doi: 10.1002/2013EO200001

Barlow, M., Zaitchik, B., Paz, S., Black, E., Evans, J., & Hoell, A. (2016). A review of drought in the Middle East and southwest Asia. *Journal of climate*, *29*(23), 8547–8574. doi: 10.1175/JCLI-D-13-00692.1

Behling, R., Roessner, S., Foerster, S., Saemian, P., Tourian, M. J., Portele, T. C., & Lorenz, C. (2022). Interrelations of vegetation growth and water scarcity in Iran revealed by satellite time series. *Scientific Reports*, *12*(1), 20784. doi: 10.1038/s41598-022-24712-6

Box, G. E., & Muller, M. E. (1958). A note on the generation of random normal deviates. *The annals of mathematical statistics*, *29*(2), 610–611. doi: 10.1214/aoms/1177706645

Boyer, J., Byrne, P., Cassman, K., Cooper, M., Delmer, D., Greene, T., … others (2013). The US drought of 2012 in perspective: a call to action. *Global Food Security*, *2*(3), 139–143. doi: 10.1016/j.gfs.2013.08.002

Coelho, C. A., & Goddard, L. (2009). El niño–induced tropical droughts in climate change projections. *Journal of Climate*, *22*(23), 6456–6476. doi: 10.1175/2009JCLI3185.1

Coumou, D., & Rahmstorf, S. (2012). A decade of weather extremes. *Nature climate change*, *2*(7), 491. doi: 10.1038/nclimate1452

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. *Nature Climate Change*, *6*(5), 508. doi: 10.1038/nclimate2941

Famiglietti, J. S., Lo, M., Ho, S. L., Bethune, J., Anderson, K., Syed, T. H., … Rodell, M. (2011). Satellites measure recent rates of groundwater depletion in California's Central

Valley. *Geophysical Research Letters*, *38*(3). doi: 10.1029/2010GL046442

FAO. (2021). Drought [Computer software manual]. (`http://www.fao.org/land-water/water/drought/en/`)

Gardner, A. S., Moholdt, G., Cogley, J. G., Wouters, B., Arendt, A. A., Wahr, J., ... others (2013). A reconciled estimate of glacier contributions to sea level rise: 2003 to 2009. *science*, *340*(6134), 852–857. doi: 10.1126/science.1234532

Gouweleeuw, B. T., Kvas, A., Gruber, C., Gain, A. K., Mayer-Gürr, T., Flechtner, F., & Güntner, A. (2018). Daily GRACE gravity field solutions track major flood events in the Ganges–Brahmaputra Delta. *Hydrology and Earth System Sciences*, *22*(5), 2867–2880. doi: 10.5194/hess-22-2867-2018

Guha-Sapir, D., Below, R., & Hoyois, P. (2016). EM-DAT: the CRED/OFDA international disaster database. *Université Catholique de Louvain*. Retrieved from `https://www.emdat.be/`

Heberger, M. (2012). Australia's millennium drought: Impacts and responses. In *The world's water* (pp. 97–125). Springer.

Heim Jr, R. R. (2002). A review of twentieth-century drought indices used in the united states. *Bulletin of the American Meteorological Society*, *83*(8), 1149–1166. doi: 10.1175/1520-0477-83.8.1149

Hisdal, H., Stahl, K., Tallaksen, L. M., & Demuth, S. (2001). Have streamflow droughts in Europe become more severe or frequent? *International Journal of Climatology*, *21*(3), 317–333. doi: 10.1002/joc.619

Hosseini-Moghari, S.-M., Araghinejad, S., Ebrahimi, K., Tang, Q., & AghaKouchak, A. (2020). Using GRACE satellite observations for separating meteorological variability from anthropogenic impacts on water availability. *Scientific Reports*, *10*(1), 15098. doi: 10.1038/s41598-020-71837-7

Hulme, M. (1992). A 1951–80 global land precipitation climatology for the evaluation of general circulation models. *Climate dynamics*, *7*, 57–72. doi: 10.1007/BF00209609

Jiang, D., Wang, J., Huang, Y., Zhou, K., Ding, X., & Fu, J. (2014). The review of GRACE data applications in terrestrial hydrology monitoring. *Advances in Meteorology*, *2014*. doi: 10.1155/2014/725131

Jones, P. D., & Hulme, M. (1996). Calculating regional climatic time series for temperature and precipitation: methods and illustrations. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, *16*(4), 361–377. doi: 10.1002/(SICI)1097-0088(199604)16:4⟨361::AID-JOC53⟩3.0.CO;2-F

Kreibich, H., Van Loon, A. F., Schröter, K., Ward, P. J., Mazzoleni, M., Sairam, N., ... others (2022). The challenge of unprecedented floods and droughts in risk management. *Nature*, *608*(7921), 80–86. doi: 10.1038/s41586-022-04917-5

Kvas, A., Behzadpour, S., Ellmer, M., Klinger, B., Strasser, S., Zehentner, N., & Mayer-Gürr, T. (2019). ITSG-Grace2018: Overview and evaluation of a new GRACE-only gravity field time series. *Journal of Geophysical Research: Solid Earth*, *124*(8), 9332–9344. doi: 10.1029/2019JB017415

Li, W., Pan, R., Jiang, Z., Chen, Y., Li, L., Luo, J.-J., ... Yu, J. (2021). Future changes in the frequency of extreme droughts over China based on two large ensemble simulations. *Journal of Climate*, *34*(14), 6023–6035.

Lorenz, C., Kunstmann, H., Devaraju, B., Tourian, M. J., Sneeuw, N., & Riegger, J. (2014). Large-scale runoff from landmasses: a global assessment of the closure of the hydrological and atmospheric water balances. *Journal of Hydrometeorology*, *15*(6), 2111–2139. doi: 10.1175/JHM-D-13-0157.1

Metropolis, N., & Ulam, S. (1949). The Monte Carlo method. *Journal of the American statistical association*, *44*(247), 335–341.

Mishra, A. K., & Singh, V. P. (2010). A review of drought concepts. *Journal of hydrology*, *391*(1-2), 202–216.

Mishra, A. K., & Singh, V. P. (2011). Drought modeling–a review. *Journal of Hydrology*, *403*(1-2), 157–175.

Mooney, C. Z. (1997). *Monte Carlo simulation* (Vol. 116). Sage publications.

Narasimhan, B., & Srinivasan, R. (2005). Development and evaluation of Soil Moisture Deficit Index (SMDI) and Evapotranspiration Deficit Index (ETDI) for agricultural drought monitoring. *Agricultural and Forest Meteorology*, *133*(1-4), 69–88. doi: 10.1016/j.agrformet.2005.07.012

Naumann, G., Cammalleri, C., Mentaschi, L., & Feyen, L. (2021). Increased economic drought impacts in Europe with anthropogenic warming. *Nature Climate Change*, *11*(6), 485–491. doi: 10.1038/s41558-021-01044-3

Nicholson, S. E. (2014). A detailed look at the recent drought situation in the Greater Horn of Africa. *Journal of Arid Environments*, *103*, 71–79. doi: 10.1016/j.jaridenv.2013.12.003

Oikonomou, P. D., Karavitis, C. A., Tsesmelis, D. E., Kolokytha, E., & Maia, R. (2020). Drought characteristics assessment in Europe over the Past 50 Years. *Water Resources Management*, *34*(15), 4757–4772. doi: 10.1007/s11269-020-02688-0

Oleson, K., Lawrence, D., Lombardozzi, D., & Wieder, D. (2019). *CLM land-only release.* (Accessed: 2021-05-30) doi: https://doi.org/10.5065/d6154fwh

Palmer, W. C. (1965). *Meteorological drought* (Vol. 30). US Department of Commerce, Weather Bureau.

Reager, J. T., & Famiglietti, J. S. (2009). Global terrestrial water storage capacity and flood potential using GRACE. *Geophysical research letters*, *36*(23). doi: 10.1029/2009GL040826

Rodell, M., Chen, J., Kato, H., Famiglietti, J. S., Nigro, J., & Wilson, C. R. (2007). Estimating groundwater storage changes in the Mississippi River basin (USA) using GRACE. *Hydrogeology Journal*, *15*(1), 159–166. doi: 0.1007/s10040-006-0103-7

Rodell, M., Velicogna, I., & Famiglietti, J. S. (2009). Satellite-based estimates of groundwater depletion in India. *Nature*, *460*(7258), 999–1002. doi: 10.1038/nature08238

Saemian, P., Elmi, O., Vishwakarma, B., Tourian, M., & Sneeuw, N. (2020). Analyzing the Lake Urmia restoration progress using ground-based and spaceborne observations. *Science of The Total Environment*, *739*, 139857. doi: 10.1016/j.scitotenv.2020.139857

Saemian, P., Tourian, M. J., AghaKouchak, A., Madani, K., & Sneeuw, N. (2022). How much water did Iran lose over the last two decades? *Journal of Hydrology: Regional*

Studies, *41*, 101095. doi: 10.1016/j.ejrh.2022.101095

Scholz, C. A., Johnson, T. C., Cohen, A. S., King, J. W., Peck, J. A., Overpeck, J. T., ... others (2007). East african megadroughts between 135 and 75 thousand years ago and bearing on early-modern human origins. *Proceedings of the National Academy of Sciences*, *104*(42), 16416–16421.

Shepherd, A., Ivins, E., Rignot, E., Smith, B., Van Den Broeke, M., Velicogna, I., ... others (2018). Mass balance of the Antarctic Ice Sheet from 1992 to 2017. *Nature*, *558*, 219–222. doi: 10.1038/s41586-018-0179-y

Sinha, D., Syed, T. H., Famiglietti, J. S., Reager, J. T., & Thomas, R. C. (2017). Characterizing drought in India using GRACE observations of terrestrial water storage deficit. *Journal of Hydrometeorology*, *18*(2), 381–396. doi: 10.1175/JHM-D-16-0047.1

Sohn, B., Yeh, S.-W., Schmetz, J., & Song, H.-J. (2013). Observational evidences of Walker circulation change over the last 30 years contrasting with GCM results. *Climate Dynamics*, *40*, 1721–1732. doi: 10.1007/s00382-012-1484-z

Spinoni, J., Naumann, G., Vogt, J. V., & Barbosa, P. (2015). The biggest drought events in Europe from 1950 to 2012. *Journal of Hydrology: Regional Studies*, *3*, 509–524. doi: 10.1016/j.ejrh.2015.01.001

Sun, A. Y., Scanlon, B. R., AghaKouchak, A., & Zhang, Z. (2017). Using GRACE satellite gravimetry for assessing large-scale hydrologic extremes. *Remote Sensing*, *9*(12), 1287. doi: 10.3390/rs9121287

Svoboda, M., Fuchs, B., et al. (2016). *Handbook of drought indicators and indices.* World Meteorological Organization Geneva, Switzerland.

Svoboda, M., Hayes, M., & Wood, D. (2012). *Standardized precipitation index: user guide* (Tech. Rep. No. NO. 1090).

Svoboda, M., LeComte, D., Hayes, M., Heim, R., Gleason, K., Angel, J., ... others (2002). The drought monitor. *Bulletin of the American Meteorological Society*, *83*(8), 1181–1190.

Tapley, B. D., Bettadpur, S., Watkins, M., & Reigber, C. (2004). The gravity recovery and climate experiment: Mission overview and early results. *Geophysical research letters*, *31*(9). doi: 10.1029/2004GL019920

Teuling, A. J. (2018). A hot future for european droughts. *Nature Climate Change*, *8*(5), 364–365.

Thomas, A. C., Reager, J. T., Famiglietti, J. S., & Rodell, M. (2014). A GRACE-based water storage deficit approach for hydrological drought characterization. *Geophysical Research Letters*, *41*(5), 1537–1545. doi: 10.1002/2014GL059323

Tourian, M., Reager, J., & Sneeuw, N. (2018). The total drainable water storage of the Amazon river basin: A first estimate using GRACE. *Water Resources Research*, *54*(5), 3290–3312. doi: 10.1029/2017WR021674

Tourian, M. J., Elmi, O., Shafaghi, Y., Behnia, S., Saemian, P., Schlesinger, R., & Sneeuw, N. (2022). Hydrosat: geometric quantities of the global water cycle from geodetic satellites. *Earth System Science Data*, *14*(5), 2463–2486.

Tourian, M. J., Saemian, P., Ferreira, V. G., Sneeuw, N., Frappart, F., & Papa, F. (2023). A copula-supported Bayesian framework for spatial downscaling of GRACE-derived terrestrial water storage flux. *Remote Sensing of Environment*, *295*, 113685. doi: 10.1016/j.rse.2023.113685

van den Broeke, M., Bamber, J., Ettema, J., Rignot, E., Schrama, E., van de Berg, W. J., ... Wouters, B. (2009). Partitioning recent Greenland mass loss. *science*, *326*(5955), 984–986. doi: 10.1126/science.1178176

Van Dijk, A. I., Beck, H. E., Crosbie, R. S., de Jeu, R. A., Liu, Y. Y., Podger, G. M., ... Viney, N. R. (2013). The Millennium Drought in southeast Australia (2001–2009): Natural and human causes and implications for water resources, ecosystems, economy, and society. *Water Resources Research*, *49*(2), 1040–1057. doi: 10.1002/wrcr.20123

Voss, K. A., Famiglietti, J. S., Lo, M., De Linage, C., Rodell, M., & Swenson, S. C. (2013). Groundwater depletion in the Middle East from GRACE with implications for transboundary water management in the Tigris-Euphrates-Western Iran region. *Water resources research*, *49*(2), 904–914. doi: 10.1002/wrcr.20078

Wilhite, D. A. (2000). Chapter 1 drought as a natural hazard: Concepts and definitions.

Drought Mitigation Center Faculty Publications.

Wilhite, D. A., Svoboda, M. D., & Hayes, M. J. (2007). Understanding the complex impacts of drought: A key to enhancing drought mitigation and preparedness. *Water resources management*, *21*(5), 763–774. doi: 10.1007/s11269-006-9076-5

Yi, H., & Wen, L. (2016). Satellite gravity measurement monitoring terrestrial water storage change and drought in the continental United States. *Scientific reports*, *6*, 19909. doi: 10.1038/srep19909

Yi, S., Saemian, P., Sneeuw, N., & Tourian, M. J. (2023). Estimating runoff from pan-Arctic drainage basins for 2002–2019 using an improved runoff-storage relationship. *Remote Sensing of Environment*, *298*, 113816. doi: 10.1016/j.rse.2023.113816

Yirdaw, S. Z., Snelgrove, K. R., & Agboma, C. O. (2008). GRACE satellite observations of terrestrial moisture changes for drought characterization in the Canadian Prairie. *Journal of Hydrology*, *356*(1-2), 84–92. doi: 10.1016/j.jhydrol.2008.04.004

Yu, M., Li, Q., Hayes, M. J., Svoboda, M. D., & Heim, R. R. (2014). Are droughts becoming more frequent or severe in China based on the standardized precipitation evapotranspiration index: 1951–2010? *International Journal of Climatology*, *34*(3), 545–558. doi: 10.1002/joc.3701

Zargar, A., Sadiq, R., Naser, B., & Khan, F. I. (2011). A review of drought indices. *Environmental Reviews*, *19*(NA), 333–349. doi: doi.org/10.1139/a11-013

Zhao, C., Brissette, F., Chen, J., & Martel, J.-L. (2020). Frequency change of future extreme summer meteorological and hydrological droughts over north america. *Journal of Hydrology*, *584*, 124316.

Zhao, M., Geruo, A., Velicogna, I., & Kimball, J. S. (2017). A global gridded dataset of GRACE drought severity index for 2002–14: Comparison with PDSI and SPEI and a case study of the Australia millennium drought. *Journal of Hydrometeorology*, *18*(8), 2117–2129. doi: 10.1175/JHM-D-16-0182.1

Zhong, Y., Hu, E., Wu, Y., An, Q., Wang, C., Bai, H., & Gao, W. (2023). Reconstructing a long-term water storage-based drought index in the yangtze river basin. *Science of The*

Total Environment, *883*, 163403.

Zscheischler, J., Westra, S., Van Den Hurk, B. J., Seneviratne, S. I., Ward, P. J., Pitman, A., . . . others (2018). Future climate risk from compound events. *Nature Climate Change*, *8*(6), 469–477. doi: 10.1038/s41558-018-0156-3

A probabilistic approach to characterizing drought using satellite gravimetry

Peyman Saemian[1], Mohammad J. Tourian[1], Omid Elmi[1]
Nico Sneeuw[1], Amir AghaKouchak[2,3]

[1]Institute of Geodesy, University of Stuttgart, Stuttgart, Germany
[2]Department of Civil and Environmental Engineering, University of California, Irvine, CA, USA.
[3]United Nations University Institute for Water, Environment and Health, Hamilton, Ontario, Canada

Key Points:

- A probabilistic framework is introduced to characterize drought using GRACE and GRACE Follow-On observations.
- Our study highlights a tendency of deterministic approaches to consistently overestimate storage-based drought severity.
- The probabilistic approach captures global droughts while delivering more realistic results suited for risk management.

Corresponding author: Peyman Saemian, `peyman.saemian@gis.uni-stuttgart.de`

Abstract

In the recent past, the Gravity Recovery and Climate Experiment (GRACE) satellite mission and its successor GRACE Follow-On (GRACE-FO), have become invaluable tools for characterizing drought through measurements of Total Water Storage Anomaly (TWSA). However, the existing approaches have often overlooked the uncertainties in TWSA that stem from GRACE orbit configuration, background models, and intrinsic data errors. Here we introduce a fresh view on this problem which incorporates the uncertainties in the data: the Probabilistic Storage-based Drought Index (PSDI). Our method leverages Monte Carlo simulations to yield realistic realizations for the stochastic process of the TWSA time series. These realizations depict a range of plausible drought scenarios that later on are used to characterize drought. This approach provides probability for each drought category instead of selecting a single final category at each epoch. We have compared PSDI with the deterministic approach (SDI) over major global basins. Our results show that the deterministic approach often leans towards an overestimation of storage-based drought severity. Furthermore, we scrutinize the performance of PSDI across diverse hydrologic events, spanning continents from the United States to Europe, the Middle East, Southern Africa, South America, and Australia. In each case, PSDI emerges as a reliable indicator for characterizing drought conditions, providing a more comprehensive perspective than traditional deterministic indices. In contrast to the common deterministic view, our probabilistic approach provides a more realistic characterization of the TWS drought, making it more suited for adaptive strategies and realistic risk management.

Plain Language Summary

Total Water Storage (TWS) is defined as the sum of water stored as surface water (e.g., lakes and rivers), groundwater, soil moisture, snow, ice, and vegetation biomass. Since its launch in 2002, the Gravity Recovery and Climate Experiment (GRACE) satellite mission has provided unique TWS change measurements with manifold applications in hydrology,

including characterizing drought events. Scientists have been using satellites like GRACE and its successor, GRACE-FO, to understand drought by measuring the Total Water Storage Anomaly (TWSA). However, previous methods didn't consider uncertainties from satellite orbits, models, and data errors. This study offers a novel probabilistic approach for characterizing drought, Probabilistic Storage-based Drought Index (PSDI), which acknowledges the uncertainties in the GRACE TWS change. We use simulations to create different drought scenarios, offering probabilities for each category instead of one fixed category. Comparing PSDI to traditional methods, we found that traditional methods tend to overestimate drought severity. We tested PSDI across different regions, and it consistently proved to be a reliable way to understand drought conditions, offering a more comprehensive perspective. Our probabilistic approach offers a more realistic view of TWS drought, making it suitable for adaptive strategies and risk management.

1 Introduction

The modern reality of human settlement is the consequence of many historical events, but perhaps none influenced human settlements as much as droughts and famine. DNA analysis indicates that a series of extreme droughts that occurred 75-135 thousand years ago may have been the reason for the first human migration out of Africa (Scholz et al., 2007). Following several consequential droughts over the past century (e.g., the 1921 drought in Europe, the 1930s Dust Bowl drought in the US, 1928-1930 drought in China, 1980s drought and famine in Africa, 2000s Millennium drought in Australia), increasingly more effort has focused on understanding, monitoring and predicting droughts and their impacts (Mishra & Singh, 2010; Heim Jr, 2002; AghaKouchak et al., 2015; Svoboda et al., 2002; Wilhite et al., 2007; Kreibich et al., 2022; AghaKouchak et al., 2021).

Compared to other hazards witnessed over the past four decades, drought impacts are often felt by a much larger number of people worldwide (Wilhite, 2000; FAO, 2021; AghaKouchak et al., 2021). Numerous nations have grappled with significant economic losses resulting from drought events. Notably, according to the NOAA's National Centers for Environmental Information (NCEI) report, the United States has experienced 26 significant droughts in the past century, amounting to a staggering economic loss of at least \$ 249 billion, equivalent to nearly \$ 10 billion per occurrence. In Europe, the southern and western regions, in particular, face an annual drought-related expenditure estimated at up to €9 billion, which could surge to over €65 billion if climate action is not taken (Naumann et al., 2021). Aside from the financial burdens, climate change, and unsustainable water management practices have amplified the frequency and severity of drought occurrences worldwide over the past two decades. This trend is projected to escalate further in the future (see e.g., Hisdal et al., 2001; Coumou & Rahmstorf, 2012; Yu et al., 2014; Donat et al., 2016; Teuling, 2018; Li et al., 2021; C. Zhao et al., 2020).

The negative consequences of drought can be effectively alleviated through the implementation of risk management strategies rather than relying on crisis management (Wilhite, 2000; Zscheischler et al., 2018). Such a proactive response may be achieved by establishing reliable drought monitoring systems, including early warning systems and forecasting capabilities, operating at both national and local levels (Wilhite et al., 2007; AghaKouchak et al., 2023). These systems trigger a series of decisions aimed at helping communities navigate the challenges posed by drought events (Mishra & Singh, 2011; Sun et al., 2017). To enhance drought monitoring efforts and provide valuable guidance to decision-makers, numerous drought indices have been developed (Mishra & Singh, 2010). These indices condense the intricacies of drought into a single numerical value, effectively characterizing its onset, intensity, frequency, and duration (Zargar et al., 2011; Wilhite, 2000; Ahmadalipour et al., 2017). Such indices offer a comprehensive representation of drought by utilizing single or multiple climatic and hydrometeorological variables such as precipitation, streamflow, evapotranspiration, temperature, and snowpack (e.g., Svoboda et al., 2016; Hosseini-Moghari et al., 2020).

A comprehensive understanding of drought dynamics necessitates the observation of Total Water Storage (TWS) including snow, surface water, soil moisture, and groundwater storage (M. Zhao et al., 2017; M. J. Tourian et al., 2023). Traditionally, TWS monitoring has relied on costly and time-consuming site measurements, providing limited regional and local coverage. While hydrological and land surface models partially address this issue, estimating TWS in regions lacking in-situ runoff data for calibrating rainfall-runoff models still yields high uncertainties (Jiang et al., 2014; S. Yi et al., 2023). Since its launch in 2002, the Gravity Recovery And Climate Experiment (GRACE) satellite mission has revolutionized the remote measurement of TWS Anomalies (TWSA) at regional to continental scales (Tapley et al., 2004; M. J. Tourian et al., 2022). The GRACE mission came to an end on 12 October 2017, due to battery failure, after more than 15 years of Earth observation. However, its successor, GRACE Follow-On (GRACE-FO), has continued the GRACE legacy since its launch on 22 May 2018. GRACE(-FO) data have been extensively utilized for manifold applications, including monitoring ice sheets and glaciers (e.g., van den Broeke et al., 2009; Gardner et al., 2013; Shepherd et al., 2018), tracking anthropogenic groundwater depletion (e.g., Rodell et al., 2007, 2009; Famiglietti et al., 2011; Voss et al., 2013; Saemian et al., 2022), forecasting flood events (e.g., Reager & Famiglietti, 2009; Gouweleeuw et al., 2018), and quantifying and comprehending hydrological processes (e.g., Lorenz et al., 2014; Saemian et al., 2020; M. Tourian et al., 2018; Behling et al., 2022), to name but a few.

GRACE-derived estimates of TWS have been employed in developing indices aimed at assessing drought on a regional to global scale. For example, Yirdaw et al. (2008) developed the Total Storage Deficit Index (TSDI), utilizing the Palmer Drought Severity Index (PDSI; Palmer, 1965) and the Soil Moisture Deficit Index (SMDI; Narasimhan & Srinivasan, 2005), to characterize the Canadian Prairie droughts of 2002/2003. Another notable endeavor by Thomas et al. (2014) presented a comprehensive framework for drought characterization based on GRACE-derived TWSA over regions including the Amazon, Zambezi, Texas, and the southeastern United States. Additionally, H. Yi & Wen (2016) devised the GRACE-based Hydrological Drought Index (GHDI) to characterize drought in the continental United

States from 2003 to 2012, building upon the foundation of the PDSI concept. Among recent indicators we can name the Drought Severity Index (DSI) by M. Zhao et al. (2017), the Water Storage Deficit Index (WSDI) by Sinha et al. (2017), and a long-term standardized GRACE reconstructed TWSA index (SGRTI) by Zhong et al. (2023).

The indices mentioned above have the potential for monitoring and assessing the TWS drought at regional to global scales. Nevertheless, they adopt a deterministic approach that disregards the intrinsic uncertainties associated with characterizing drought using GRACE observations. These uncertainties are inherent in the GRACE data due to factors such as its orbit configuration, measurement concept, various post-processing approaches of GRACE data, and different options for de-aliasing products. Besides, the estimation of GRACE uncertainty varies among different GRACE level-2 products, in both magnitude and spatial pattern. Most of the centers offer an uncertainty measure (known as formal errors) in the form of spherical harmonic coefficients. Figure 1 shows the coefficient-wise ratio of average formal errors and empirical errors of the GRACE solution following the approach suggested by Kvas et al. (2019). The ideal ratio is set at one, with values below indicating an underestimation of empirical errors, whereas values exceeding one signify an overestimation. The three official centers (JPL, CSR, and GFZ) together with AIUB, HUST, and SWPU exhibit a similar pattern. The SWPU solution demonstrates a pronounced overestimation of formal errors, particularly for low d/o (under 30). ITSG and Tongji display comparable patterns, although Tongji tends to lean slightly towards a more pessimistic estimation of errors. In contrast, COST-G reflects more realistic formal errors in comparison to empirical errors but appears overly optimistic for lower d/o values. The distinctive pattern observed in the CNES product can be attributed to the regularization applied during the derivation of the gravity field from the Level-1 dataset. The disparities in formal error performance among GRACE Level-2 products underscore the inherent uncertainty in GRACE data and consequently, the necessity for a comprehensive approach that effectively considers GRACE uncertainty while characterizing drought.

To address this gap in conventional methods, we propose a probabilistic approach that considers all possible scenarios and associated impacts. Our approach leverages Monte Carlo simulations to obtain realistic realizations of TWSA compatible with GRACE-TWSA and its corresponding uncertainties. We then characterize drought based on TWSA using our proposed drought index, the Probabilistic Storage-based Drought Index (PSDI). This index indicates not only drought but also its corresponding occurrence probability. We compare our results with those from the conventional deterministic approaches over the major river basins. Moreover, the performance of PSDI in capturing the main hydrological drought extremes is examined within the GRACE era. PSDI facilitates more informed and proactive responses to water resource challenges and serves as a practical tool for decision-makers and water resource managers to assess and manage drought-related risks more realistically.

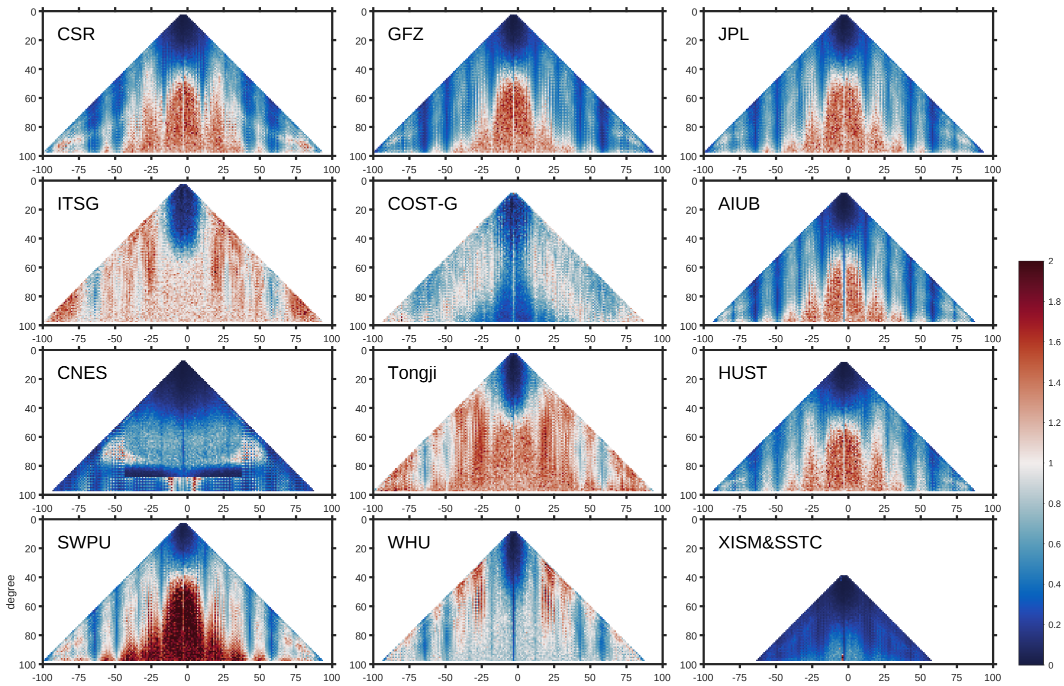

Figure 1. The coefficient-wise ratio of average formal errors and empirical errors of the GRACE solutions following the approach suggested by Kvas et al. (2019). For the average formal errors, the mean of the reported variance of the spherical harmonics coefficients is computed for monthly solutions from January 2005 to December 2010, which is assumed to hold a homogeneous data quality. To estimate the empirical errors, we compute the standard deviation of the coefficients after removing the mean, linear trend, and annual and semi-annual signals. The optimal value of the ratio is one, and values below one indicate an underestimation of the empirical errors, while values bigger than one show overestimation. We have only included solutions that provide formal errors.

2 Data and Method

2.1 GRACE data

The GRACE TWSA can be obtained from the two main approaches, namely Spherical Harmonics (SHs) and mass concentration blocks (mascons). In the former, one needs to apply post-processing steps including noise reduction and signal restoration while the latter is already the Level-3 product (gridded TWSA over the globe). These approaches are briefly described in section 1 of the supplementary file. In line with the common practice within the hydrology community, we have utilized the mascons solutions. The probabilistic approach for characterizing storage-based drought index, however, can readily be applied to any level-3 products that provide estimations for GRACE TWSA and its corresponding uncertainty. Among the mascon products, we have employed the one from Goddard Space Flight Center (GSFC), NASA. The GSFC mascon product has been widely used in the geodesy and Earth science communities to investigate a range of phenomena, including hydrology, glaciology, and solid Earth dynamics, and can be downloaded from `https://earth.gsfc.nasa.gov/geo/data/grace-mascons`. We used the latest version of the dataset available at the time of our analysis, which covers the period from August 2002 to November 2022. The dataset includes monthly gravity field solutions with a grid size of $0.5\,^\circ$.

2.2 Methodology

We propose a probabilistic framework to characterize storage-based drought. The framework is illustrated in Figure 2, Figure 3, and Figure 4 using TWSA over the Death Valley basin in the US as an example. To characterize drought, we must first define a reference, based on which a prolonged relative water deficiency is determined. It is common to consider the long-term monthly average, also known as the *climatology*, as the reference or *normal* condition in a region. Obtaining accurate climatology from short time series can be challenging. Calculating the climatology over at least 30 years, preferably 60 years, is standard practice, as this time frame allows us to average out the effects of short-term variability, resulting in a more robust estimate of the long-term average conditions (e.g., Hulme, 1992; Jones & Hulme, 1996; Svoboda et al., 2012). The GRACE and GRACE-FO missions, with their approximately 20-year duration, fall short of providing sufficient data for calculating long-term climatology. In this study, we utilized a combination of different models to estimate TWSA dating back to 1980. To this end, we incorporated a total of 13 state-of-the-art datasets including Global Hydrological Models (GHMs), Land Surface Models (LSMs), and atmospheric reanalysis models. To combine models, we employed the Multivariate Linear Regression (MLR) method. We compared the results of Multiple Linear Regression (MLR) in reconstructing TWSA with GRACE time series (see Supplementary section 2). The MLR exhibits a strong capability to capture the features and effectively reconstruct TWSA data as far back as 1980. It demonstrated superior performance compared to the ensemble mean of the models, as indicated by a substantial improvement in both the correlation coefficient (on average from 0.87 to 0.97) and the Kling-Gupta Efficiency (KGE) score (on average from 0.27 to 0.95) across major river basins. For more details about the datasets and long-term TWSA, please refer to Supplementary section 2. This extended time frame enables us to capture significant climate events and phenomena that influence long-term climate, such as the El Niño-Southern Oscillation (ENSO) and the North Atlantic Oscillation (NAO) (Sohn et al., 2013; Coelho & Goddard, 2009).

The climatology, along with its corresponding uncertainty (see Figure 2 (a)), is obtained by:

$$\overline{\mathrm{TWSA}}[t_m] = \frac{1}{N}\sum_{y=y_1}^{y_N} \mathrm{TWSA}[t_{y,m}] \tag{1}$$

$$\sigma^2_{\overline{\mathrm{TWSA}}[t_m]} = \frac{1}{N}\sqrt{\sum_{y=y_1}^{y_N} \sigma^2_{\mathrm{TWSA}[t_{y,m}]}} \tag{2}$$

where $\overline{\mathrm{TWSA}}[t_m]$ represents the TWSA climatology for month m, y denotes the year and can vary from y_1 to y_N, m corresponds to the month within a year, taking values $1, 2, 3, \ldots, 12$, and N is the number of years in the long-term dataset. Note that we deliberately retain the trend in the time series. We reason that the trend reflects long-term changes in climate, such as temperature increases or precipitation pattern alterations, and that it affects the frequency and severity of droughts (see Supplementary section 3 for more details). We then subtract the climatology from the GRACE TWSA time series to obtain TWS residual (S):

$$S[t_{y,m}] = \mathrm{TWSA}[t_{y,m}] - \overline{\mathrm{TWSA}}[t_m] \tag{3}$$

$$\sigma_{S[t_{y,m}]} = \sqrt{\sigma^2_{\mathrm{TWSA}[t_{y,m}]} + \sigma^2_{\overline{\mathrm{TWSA}}[t_m]}} \tag{4}$$

where negative values of S represent water storage deficits.

To reduce the effects of short-term fluctuations due to precipitation and other factors, we chose to use a 3-month moving average to smooth the TWS residual (see Figure 2 (b)):

$$S[t] = f * S[t]_{\mathrm{unsmoothed}} \tag{5}$$

where $*$ denotes the convolution operation, and f is the kernel [1/3, 1/3, 1/3] which is convolved with the $S[t]_{\mathrm{unsmoothed}}$ time series.

To address the inherent uncertainty, it becomes essential to employ a stochastic approach that incorporates S along with its associated uncertainty. At each epoch, we postulate a normal distribution with the mean being the obtained TWS residual $S[t]$ and $\sigma = \sigma_{S[t]}$. Sampling from this distribution in each time step allows us to create realizations of the S time series. Several methods exist for sampling from a Gaussian distribution, and one widely used technique is the Box-Muller transform (Box & Muller, 1958). This method guarantees generating a realization of S, which is independent epoch-wise with no artificial correlations. We then use Monte Carlo Simulation (Mooney, 1997; Metropolis & Ulam, 1949) to generate

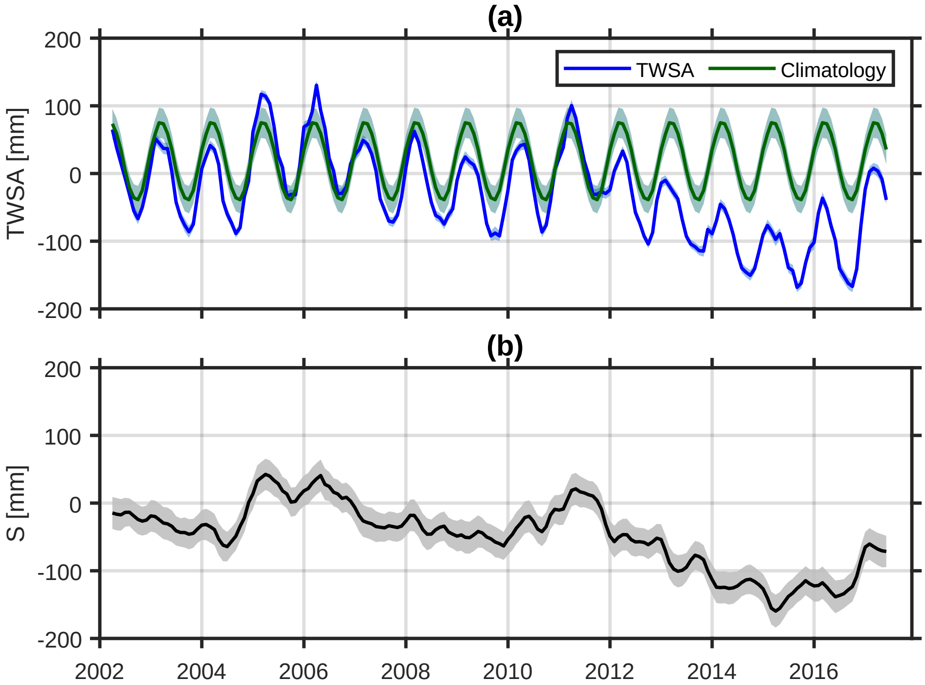

Figure 2. (a) Time series of the long-term TWSA from GRACE and the long-term climatology (1980–2012) from the hindcasted TWSA together with their uncertainties. At each epoch, we assume a Gaussian distribution for the uncertainties and the depicted uncertainty corresponds to the 1-σ level. Here the results are shown for the Death Valley basin in the US.

multiple realizations of S. Figure 3 shows 10 000 realization of TWS residual considering the 3-σ uncertainty. The density of the realizations is highest around the mean signal and decays following a Gaussian distribution. The colored lines overlaid on the time series depict the distribution of outcomes for three specific epochs: July 2004, May 2013, and April 2015.

To characterize drought within each time epoch, one common approach is to use the percentile rank method and the U.S. Drought Monitor (USDM) criteria. To this end, a set of five drought categories is defined (Table 1).

The quantile values for the S time series can be extracted as the inverse of the Cumulative Distribution Function (CDF), also known as the quantile function:

$$Q(p) = F^{-1}(p) \tag{6}$$

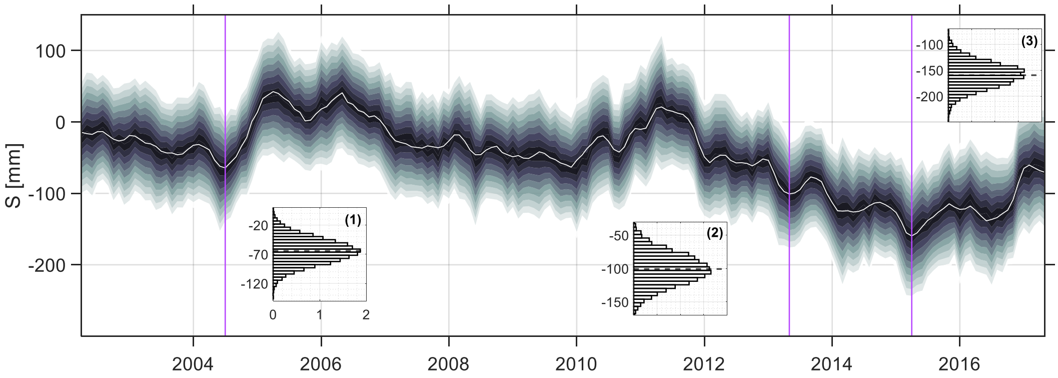

Figure 3. The TWS residual (S) together with its 10 000 realizations, calculated using Monte Carlo simulation. Here the results are shown for the Death Valley basin in the US. The distribution of realizations for three epochs, namely July 2004, May 2013, and April 2015 are marked with colored dots over the time series and are shown in sub-figures.

Table 1. Drought categories and corresponding percentile ranges as defined by the U.S. Drought Monitor (USDM).

Drought Category	Description	Percentile Range
D0	Abnormally dry	20–30%
D1	Moderate drought	10–20%
D2	Severe drought	5–10%
D3	Extreme drought	2–5%
D4	Exceptional drought	Less than 2%

where $Q(p)$ represents the quantile function and $F^{-1}(p)$ denotes the inverse CDF evaluated at probability p. In a conventional deterministic approach, the drought category for each epoch is determined based on its quantile value of S. For example in Figure 4 (a), the dark solid line represents the quantile function of the S time series. Using such a function one can characterize drought for case (1), case (2), and case (3) as D4, D1, and no drought, respectively.

Such an approach overlooks the uncertainty in TWS residual S. However, accounting for uncertainty would entail obtaining the quantile function for all realizations of S. These functions form a cloud of points rather than a single line as it has been illustrated in Figure 4 (a). The quantile functions are shown in grayscale representing the probability $\Pr(p, S)$ for a given percentile p and TWS residual S. Already at this stage, a glance at Figure 4 (a) reveals the complexity introduced by the uncertainty envelope, challenging the conventional approach to assigning a specific class to a particular measurement. It is noteworthy that the uncertainty envelope depicted in Figure 4 (a) exhibits a stationarity character, indicating general uncertainty in the data regardless of the specific time of measurements. This characteristic is reflective of the general uncertainty of S, emphasizing the broader statistical context rather than being tied to specific instances in time.

Now, let's delve into the characterization of drought for one of the measurements illustrated in Figure 3. In this context, alongside the consideration of the stationary uncertainty as reflected in the quantile envelope and represented by $\Pr(p, S)$, it becomes essential to account for the uncertainty associated with the measurement at that specific epoch. This is fundamentally crucial because two GRACE measurements with the same value of S may exhibit varying levels of uncertainty. Therefore, we incorporate the probability density function of the value S_t, denoted by $f(S_t)$, obtained from the mean and uncertainty of that epoch. $f(S_t)$ is shown for the three sample epochs on the top right panel of Figure 4. At each epoch, we multiply this probability density function with the entire distribution $\Pr(p, S)$, as illustrated in Figure 4 (c). Essentially, this multiplication results in a down-weighting of probabilities located in the tails of $f(S_t)$.

Once $\Pr(p, S)\, f(S_t)$ is achieved, to obtain PSDI at each epoch and for each drought category D_i, we can integrate the probabilities both in S and p domains and normalize it with the integral over the entire domain:

$$\mathrm{PSDI}(t, D_i) = \frac{\int_S \int_{D_i} \Pr(p, S) f(S_t)\, dp\, ds}{\int_S \int_0^1 \Pr(p, S) f(S_t)\, dp\, ds} \tag{7}$$

By performing this process for all drought categories and time epochs, we generate a comprehensive probabilistic representation of drought severity over time. For decision-making purposes, the highest-probability category can be judiciously chosen as the definitive drought classification for a particular month. The flowchart of the proposed probabilistic approach is shown in Figure 5.

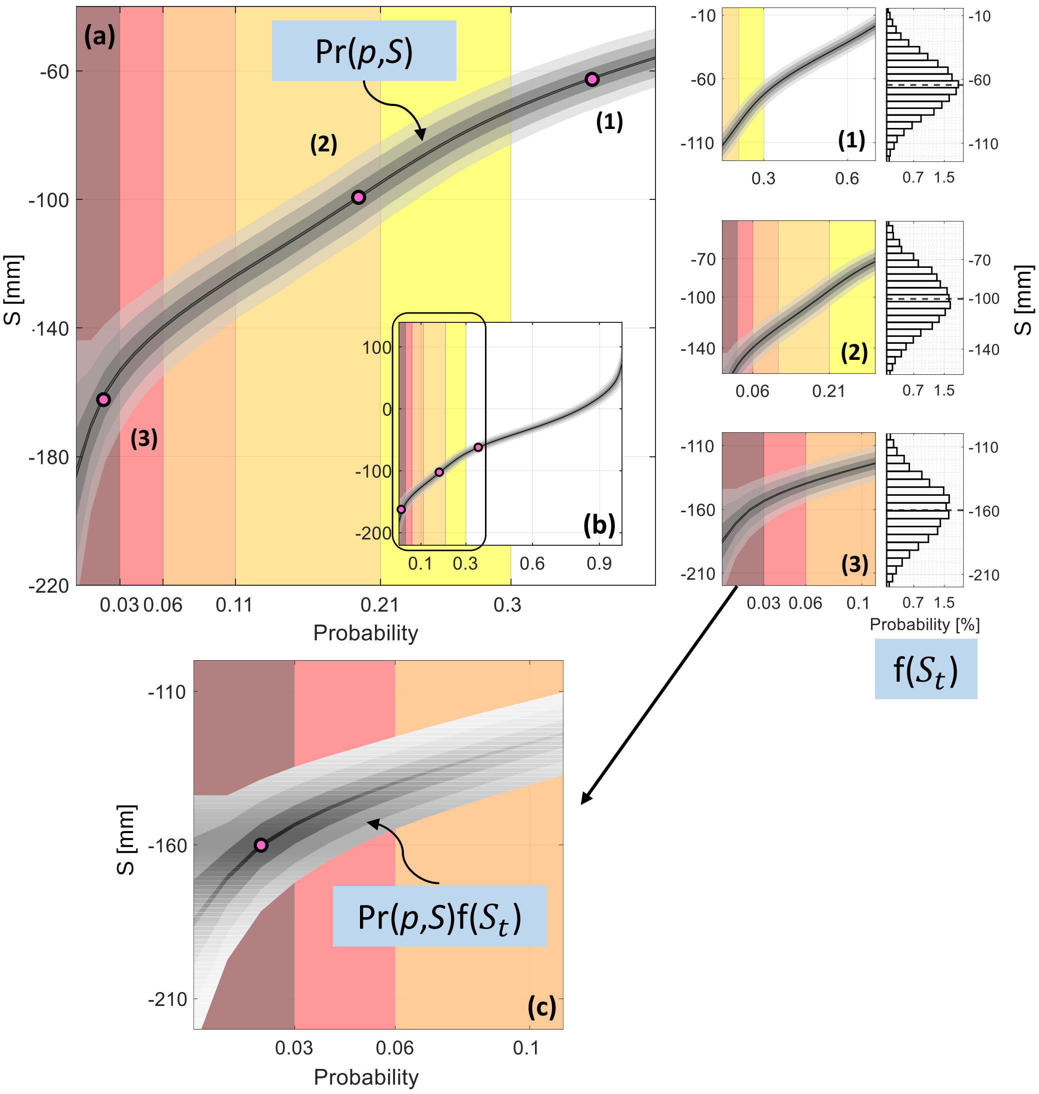

Figure 4. (a) The quantile functions (inverse of the Cumulative Distribution Function (CDF)) of the S realizations are depicted. The varying shades of gray signify the density of data points, with darker shades indicating higher density. The drought categories, ranging from D0 to D4, are delineated within their respective percentile ranges, each denoted by its corresponding color. Colored dots illustrate the positions of the three cases from Figure 3 on the quantile functions plot. These cases are further elaborated in a magnified view, accompanied by the corresponding probability distribution derived from S and its associated uncertainty. (b) Similarly, as in (a), this visualization portrays the quantile functions plot, but encompasses the complete range of quantile values. (c) The density of the counted points after integrating the probability distribution stemming from the S and its corresponding uncertainty. It's important to note that the presented results are centered on the Death Valley basin within the United States.

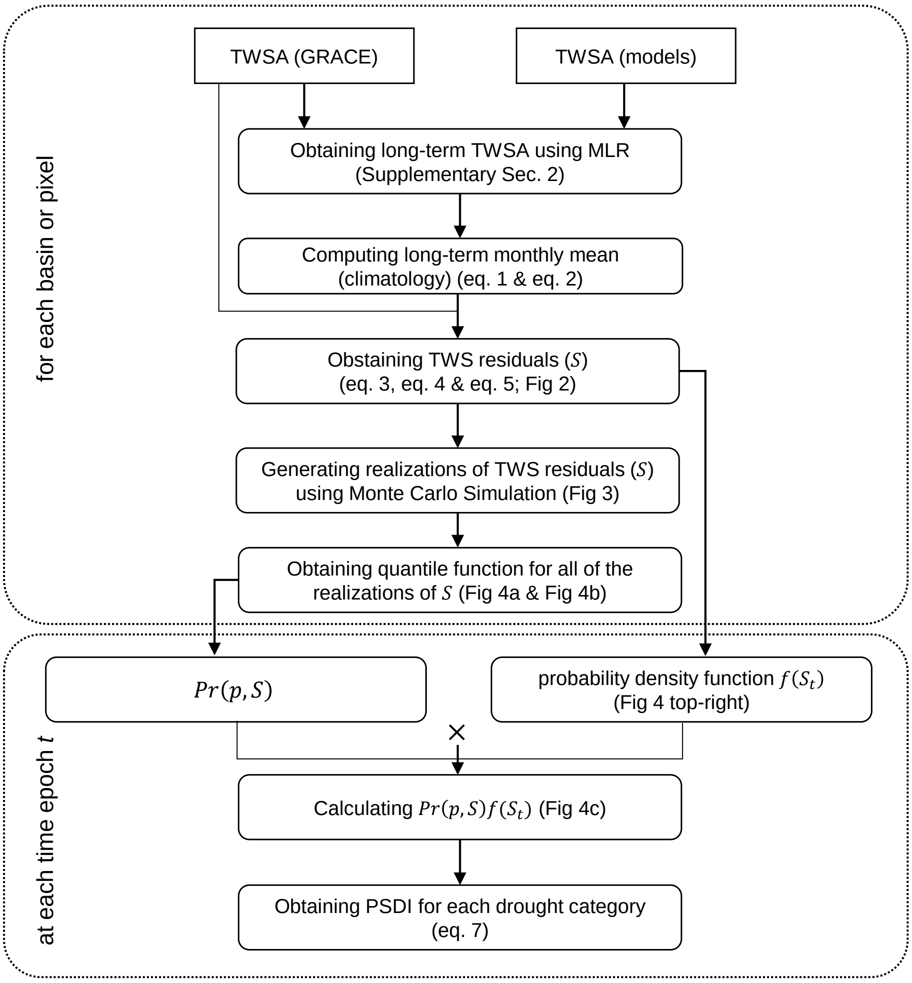

Figure 5. Flowchart of the proposed PSDI framework.

3 Results and Discussion

3.1 PSDI vs SDI

The PSDI approach offers a more nuanced understanding of drought conditions compared to the SDI approach. This is because PSDI captures the uncertainty associated with drought severity, while the SDI approach may oversimplify the classification of drought conditions. Although the SDI categorization is often the most probable category according to the PSDI, the neighboring categories may also have significant probabilities. This tendency becomes more pronounced as the intensity of the drought increases. This can be attributed to the lower slope of the CDF curve over more severe droughts and the wider range of quantile values.

To delve deeper into the analysis, we have quantified the disparities between drought categorizations as defined by SDI and $\mathrm{PSDI}_{\mathrm{max}}$—the category of drought with the highest probability in PSDI—across the world's major river basins, with the exclusion of Greenland and Antarctica. The findings, illustrated in Figure 6 (a), shed light on the prevalence of these discrepancies throughout the study period spanning from 2003 to 2016. The ratio exhibits a range of variations, hovering near zero for basins such as Lake Balkhash in southeastern Kazakhstan or Po in Italy to a significant value of 30 % over Highland of Ethiopia and Somalia in Africa or Sao Francisco in Brazil. In general, the risk of mischaracterizing storage-based drought through the deterministic approach is notably high (exceeding 10 % in Figure 6 (a)) across Africa (excluding the northern region), Eastern Europe, Mongolia, Russia, and within the river basins of Nelson river, St. Lawrence, and Colorado (Argentina). In instances where discrepancies arise between SDI and $\mathrm{PSDI}_{\mathrm{max}}$, a predominant tendency is for SDI to overestimate the drought category. This is evident when comparing Figure 6 (b) and Figure 6 (c).

To investigate further, Figure 7 provides a visual comparison between two approaches for characterizing drought: probabilistic (PSDI) and deterministic (SDI), over several selected basins. The distribution of the basins is shown in the top panel of the Figure 7. For each basin, the drought categories, ranging from the status of no drought to exceptional drought (D4), are displayed in columns. The probability assigned to each category at every time step is depicted using gray scale. The deterministic perspective is illustrated with red boxes, allowing for a direct comparison of the two approaches.

The Danube and Ganges basins exhibited no disparity between SDI and $\mathrm{PSDI}_{\mathrm{max}}$ from 2015 to 2016. In contrast, the Mississippi basin displayed the most substantial mismatch between SDI and $\mathrm{PSDI}_{\mathrm{max}}$. It's noteworthy that these mismatches were confined to adjacent categories. Specifically, when considering mismatches spanning more than one category, only

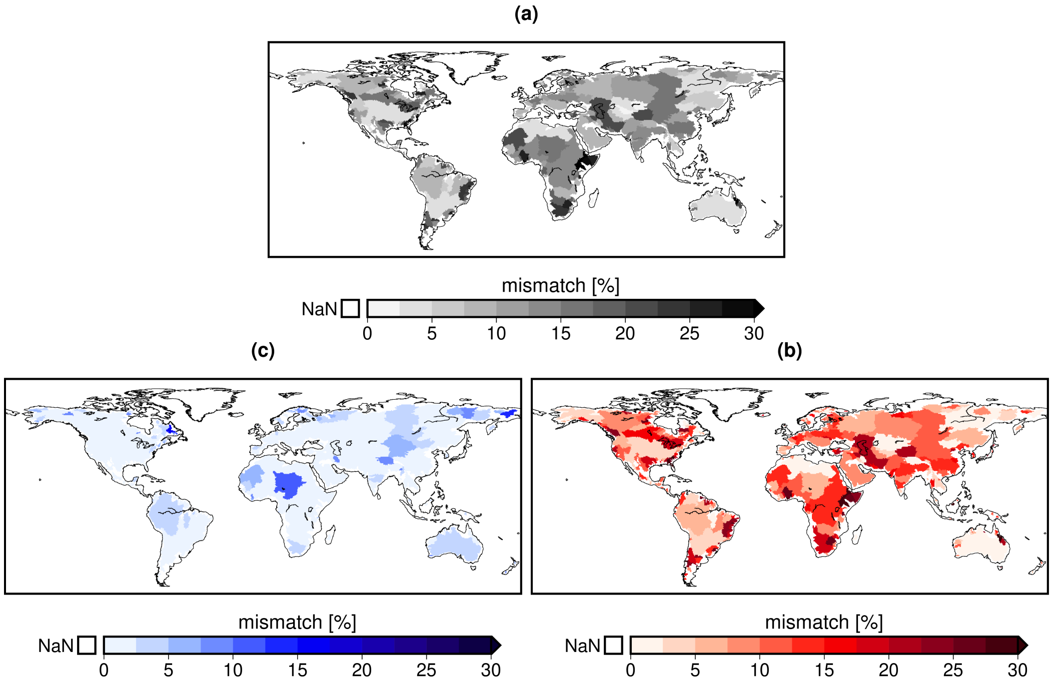

Figure 6. (a) Basin-wise distribution of the discrepancies between SDI and $PSDI_{max}$. The values represent the percentage of epochs where $PSDI_{max}$ differs from SDI by at least two drought categories. (b) the percentage of epochs with a discrepancy of more than one category higher in SDI compared to $PSDI_{max}$. (b) the percentage of epochs with a discrepancy of more than one category lower in SDI compared to $PSDI_{max}$. Greenland and Antarctica are excluded from the maps.

four basins had such occurrences: one month in Amazonas and Nile, two months in Niger, and five months in Murray Darling. Across all basins, when a discrepancy arose between SDI and $PSDI_{max}$, the SDI category consistently indicated a higher severity of drought.

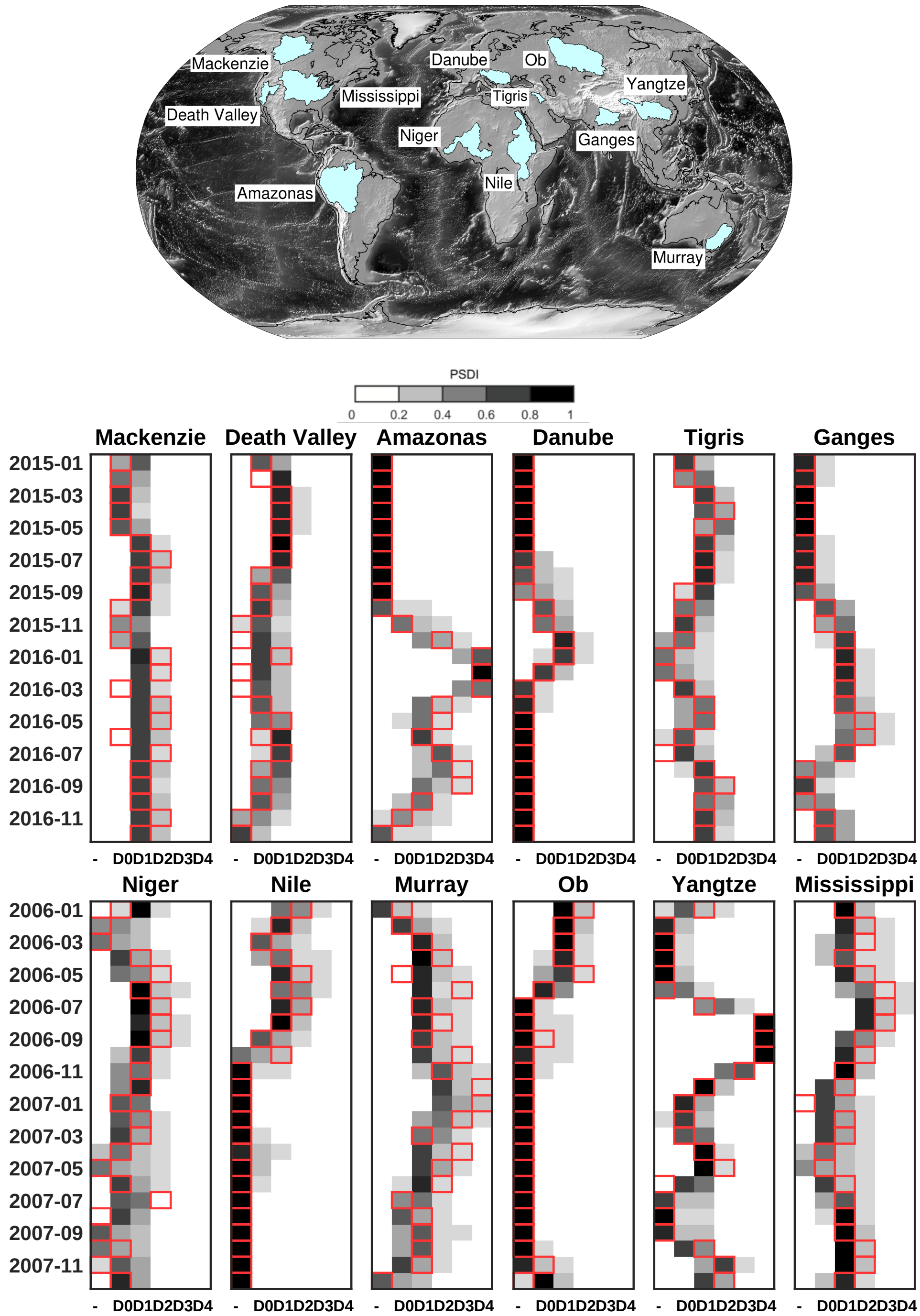

Figure 7. Top: The global distribution of the selected basins. Bottom: The SDI (red boxes) together with PSDI (gray scale probability range) for selected basins. The basins are shown in two groups considering the period with more frequency of drought, the first row between 2015–2016 and the second row between 2006–2007. The "-" represents "no drought" or "normal state" of the water storage.

We have investigated further the sensitivity of different categories of drought to incorporating uncertainties into drought characterization. Figure 8 (a) visualizes the percentage of epochs where $PSDI_{max}$ differs from SDI by at least two drought categories. The results suggest that such discrepancies can diverge significantly in the categorization of drought conditions, especially in the D1, D2, and D3 categories, especially in D2. We have also compared the ratio of the mismatch period with respect to different climate categories (Figure 8 (b)). For more detailed information about the categories and the method of classification, please see section 4 in the supplementary file. Although the mismatch range can vary from arid to humid climate, the average value of the mismatch is the same over different climate regions, with a slightly higher value for the Dry sub-humid regions (Dry sH).

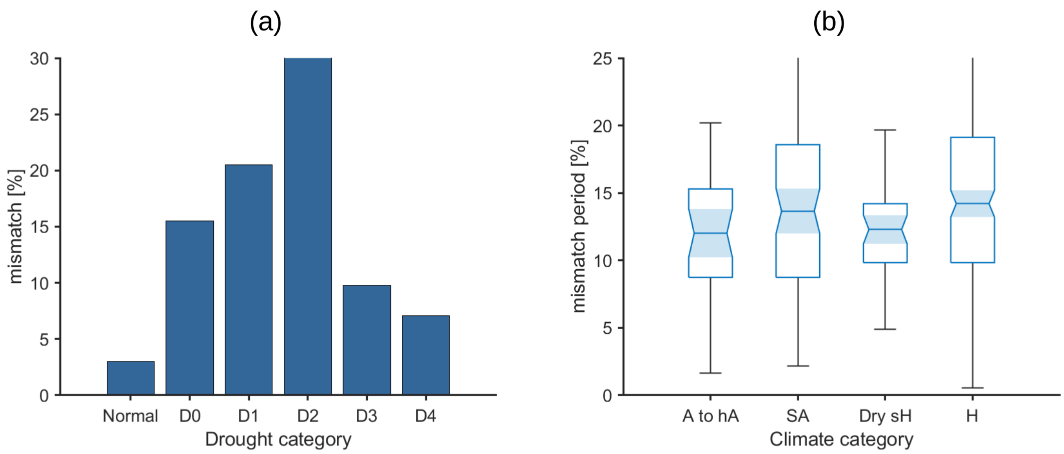

Figure 8. (a) A barplot illustrates the percentage of epochs where $PSDI_{max}$ diverges from SDI by at least two drought categories. (b) Boxplot of the mismatch between the $PSDI_{max}$ and SDI over different climate categories, namely, arid to hyper-arid (A to hA), semi-arid (SA), dry sub-humid (Dry sH), and humid (H). It is noteworthy that to count the number of months, we have considered those with more than one category difference between the $PSDI_{max}$ and SDI.

3.2 Performance of the PSDI during extreme hydrologic events

To assess the PSDI's reliability, we analyzed its performance during several well-documented extreme hydrologic events between 2002 and 2016. The drought events during 2012 included the moderate to exceptional drought over the United state (Boyer et al., 2013; Ault et al., 2013), southern Europe (Oikonomou et al., 2020; Spinoni et al., 2015). The drought affected many Middle East regions between 2007 and 2008 (Barlow et al., 2016). Southern Africa suffered from a severe to exceptional drought between 2005 and early 2006 (Nicholson, 2014), while central Argentina and Paraguay were affected by drought throughout 2009 (Guha-Sapir et al., 2016). Moreover, Australia experienced the worst drought recorded since European settlement in the 2000s, called the *Millennium drought*, with a peak in 2006 that affected many regions of the south to the east, including agricultural lands of the Murray-Darling basin (Van Dijk et al., 2013; Heberger, 2012). Figure 9 illustrates the performance of the PSDI over the events mentioned above. For each region, the category with the maximum probability and the estimated probability is shown for the selected date. Generally, the PSDI shows high performance in characterizing drought in the selected drought events (Figure 9). Comparing the SDI with PSDI_{max} reveals that SDI categorizes higher drought intensities.

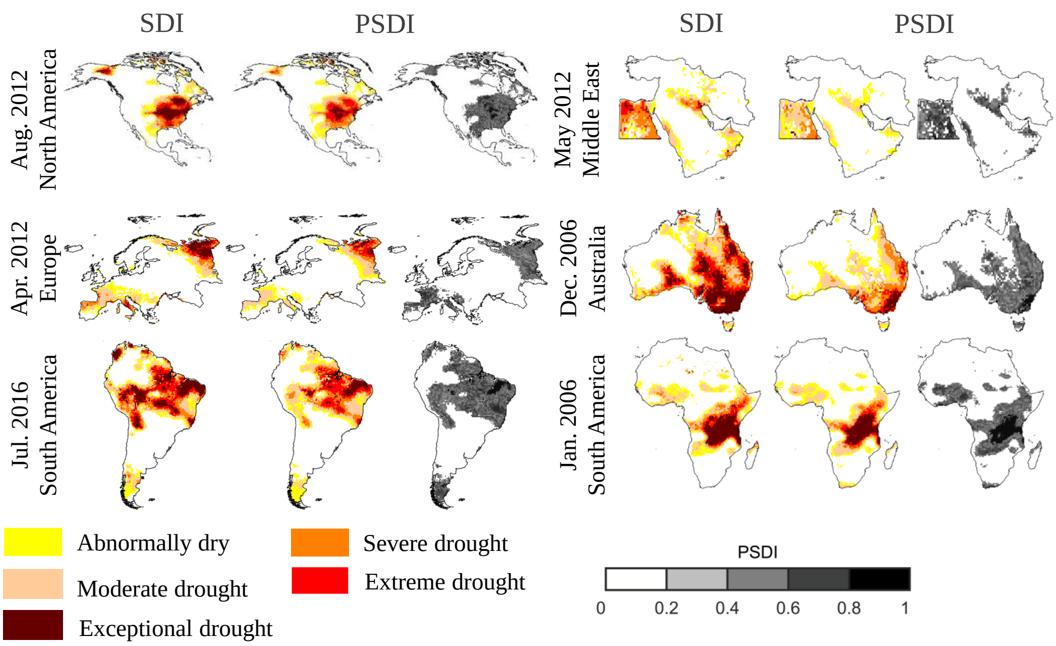

Figure 9. Comparing SDI with PSDI during some reported drought events.

4 Conclusions

For the first time, this study presents a probabilistic approach to characterizing TWS drought using time-variable gravity from satellite gravimetry. Our proposed framework acknowledges and addresses the inherent uncertainties associated with GRACE data. Our approach leverages Monte Carlo simulations to generate realistic realizations, capturing the stochastic nature of the TWSA time series. This ensemble reflects the diverse possible scenarios and their associated uncertainties, paving the way for a more insightful understanding of drought conditions. We have monitored the results of the proposed PSDI over major river basins and compared the result with SDI (deterministic approach). Our spatial analysis underscores the significance of adopting a probabilistic approach. It becomes evident that deterministic methodologies, in certain regions, tend to overestimate the severity of storage-based drought, potentially leading to misleading conclusions. While deterministic indices may tend to oversimplify drought categorization, PSDI accounts for uncertainty, thereby offering a more accurate representation of drought severity, particularly during extreme events.

Furthermore, our study assesses the performance of PSDI during well-documented extreme hydrologic events, spanning from the United States to Europe, the Middle East, Southern Africa, South America, and Australia. In each case, PSDI demonstrates its robustness in characterizing drought conditions. Comparing the SDI with $\mathrm{PSDI}_{\mathrm{max}}$ reveals that the drought can be categorized with more intensity using SDI with respect to the PSDI. We also address the uncertainties associated with different GRACE mascon products, emphasizing the importance of selecting the appropriate data source for reliable drought characterization. Variations in uncertainty estimates among different centers and processing methods highlight the need for caution when utilizing GRACE-derived data for drought analysis. We also shed light on the formal errors associated with GRACE data, highlighting the overestimation and underestimation tendencies of various solutions. This insight serves as a valuable reference for researchers and institutions relying on GRACE data for drought monitoring and assessment.

The findings of this study underscore the importance of a probabilistic approach in characterizing drought over various regions and during several drought events. The new approach provides a more realistic characterization of drought by accounting for the uncertainties in the GRACE(-FO) TWSA data in contrast to the common deterministic approach. By embracing uncertainty and providing a comprehensive ensemble of drought scenarios, PSDI advances the field of drought assessment, offering improved accuracy and insight for decision-makers and researchers alike. In an era marked by changing climate patterns and increasing water stress, our probabilistic approach represents a significant step toward more effective drought management and adaptation strategies.

Author Contribution Statement

Peyman Saemian and Mohammad J. Tourian developed the method, conducted the data analysis, and wrote the paper. Omid Elmi contributed to the analysis and assisted in producing graphics. Amir Aghakouchak and Nico Sneeuw supported the study with discussions on algorithm development. All authors commented on and reviewed the manuscript, and contributed to the final version.

Open Research

In this study, we employed a diverse set of datasets. The GRACE data, the GSFC mascon product, is available at `https://earth.gsfc.nasa.gov/geo/data/grace-mascons`. Nine global water resources datasets, including PCR-GLOBWB, SURFEX-TRIP, HBV-SIMREG, HTESSEL, JULES, LISFLOOD, ORCHIDEE, SWBM, and W3RA, were obtained from the eartH2Observe Water Cycle Integrator (`ftp://wci.earth2observe.eu`). CLM5 products are accessible via Earth System Grid (Oleson et al., 2019). The WaterGAP Global Hydrology Model (WaterGAP v2.2d) data is accessible at `https://doi.pangaea`

.de/10.1594/PANGAEA.918447. Additionally, the fifth generation ECMWF atmospheric reanalysis (ERA5) data can be downloaded from the Copernicus Climate Change Service (C3S) at ECMWF (https://cds.climate.copernicus.eu). For the long-term TWSA from the MLR approach, the data is available in mat format at DaRUS "Data for: A probabilistic approach to characterizing drought using satellite gravimetry", https://doi.org/10.18419/darus-3832.

Acknowledgments

We extend our gratitude to the numerous organizations and data providers whose contributions were essential to the success of this research. The datasets utilized in this study were sourced from various centers and repositories, as comprehensively listed in Tables 1 and 2 of the Supplementary section 1. We acknowledge the invaluable support and data access made possible by these sources. P. Saemian is funded by the DFG funding in project RASLyBoCa within the Special Priority Program Regional Sea Level Change & Society 'SeaLevel', Germany (SPP1889)—Project number 313883668, Grant TO 1042/3-1. O. Elmi is funded by DFG (Deutsche Forschungsgemeinschaft) (project Number: 324641997) within the framework of the Research Unit 2630, GlobalCDA: understanding the global freshwater system by combining geodetic and remote sensing information with modeling using a calibration/data assimilation approach (https://globalcda.de).

References

AghaKouchak, A., Farahmand, A., Melton, F., Teixeira, J., Anderson, M., Wardlow, B., & Hain, C. (2015). Remote sensing of drought: progress, challenges and opportunities. *Reviews of Geophysics*, *53*(2), 452–480. (doi: 10.1002/2014RG000456)

AghaKouchak, A., Huning, L. S., Sadegh, M., Qin, Y., Markonis, Y., Vahedifard, F., ... others (2023). Toward impact-based monitoring of drought and its cascading hazards. *Nature Reviews Earth & Environment*, 1–14.

AghaKouchak, A., Mirchi, A., Madani, K., Di Baldassarre, G., Nazemi, A., Alborzi, A., ...

others (2021). *Anthropogenic Drought: Definition, Challenges, and Opportunities.* Wiley Online Library. doi: 10.1029/2019RG000683

Ahmadalipour, A., Moradkhani, H., Yan, H., & Zarekarizi, M. (2017). Remote sensing of drought: vegetation, soil moisture, and data assimilation. In *Remote sensing of hydrological extremes* (pp. 121–149). Springer.

Ault, T., Henebry, G., De Beurs, K., Schwartz, M., Betancourt, J. L., & Moore, D. (2013). The false spring of 2012, earliest in North American record. *Eos, Transactions American Geophysical Union*, *94*(20), 181–182. doi: 10.1002/2013EO200001

Barlow, M., Zaitchik, B., Paz, S., Black, E., Evans, J., & Hoell, A. (2016). A review of drought in the Middle East and southwest Asia. *Journal of climate*, *29*(23), 8547–8574. doi: 10.1175/JCLI-D-13-00692.1

Behling, R., Roessner, S., Foerster, S., Saemian, P., Tourian, M. J., Portele, T. C., & Lorenz, C. (2022). Interrelations of vegetation growth and water scarcity in Iran revealed by satellite time series. *Scientific Reports*, *12*(1), 20784. doi: 10.1038/s41598-022-24712-6

Box, G. E., & Muller, M. E. (1958). A note on the generation of random normal deviates. *The annals of mathematical statistics*, *29*(2), 610–611. doi: 10.1214/aoms/1177706645

Boyer, J., Byrne, P., Cassman, K., Cooper, M., Delmer, D., Greene, T., … others (2013). The US drought of 2012 in perspective: a call to action. *Global Food Security*, *2*(3), 139–143. doi: 10.1016/j.gfs.2013.08.002

Coelho, C. A., & Goddard, L. (2009). El niño–induced tropical droughts in climate change projections. *Journal of Climate*, *22*(23), 6456–6476. doi: 10.1175/2009JCLI3185.1

Coumou, D., & Rahmstorf, S. (2012). A decade of weather extremes. *Nature climate change*, *2*(7), 491. doi: 10.1038/nclimate1452

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. *Nature Climate Change*, *6*(5), 508. doi: 10.1038/nclimate2941

Famiglietti, J. S., Lo, M., Ho, S. L., Bethune, J., Anderson, K., Syed, T. H., … Rodell, M. (2011). Satellites measure recent rates of groundwater depletion in California's Central

Valley. *Geophysical Research Letters*, *38*(3). doi: 10.1029/2010GL046442

FAO. (2021). Drought [Computer software manual]. (`http://www.fao.org/land-water/water/drought/en/`)

Gardner, A. S., Moholdt, G., Cogley, J. G., Wouters, B., Arendt, A. A., Wahr, J., ... others (2013). A reconciled estimate of glacier contributions to sea level rise: 2003 to 2009. *science*, *340*(6134), 852–857. doi: 10.1126/science.1234532

Gouweleeuw, B. T., Kvas, A., Gruber, C., Gain, A. K., Mayer-Gürr, T., Flechtner, F., & Güntner, A. (2018). Daily GRACE gravity field solutions track major flood events in the Ganges–Brahmaputra Delta. *Hydrology and Earth System Sciences*, *22*(5), 2867–2880. doi: 10.5194/hess-22-2867-2018

Guha-Sapir, D., Below, R., & Hoyois, P. (2016). EM-DAT: the CRED/OFDA international disaster database. *Université Catholique de Louvain*. Retrieved from `https://www.emdat.be/`

Heberger, M. (2012). Australia's millennium drought: Impacts and responses. In *The world's water* (pp. 97–125). Springer.

Heim Jr, R. R. (2002). A review of twentieth-century drought indices used in the united states. *Bulletin of the American Meteorological Society*, *83*(8), 1149–1166. doi: 10.1175/1520-0477-83.8.1149

Hisdal, H., Stahl, K., Tallaksen, L. M., & Demuth, S. (2001). Have streamflow droughts in Europe become more severe or frequent? *International Journal of Climatology*, *21*(3), 317–333. doi: 10.1002/joc.619

Hosseini-Moghari, S.-M., Araghinejad, S., Ebrahimi, K., Tang, Q., & AghaKouchak, A. (2020). Using GRACE satellite observations for separating meteorological variability from anthropogenic impacts on water availability. *Scientific Reports*, *10*(1), 15098. doi: 10.1038/s41598-020-71837-7

Hulme, M. (1992). A 1951–80 global land precipitation climatology for the evaluation of general circulation models. *Climate dynamics*, *7*, 57–72. doi: 10.1007/BF00209609

Jiang, D., Wang, J., Huang, Y., Zhou, K., Ding, X., & Fu, J. (2014). The review of GRACE data applications in terrestrial hydrology monitoring. *Advances in Meteorology*, *2014*. doi: 10.1155/2014/725131

Jones, P. D., & Hulme, M. (1996). Calculating regional climatic time series for temperature and precipitation: methods and illustrations. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, *16*(4), 361–377. doi: 10.1002/(SICI)1097-0088(199604)16:4⟨361::AID-JOC53⟩3.0.CO;2-F

Kreibich, H., Van Loon, A. F., Schröter, K., Ward, P. J., Mazzoleni, M., Sairam, N., ... others (2022). The challenge of unprecedented floods and droughts in risk management. *Nature*, *608*(7921), 80–86. doi: 10.1038/s41586-022-04917-5

Kvas, A., Behzadpour, S., Ellmer, M., Klinger, B., Strasser, S., Zehentner, N., & Mayer-Gürr, T. (2019). ITSG-Grace2018: Overview and evaluation of a new GRACE-only gravity field time series. *Journal of Geophysical Research: Solid Earth*, *124*(8), 9332–9344. doi: 10.1029/2019JB017415

Li, W., Pan, R., Jiang, Z., Chen, Y., Li, L., Luo, J.-J., ... Yu, J. (2021). Future changes in the frequency of extreme droughts over China based on two large ensemble simulations. *Journal of Climate*, *34*(14), 6023–6035.

Lorenz, C., Kunstmann, H., Devaraju, B., Tourian, M. J., Sneeuw, N., & Riegger, J. (2014). Large-scale runoff from landmasses: a global assessment of the closure of the hydrological and atmospheric water balances. *Journal of Hydrometeorology*, *15*(6), 2111–2139. doi: 10.1175/JHM-D-13-0157.1

Metropolis, N., & Ulam, S. (1949). The Monte Carlo method. *Journal of the American statistical association*, *44*(247), 335–341.

Mishra, A. K., & Singh, V. P. (2010). A review of drought concepts. *Journal of hydrology*, *391*(1-2), 202–216.

Mishra, A. K., & Singh, V. P. (2011). Drought modeling–a review. *Journal of Hydrology*, *403*(1-2), 157–175.

Mooney, C. Z. (1997). *Monte Carlo simulation* (Vol. 116). Sage publications.

Narasimhan, B., & Srinivasan, R. (2005). Development and evaluation of Soil Moisture Deficit Index (SMDI) and Evapotranspiration Deficit Index (ETDI) for agricultural drought monitoring. *Agricultural and Forest Meteorology*, *133*(1-4), 69–88. doi: 10.1016/j.agrformet.2005.07.012

Naumann, G., Cammalleri, C., Mentaschi, L., & Feyen, L. (2021). Increased economic drought impacts in Europe with anthropogenic warming. *Nature Climate Change*, *11*(6), 485–491. doi: 10.1038/s41558-021-01044-3

Nicholson, S. E. (2014). A detailed look at the recent drought situation in the Greater Horn of Africa. *Journal of Arid Environments*, *103*, 71–79. doi: 10.1016/j.jaridenv.2013.12.003

Oikonomou, P. D., Karavitis, C. A., Tsesmelis, D. E., Kolokytha, E., & Maia, R. (2020). Drought characteristics assessment in Europe over the Past 50 Years. *Water Resources Management*, *34*(15), 4757–4772. doi: 10.1007/s11269-020-02688-0

Oleson, K., Lawrence, D., Lombardozzi, D., & Wieder, D. (2019). *CLM land-only release.* (Accessed: 2021-05-30) doi: https://doi.org/10.5065/d6154fwh

Palmer, W. C. (1965). *Meteorological drought* (Vol. 30). US Department of Commerce, Weather Bureau.

Reager, J. T., & Famiglietti, J. S. (2009). Global terrestrial water storage capacity and flood potential using GRACE. *Geophysical research letters*, *36*(23). doi: 10.1029/2009GL040826

Rodell, M., Chen, J., Kato, H., Famiglietti, J. S., Nigro, J., & Wilson, C. R. (2007). Estimating groundwater storage changes in the Mississippi River basin (USA) using GRACE. *Hydrogeology Journal*, *15*(1), 159–166. doi: 0.1007/s10040-006-0103-7

Rodell, M., Velicogna, I., & Famiglietti, J. S. (2009). Satellite-based estimates of groundwater depletion in India. *Nature*, *460*(7258), 999–1002. doi: 10.1038/nature08238

Saemian, P., Elmi, O., Vishwakarma, B., Tourian, M., & Sneeuw, N. (2020). Analyzing the Lake Urmia restoration progress using ground-based and spaceborne observations. *Science of The Total Environment*, *739*, 139857. doi: 10.1016/j.scitotenv.2020.139857

Saemian, P., Tourian, M. J., AghaKouchak, A., Madani, K., & Sneeuw, N. (2022). How much water did Iran lose over the last two decades? *Journal of Hydrology: Regional*

Studies, *41*, 101095. doi: 10.1016/j.ejrh.2022.101095

Scholz, C. A., Johnson, T. C., Cohen, A. S., King, J. W., Peck, J. A., Overpeck, J. T., ... others (2007). East african megadroughts between 135 and 75 thousand years ago and bearing on early-modern human origins. *Proceedings of the National Academy of Sciences*, *104*(42), 16416–16421.

Shepherd, A., Ivins, E., Rignot, E., Smith, B., Van Den Broeke, M., Velicogna, I., ... others (2018). Mass balance of the Antarctic Ice Sheet from 1992 to 2017. *Nature*, *558*, 219–222. doi: 10.1038/s41586-018-0179-y

Sinha, D., Syed, T. H., Famiglietti, J. S., Reager, J. T., & Thomas, R. C. (2017). Characterizing drought in India using GRACE observations of terrestrial water storage deficit. *Journal of Hydrometeorology*, *18*(2), 381–396. doi: 10.1175/JHM-D-16-0047.1

Sohn, B., Yeh, S.-W., Schmetz, J., & Song, H.-J. (2013). Observational evidences of Walker circulation change over the last 30 years contrasting with GCM results. *Climate Dynamics*, *40*, 1721–1732. doi: 10.1007/s00382-012-1484-z

Spinoni, J., Naumann, G., Vogt, J. V., & Barbosa, P. (2015). The biggest drought events in Europe from 1950 to 2012. *Journal of Hydrology: Regional Studies*, *3*, 509–524. doi: 10.1016/j.ejrh.2015.01.001

Sun, A. Y., Scanlon, B. R., AghaKouchak, A., & Zhang, Z. (2017). Using GRACE satellite gravimetry for assessing large-scale hydrologic extremes. *Remote Sensing*, *9*(12), 1287. doi: 10.3390/rs9121287

Svoboda, M., Fuchs, B., et al. (2016). *Handbook of drought indicators and indices.* World Meteorological Organization Geneva, Switzerland.

Svoboda, M., Hayes, M., & Wood, D. (2012). *Standardized precipitation index: user guide* (Tech. Rep. No. NO. 1090).

Svoboda, M., LeComte, D., Hayes, M., Heim, R., Gleason, K., Angel, J., ... others (2002). The drought monitor. *Bulletin of the American Meteorological Society*, *83*(8), 1181–1190.

Tapley, B. D., Bettadpur, S., Watkins, M., & Reigber, C. (2004). The gravity recovery and climate experiment: Mission overview and early results. *Geophysical research letters*, *31*(9). doi: 10.1029/2004GL019920

Teuling, A. J. (2018). A hot future for european droughts. *Nature Climate Change*, *8*(5), 364–365.

Thomas, A. C., Reager, J. T., Famiglietti, J. S., & Rodell, M. (2014). A GRACE-based water storage deficit approach for hydrological drought characterization. *Geophysical Research Letters*, *41*(5), 1537–1545. doi: 10.1002/2014GL059323

Tourian, M., Reager, J., & Sneeuw, N. (2018). The total drainable water storage of the Amazon river basin: A first estimate using GRACE. *Water Resources Research*, *54*(5), 3290–3312. doi: 10.1029/2017WR021674

Tourian, M. J., Elmi, O., Shafaghi, Y., Behnia, S., Saemian, P., Schlesinger, R., & Sneeuw, N. (2022). Hydrosat: geometric quantities of the global water cycle from geodetic satellites. *Earth System Science Data*, *14*(5), 2463–2486.

Tourian, M. J., Saemian, P., Ferreira, V. G., Sneeuw, N., Frappart, F., & Papa, F. (2023). A copula-supported Bayesian framework for spatial downscaling of GRACE-derived terrestrial water storage flux. *Remote Sensing of Environment*, *295*, 113685. doi: 10.1016/j.rse.2023.113685

van den Broeke, M., Bamber, J., Ettema, J., Rignot, E., Schrama, E., van de Berg, W. J., ... Wouters, B. (2009). Partitioning recent Greenland mass loss. *science*, *326*(5955), 984–986. doi: 10.1126/science.1178176

Van Dijk, A. I., Beck, H. E., Crosbie, R. S., de Jeu, R. A., Liu, Y. Y., Podger, G. M., ... Viney, N. R. (2013). The Millennium Drought in southeast Australia (2001–2009): Natural and human causes and implications for water resources, ecosystems, economy, and society. *Water Resources Research*, *49*(2), 1040–1057. doi: 10.1002/wrcr.20123

Voss, K. A., Famiglietti, J. S., Lo, M., De Linage, C., Rodell, M., & Swenson, S. C. (2013). Groundwater depletion in the Middle East from GRACE with implications for transboundary water management in the Tigris-Euphrates-Western Iran region. *Water resources research*, *49*(2), 904–914. doi: 10.1002/wrcr.20078

Wilhite, D. A. (2000). Chapter 1 drought as a natural hazard: Concepts and definitions.

Drought Mitigation Center Faculty Publications.

Wilhite, D. A., Svoboda, M. D., & Hayes, M. J. (2007). Understanding the complex impacts of drought: A key to enhancing drought mitigation and preparedness. *Water resources management*, *21*(5), 763–774. doi: 10.1007/s11269-006-9076-5

Yi, H., & Wen, L. (2016). Satellite gravity measurement monitoring terrestrial water storage change and drought in the continental United States. *Scientific reports*, *6*, 19909. doi: 10.1038/srep19909

Yi, S., Saemian, P., Sneeuw, N., & Tourian, M. J. (2023). Estimating runoff from pan-Arctic drainage basins for 2002–2019 using an improved runoff-storage relationship. *Remote Sensing of Environment*, *298*, 113816. doi: 10.1016/j.rse.2023.113816

Yirdaw, S. Z., Snelgrove, K. R., & Agboma, C. O. (2008). GRACE satellite observations of terrestrial moisture changes for drought characterization in the Canadian Prairie. *Journal of Hydrology*, *356*(1-2), 84–92. doi: 10.1016/j.jhydrol.2008.04.004

Yu, M., Li, Q., Hayes, M. J., Svoboda, M. D., & Heim, R. R. (2014). Are droughts becoming more frequent or severe in China based on the standardized precipitation evapotranspiration index: 1951–2010? *International Journal of Climatology*, *34*(3), 545–558. doi: 10.1002/joc.3701

Zargar, A., Sadiq, R., Naser, B., & Khan, F. I. (2011). A review of drought indices. *Environmental Reviews*, *19*(NA), 333–349. doi: doi.org/10.1139/a11-013

Zhao, C., Brissette, F., Chen, J., & Martel, J.-L. (2020). Frequency change of future extreme summer meteorological and hydrological droughts over north america. *Journal of Hydrology*, *584*, 124316.

Zhao, M., Geruo, A., Velicogna, I., & Kimball, J. S. (2017). A global gridded dataset of GRACE drought severity index for 2002–14: Comparison with PDSI and SPEI and a case study of the Australia millennium drought. *Journal of Hydrometeorology*, *18*(8), 2117–2129. doi: 10.1175/JHM-D-16-0182.1

Zhong, Y., Hu, E., Wu, Y., An, Q., Wang, C., Bai, H., & Gao, W. (2023). Reconstructing a long-term water storage-based drought index in the yangtze river basin. *Science of The*

Total Environment, *883*, 163403.

Zscheischler, J., Westra, S., Van Den Hurk, B. J., Seneviratne, S. I., Ward, P. J., Pitman, A., . . . others (2018). Future climate risk from compound events. *Nature Climate Change*, *8*(6), 469–477. doi: 10.1038/s41558-018-0156-3

Supplementary Material for: A probabilistic approach to characterizing drought using satellite gravimetry

Peyman Saemian[1], Mohammad J. Tourian[1], Omid Elmi[1]
Nico Sneeuw[1], Amir AghaKouchak[2,3]

[1]Institute of Geodesy, University of Stuttgart, Stuttgart, Germany
[1]Department of Civil and Environmental Engineering, University of California, Irvine, CA, USA.
[3]United Nations University Institute for Water, Environment and Health, Hamilton, Ontario, Canada

The supplementary material includes:

1 TWSA from GRACE observations

Two main approaches have been developed to process GRACE range-rate observations. In the first approach, the Earth's gravity field is parameterized using the global Spherical Harmonics (SHs) basis functions (see Wahr et al. (1998) for details). Within the past couple of years, an alternative approach for processing GRACE level 1 (L1) has been proposed which considers parameterizing with regional mass concentration functions (mascons) (Watkins et al., 2015a; Scanlon et al., 2016). In this study, we have used the latest version (version 2) of the Goddard Space Flight Center (GSFC) which can be accessed via `https://earth.gsfc.nasa.gov/geo/data/grace-mascons`. We have compared the uncertainty estimation from GSFC with the Jet Propulsion Laboratory (JPL) mascon solutions. The latest version (Release 6.1 Version 03) of the JPL mascon solutions used in the comparison can be obtained from `https://podaac.jpl.nasa.gov/dataset/TELLUS_GRAC-GRFO_MASCON_CRI_GRID_RL06.1_V3`. Moreover, we have compared the error estimation in the level-2 products, also known as formal errors, in Figure 1. Table S1 and Table S2 list all the mascons and level-2 products of GRACE and GRACE-FO used in this study, respectively.

Table S1. List of centers which provide Level-3 TWSA from GRACE and GRACE-FO.

Product	Sensor(s)	Source/Reference
GSFC v02 mascons	GRACE/GRACE-FO	Loomis et al. (2019)
JPL RL06.1 v03 L3 mascons	GRACE/GRACE-FO	Landerer et al. (2020); D. Wiese et al. (2018); Watkins et al. (2015b); D. N. Wiese et al. (2016)

The mascons products, like the one used in this study, estimate the uncertainty in the TWSA estimation, either in the form of spherical harmonics or global grids. Despite the same level-1 product, the errors in the mascons approaches vary among different centers, as they would use different processing approaches and background models. Figure S1 illustrates a spatio-temporal comparison comparison between two widely used mascons datasets, namely, JPL RL06-v02 and GSFC RL06-v02. the Figure S1 (a) shows the mean TWSA uncertainty from the above products from April 2002 to November 2022. The highest values belong to Greenland, the Amazonas, the Indian sub-continent, and the northwest of Canada. The Figure S1 (b) compares the time series of the global land averaged TWSA uncertainty from April 2002 to November 2022. The time series shows a sharp pick in 2015, followed by a positive trend related to the battery failure (Save, 2016; Mayer-Gürr et al., 2018; Bandikova et al., 2019). The two mascon solutions exhibit consistent uncertainty estimates ($\overline{\sigma} = 2.5\,\mathrm{cm}$) throughout the GRACE observation period, except for the initial year (April 2002 to June 2003). The elevated uncertainties in JPL solutions from April 2002 to June 2003, as well as at the last year of the GRACE-FO mission, stem from the application of a Kalman filter in the solution methodology, facilitating the temporal connection of adjacent months (D. Wiese et al., 2016). Notably, during the GRACE-FO mission, GSFC's uncertainty values are significantly higher ($\overline{\sigma} = 3.2\,\mathrm{cm}$ for GSFC compared to $\overline{\sigma} = 1.8\,\mathrm{cm}$ for JPL).

Table S2. List of GRACE(-FO) Level-2 solutions.

Center	Product	Sensor(s)	Time span
Solutions that include GRACE and GRACE-FO			
CSR	CSR RL06	GRACE	200204–201706
	CSR RL06	GRACE-FO	201806–present
GFZ	GFZ RL06	GRACE	200204–201706
	CSR RL06 (GFO)	GRACE-FO	201806–present
JPL	JPL RL06	GRACE	200204–201706
	CSR RL06 (GFO)	GRACE-FO	201806–present
ITSG	ITSG-Grace2018	GRACE	200204–201706
	ITSG-Grace_op	GRACE-FO	201806–present
LUH	LUH-Grace2018	GRACE	200301–201603
	LUH-GRACE-FO-2020	GRACE-FO	201806–present
COST-G**	Grace	GRACE	200204–201706
	Grace-FO	GRACE-FO	201806–present
AIUB	AIUB-RL02	GRACE	200302–201403
	AIUB-GRACE-FO_op	GRACE-FO	201806–present
CNES	CNES_GRGS_RL05	GRACE & GRACE-FO	200209–present
Solutions that include only GRACE			
Tongji	Tongji-Grace2018	GRACE	200204–201608
HUST	HUST-Grace2020	GRACE	200301–201607
IGG	IGG-RL01	GRACE	200204–201607
SWJTU	SWJTU-GRACE-RL01	GRACE	200303–201110
SWPU	SWPU-GRACE2021	GRACE	200204–201705
WHU	WHU RL01	GRACE	200204–201607
XISM&SSTC	GRACE01	GRACE	200204-201603

2 Long-term TWSA dataset

In this study, we have used a combination of various models to estimate TWSA for the pre-GRACE era, back to 1980. Models, from a simple box model to a recent sophisticated deep learning model, have been designed to enhance our understanding and acuity of the Earth's water system that occurs as an exchange between the terrestrial biosphere and atmosphere. In general, three different groups of models have been developed, namely Land Surface Mod-

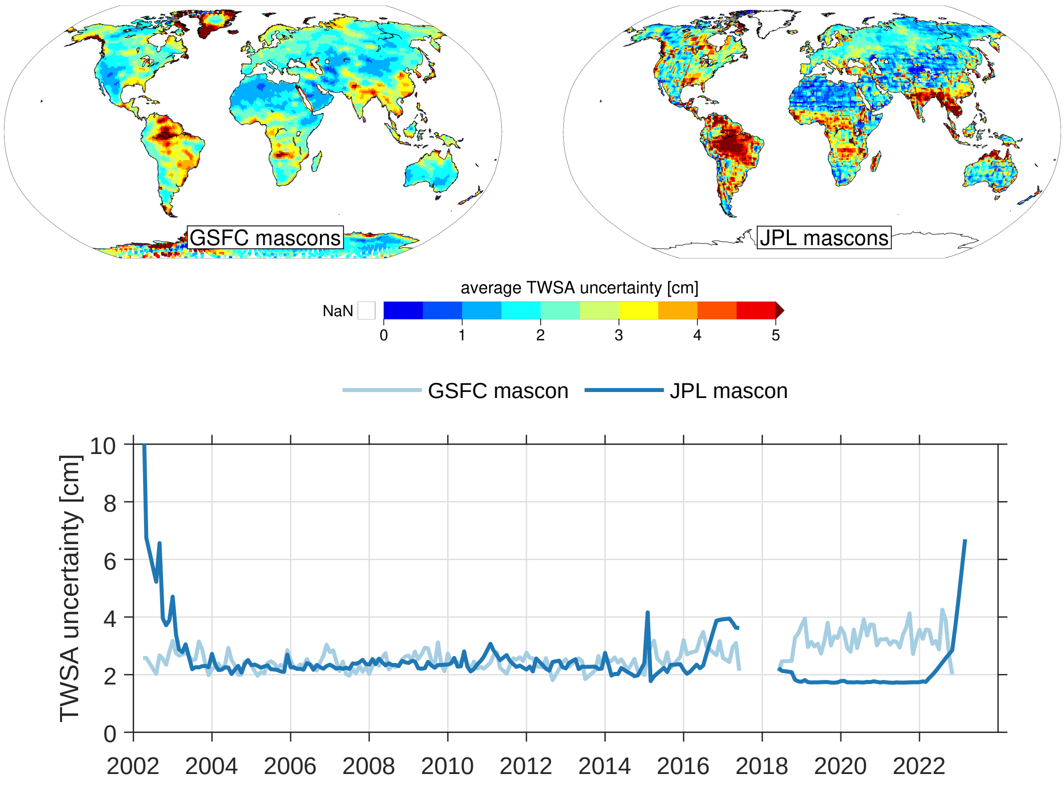

Figure S1. Top: Global distribution of the averaged TWSA uncertainty spanning from April 2002 to November 2022. Bottom: Time series of the global averaged TWSA uncertainty. The data is obtained from two distinct mascon datasets: JPL RL06-v02 and GSFC RL06-v02.

els (LSMs), Global Hydrological Models (GHMs), and global atmospheric reanalysis models. In this study, we have employed in total of 13 state-of-the-art datasets of Global Hydrological Models (GHMs), Land Surface Models (LSMs), and atmospheric reanalysis models (Table S3). Nine multi-decadal global water resources datasets were obtained from the eartH2Observe Water Cycle Integrator (WCI; `ftp://wci.earth2observe.eu` (last access: 31 May 2021)), including PCR-GLOBWB, SURFEX-TRIP, HBV-SIMREG, HTESSEL-CaMa, JULES, LISFLOOD, ORCHIDEE, SWBM, and W3RA. The output of these datasets is available at 0.5° spatial resolution over the period 1979–2012. Besides datasets from eartH2Observe, we have included the Community Land Model Version 5 (CLM5) with two standard forcing datasets, namely the Global Soil Wetness Project forcing data set (GSWP3) and CRUNCEP (the combination of the Climate Research Unit (CRU) and the National Centers for Environmental Prediction (NCEP)). The CLM5 datasets are at 0.5° spatial resolution covering the period 1901–2014 (for more detail about the CLM5 model, please see Lawrence et al. (2019)). The CLM5 products are accessible via Earth

System Grid (ESG) (Oleson et al., 2019). We have also included the latest version of the WaterGAP Global Hydrology Model (WaterGAP v2.2d) (Müller Schmied et al., 2021), covering the period 1901–2016 and at 0.5° spatial resolution. The outputs of the WaterGap v2.d are available at (`https://doi.pangaea.de/10.1594/PANGAEA.918447`). Finally, we have included the fifth generation ECMWF atmospheric reanalysis of the global climate (ERA5) at 0.25° spatial resolution which provides data from 1979 to the present. The data is downloaded from the Copernicus Climate Change Service (C3S) at ECMWF (`https://cds.climate.copernicus.eu`)(last access: 30 May 2021). TWSA from models carries a higher spatial resolution and therefore values with higher frequency. To set the same spectral content in models compared to GRACE TWSA, we have transferred the model outputs into the spectral domain and truncated the SHs to the maximum degree and order 96. Finally, we recovered the TWSA fields from the truncated SHs.

Table S3. Summary of global models used in this study. GHM: Global Hydrological Model; LSM: Land Surface Model; ReA: Reanalysis Model.

	Model	Time Period	Data Provider	Reference
GHM	WGHM	1901–2016	Goethe University Frankfurt	Müller Schmied et al. (2021)
	PCRGLOB-WB	1979–2012	Utrecht University (UU)	Wada et al. (2014)
				Sutanudjaja et al. (2018)
	HBV-SIMREG	1979–2012	Joint Research Centre (JRC)	Lindström et al. (1997)
	LISFLOOD	1979–2012	Joint Research Centre (JRC)	Van Der Knijff et al. (2010)
	W3RA	1979–2012	CSIRO**	Van Dijk (2010)
	SWBM	1979–2012	Simple Water Balance Model	Koster & Mahanama (2012)
				Orth & Seneviratne (2013)
LSM	CLM5	1940–2014	The Earth System Grid (ESG) at NCAR	Lawrence et al. (2019)
	HTESSEL	1979–2012	ECMWF	Balsamo et al. (2015)
	JULES	1979–2012	Centre for Ecology and Hydrology (CEH)	Best et al. (2011)
				Clark et al. (2011)
	ORCHIDEE	1979–2012	French National Centre for Scientific Research	Polcher et al. (2011)
	SURFEX-TRIP	1979–2012	Meteo France	Decharme et al. (2013)
ReA	ERA5	1979–2016	ECMWF*	Hersbach et al. (2020)

* ECMWF: European Centre for Medium-Range Weather Forecasts

** CSIRO: Commonwealth Scientific and Industrial Research Organisation

2.1 Multivariate Linear Regression

To combine models, we have used the Multivariate Linear Regression (MLR) method. MLR is a statistical method used for estimating the parameters of a linear regression model with multiple independent variables. MLR has several advantages, including its ability to handle multiple independent variables and to model complex relationships between variables. It also provides estimates of the coefficients and their standard errors, which can be used to test hypotheses and construct confidence intervals. However, MLR assumes that the errors are normally distributed and have constant variance, which may not always be true in practice. Additionally, it can be sensitive to outliers and multicollinearity among the independent variables. The basic idea behind MLR is to find the coefficients that minimize the sum of squared errors between the predicted and actual values of the dependent variable. The formula for MLR is as follows:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{1}$$

Here $\mathbf{y}$ is the vector of dependent variable values, $\mathbf{X}$ is the matrix of independent variable values, $\boldsymbol{\beta}$ is the vector of coefficients to be estimated, and $\boldsymbol{\epsilon}$ is the vector of errors, which are assumed to be normally distributed with mean zero and constant variance.

2.2 Compare with GRACE

To evaluate the performance of the long-term TWSA dataset from the MLR method ($\mathrm{TWSA}_{\mathrm{MLR}}$), we have compared the results with GRACE estimation within the GRACE era (April 2002 to December 2012).

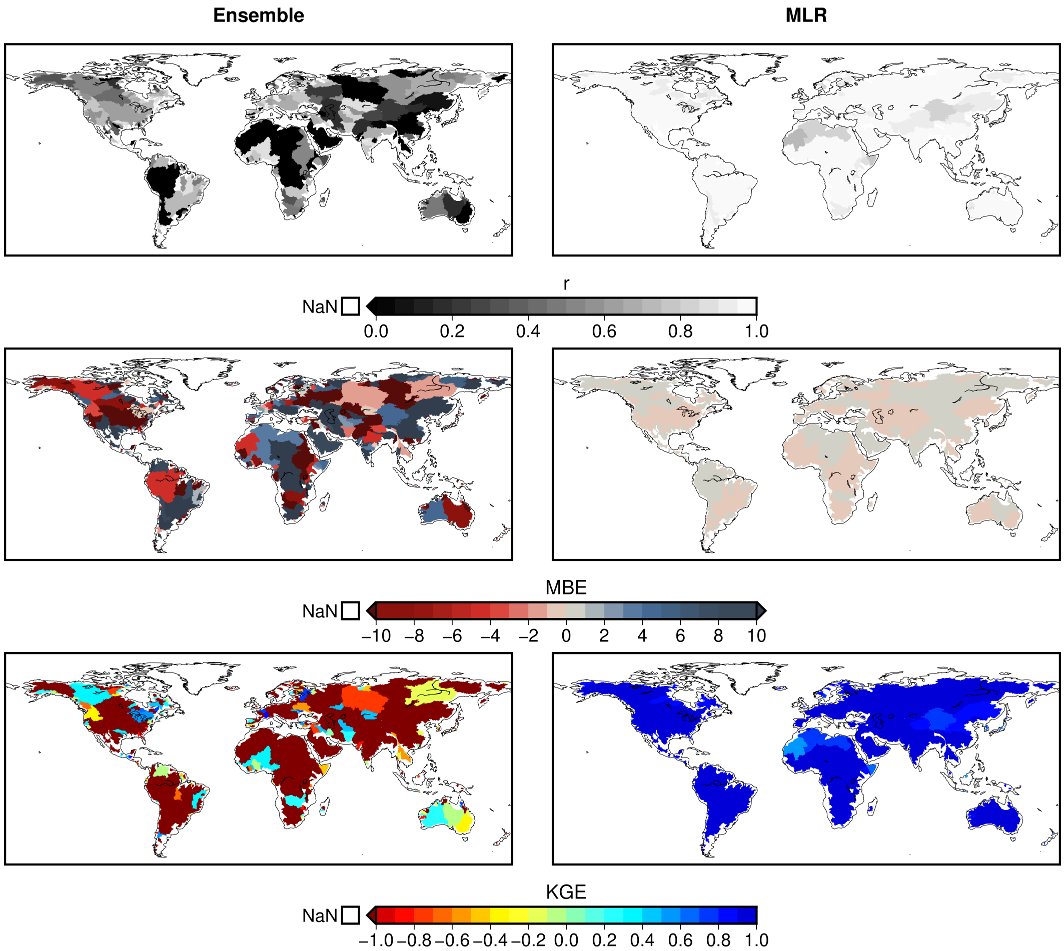

Figure S2. Global distribution of the correlation coefficient (r), Mean Bias Error (MBE), and Kling-Gupta Efficiency (KGE) values for major river basins (excluding Greenland and Antarctica) obtained from the comparison between the reconstructed TWSA from ensemble mean and Multivariate Linear Regression (MLR) and GRACE during 2003–2012.

3 Handling trends

Several studies have suggested that before investigating drought indices using the GRACE TWSA time series, detrending is necessary (e.g., Liu et al., 2020; Khorrami & Gunduz, 2021). Liu et al. (2020), for instance, have demonstrated that without detrending TWSA time series drought severity can be overestimated over some basins in China after 2013. While the soil moisture data suggests that the drought ceased in September 2014, their GRACE indices (GRACE-DSI) show a continuous drought condition. In contrast to the aforementioned studies, we deliberately retain the trend in the time series. We reason that the trend reflects long-term changes in climate, such as temperature increases or precipitation pattern alterations, which can affect the frequency and severity of droughts. Eliminating the trend would essentially omit these long-term changes from the analysis, providing an incomplete understanding of the hydrological system.

To demonstrate the impact of detrending, we calculated the TWSA time series in two real cases using the SSA approach with a 24-month window to remove the trend in the data. The two cases, the Tigris basin in the Middle East with a negative trend and the Niger basin in Africa with a positive trend are presented in Figure S3 and Figure S4, respectively. In each case, we compared the results from two scenarios: one without detrending, denoted by the solid line in (c) and (d) and labeled as (a), and one with detrending, shown as the dashed line in (c) and (e) and labeled as (b).

The Tigris basin experienced a prolonged period of water loss, particularly after 2007, which is apparent in the red area in Figure S3(d). Detrending the data resulted in higher values for the climatology compared to the non-detrended data, as shown in Figure S3(c), and caused oscillations between wet and dry years, as seen in Figure S3(e). On the other hand, the Niger basin exhibited a positive trend mainly after 2010, resulting in wetter years in the basin, as depicted in Figure S4(d). Although detrending did not significantly alter the climatology, as illustrated in Figure S4(c), it did reveal dry years after 2010, which is inconsistent with actual conditions.

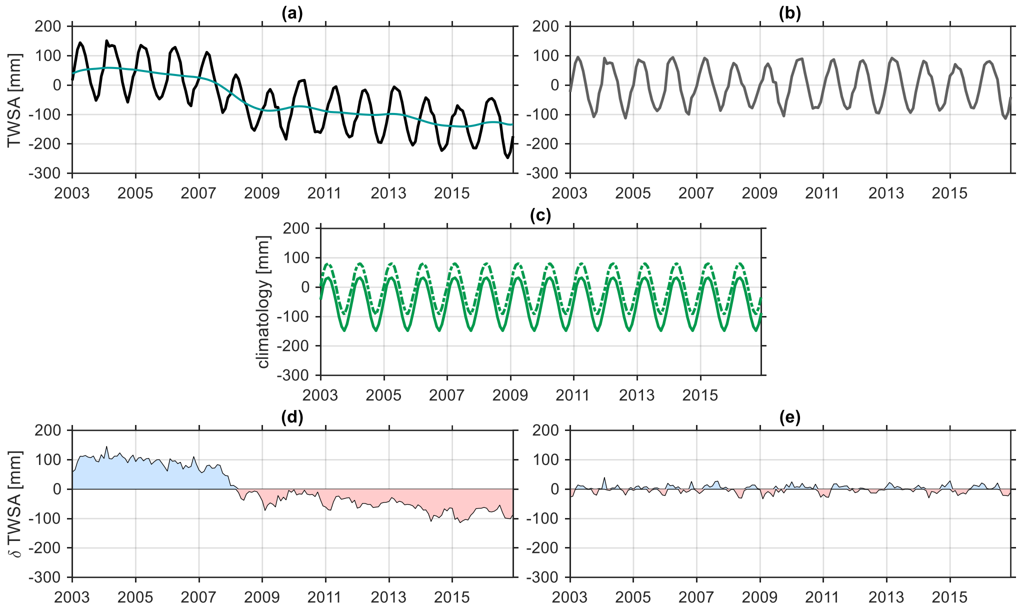

Figure S3. This figure presents a comprehensive analysis of TWSA for the Tigris River basin in the Middle East, using data from GRACE satellite mission. (a) shows the time series of TWSA from GRACE, along with its inter-annual variations which are extracted using the Singular Spectrum Analysis (SSA) approach with a 24-month window. (b) displays the TWSA after removing the inter-annual variations, highlighting the long-term trends. (c) illustrates the climatology of TWSA, which represents the long-term monthly mean. The solid and dashed lines represent the climatology obtained from (a) and (b), respectively. (d) and (e) show the TWSA residuals, obtained by subtracting the corresponding climatology from panels (a) and (b), respectively. These residual plots reveal the short-term fluctuations in TWSA that are not captured by the climatology.

4 Major river basins

In this study, we have presented and analyzed the results of the global major river basins. The border of the basins follows the HydroSHEDS database (`https://www.hydrosheds.org/`). Moreover, the climate of the basins is determined using the Aridity Index (AI), which is the ratio of total annual precipitation to potential evapotranspiration. To compute the aridity index, we have employed the latest version of the European Center for Medium-Range Weather Forecasts (ECMWF) Reanalysis (ERA), namely ERA5 (Hersbach et al., 2020). Based on AI, the climate of the basins can be categorized into humid (AI > 0.65), sub-humid (AI ≤ 0.65, and AI > 0.5), semi-arid (AI ≤ 0.5 and > 0.2), arid (AI ≤ 0.2 and > 0.05), and hyperarid (AI ≤ 0.05). This study grouped arid and hyper-arid into one group,

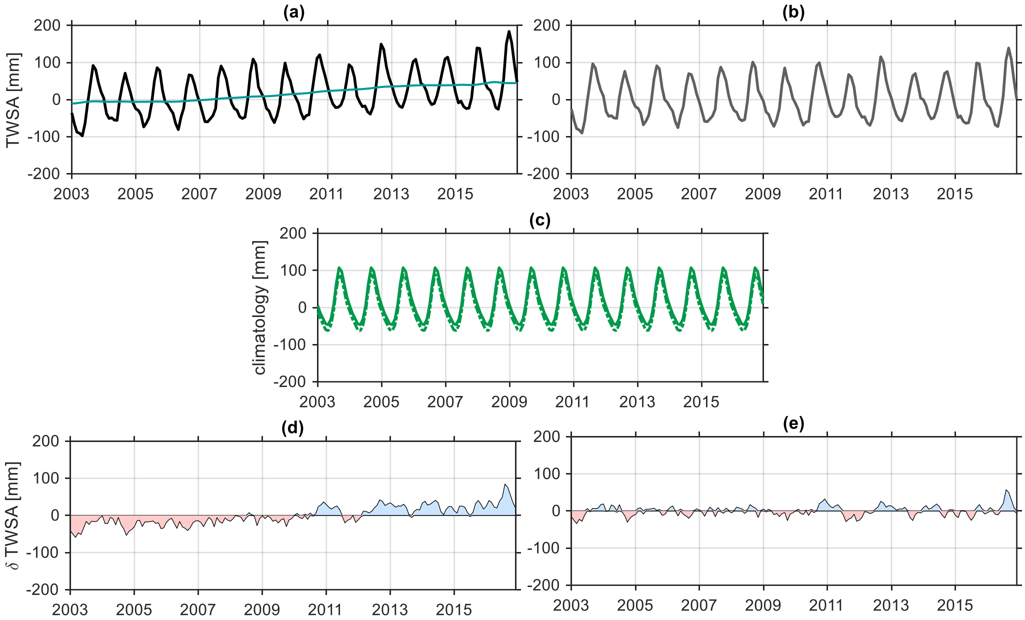

Figure S4. Same as Figure S3 but for Niger river basin in West Africa, flowing through 10 countries: Guinea, Mali, Niger, Benin, Burkina Faso, Cote d'Ivoire, Ghana, Togo, Cameroon, and Nigeria.

Arid-hyper Arid (Figure S5). Based on AI criteria, 60 % of the river basins are categorized as humid, ∼ 10 % as sub-humid, 22 d% as semiarid, and ∼ 8 % as arid to hyper-arid).

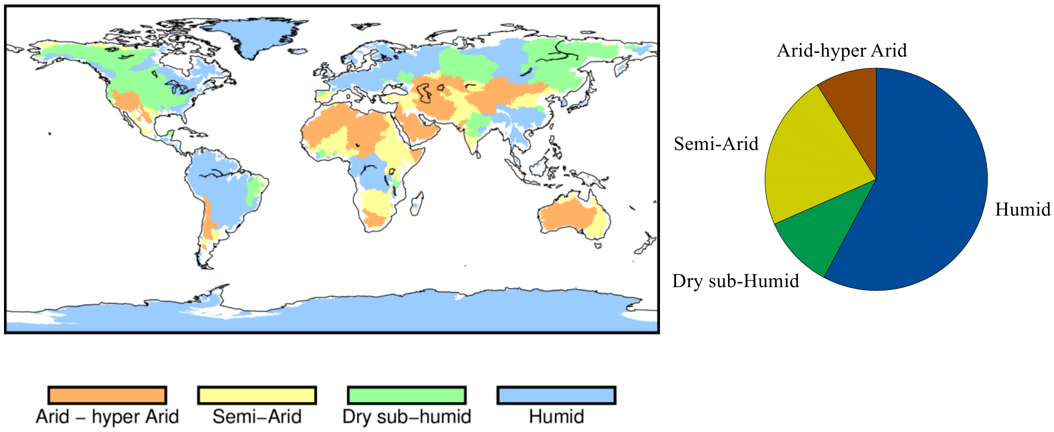

Figure S5. Global distribution of the major river basins together with their corresponding climate category. Besides, a pie chart illustrates the worldwide share of each category in terms of area.

References

Balsamo, G., Albergel, C., Beljaars, A., Boussetta, S., Brun, E., Cloke, H., ... others (2015). ERA-Interim/Land: a global land surface reanalysis data set. *Hydrology and Earth System Sciences*, *19*(1), 389–407. doi: 10.5194/hess-19-389-2015

Bandikova, T., McCullough, C., Kruizinga, G. L., Save, H., & Christophe, B. (2019). GRACE accelerometer data transplant. *Advances in Space Research*, *64*(3), 623–644.

Best, M., Pryor, M., Clark, D., Rooney, G., Essery, R., Ménard, C., ... others (2011). The Joint UK Land Environment Simulator (JULES), model description–Part 1: energy and water fluxes. *Geoscientific Model Development*, *4*(3), 677–699. doi: 10.5194/gmd-4-677-2011

Clark, D., Mercado, L., Sitch, S., Jones, C., Gedney, N., Best, M., ... others (2011). The Joint UK Land Environment Simulator (JULES), model description–Part 2: carbon fluxes and vegetation dynamics. *Geoscientific Model Development*, *4*(3), 701–722. doi: 10.5194/gmd-4-701-2011

Decharme, B., Martin, E., & Faroux, S. (2013). Reconciling soil thermal and hydrological lower boundary conditions in land surface models. *Journal of Geophysical Research: Atmospheres*, *118*(14), 7819–7834. doi: 10.1002/jgrd.50631

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., ... others (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological*

Society, *146*(730), 1999–2049. doi: 10.1002/qj.3803

Khorrami, B., & Gunduz, O. (2021). An enhanced water storage deficit index (EWSDI) for drought detection using GRACE gravity estimates. *Journal of Hydrology*, *603*, 126812. doi: 10.1016/j.jhydrol.2021.126812

Koster, R. D., & Mahanama, S. P. (2012). Land surface controls on hydroclimatic means and variability. *Journal of Hydrometeorology*, *13*(5), 1604–1620. doi: 10.1175/JHM-D-12-050.1

Landerer, F., Flechtner, F. M., Save, H., Webb, F. H., Bandikova, T., Bertiger, W. I., ... others (2020). Extending the global mass change data record: GRACE Follow-On instrument and science data performance. *Geophysical Research Letters*, *47*(12), e2020GL088306. doi: 10.1029/2020GL088306

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., ... others (2019). The Community Land Model version 5: Description of new features, benchmarking, and impact of forcing uncertainty. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4245–4287. doi: 10.1029/2018MS001583

Lindström, G., Johansson, B., Persson, M., Gardelin, M., & Bergström, S. (1997). Development and test of the distributed HBV-96 hydrological model. *Journal of hydrology*, *201*(1-4), 272–288. doi: 10.1016/S0022-1694(97)00041-3

Liu, X., Feng, X., Ciais, P., Fu, B., Hu, B., & Sun, Z. (2020). GRACE satellite-based drought index indicating increased impact of drought over major basins in China during 2002–2017. *Agricultural and Forest Meteorology*, *291*, 108057. doi: 10.1016/j.agrformet.2020.108057

Loomis, B., Luthcke, S., & Sabaka, T. (2019). Regularization and error characterization of GRACE mascons. *Journal of geodesy*, *93*(9), 1381–1398. doi: 10.1007/s00190-019-01252-y

Mayer-Gürr, T., Behzadpour, S., Kvas, A., Ellmer, M., Klinger, B., Strasser, S., & Zehentner, N. (2018). ITSG-Grace2018: Monthly, Daily and Static Gravity Field Solutions from GRACE. *Graz University of Technology*. doi: 10.5880/ICGEM.2018.003

Müller Schmied, H., Cáceres, D., Eisner, S., Flörke, M., Herbert, C., Niemann, C., ... others (2021). The global water resources and use model WaterGAP v2. 2d: Model description and evaluation. *Geoscientific Model Development*, *14*(2), 1037–1079. doi: 10.5194/gmd-14-1037-2021

Oleson, K., Lawrence, D., Lombardozzi, D., & Wieder, D. (2019). *CLM land-only release.* (Accessed: 2021-05-30) doi: https://doi.org/10.5065/d6154fwh

Orth, R., & Seneviratne, S. I. (2013). Predictability of soil moisture and streamflow on subseasonal timescales: A case study. *Journal of Geophysical Research: Atmospheres*, *118*(19), 10–963. doi: 10.1002/jgrd.50846,2013

Polcher, J., Bertrand, N., Biemans, H., Clark, D. B., Floerke, M., Gedney, N., ... Voss, F. (2011). Improvements in hydrological processes in general hydrological models and land surface models within WATCH.

Save, H. (2016). Grace battery status and science impact. In *Grace science team meeting, potsdam, germany oct* (pp. 5–7).

Scanlon, B. R., Zhang, Z., Save, H., Wiese, D. N., Landerer, F. W., Long, D., ... Chen, J. (2016). Global evaluation of new GRACE mascon products for hydrologic applications. *Water Resources Research*, *52*(12), 9412–9429. doi: 10.1002/2016WR019494

Sutanudjaja, E. H., Van Beek, R., Wanders, N., Wada, Y., Bosmans, J. H., Drost, N., ... others (2018). PCR-GLOBWB 2: a 5 arcmin global hydrological and water resources model. *Geoscientific Model Development*, *11*(6), 2429–2453. doi: 10.5194/gmd-11-2429-2018

Van Der Knijff, J., Younis, J., & De Roo, A. (2010). LISFLOOD: a GIS-based distributed model for river basin scale water balance and flood simulation. *International Journal of Geographical Information Science*, *24*(2), 189–212. doi: 10.1080/13658810802549154

Van Dijk, A. (2010). The Australian Water Resources Assessment System Technical Report 3. Landscape Model (version 0.5) Technical Description. doi: 10.4225/08/5852dd9bb578c

Wada, Y., Wisser, D., & Bierkens, M. F. (2014). Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources. *Earth System Dynamics*, *5*(1), 15–40. doi: 10.5194/esd-5-15-2014

Wahr, J., Molenaar, M., & Bryan, F. (1998). Time variability of the Earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE. *Journal of Geophysical Research: Solid Earth*, *103*(B12), 30205–30229. doi: 10.1029/98JB02844

Watkins, M. M., Wiese, D. N., Yuan, D.-N., Boening, C., & Landerer, F. W. (2015a). 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. doi: 10.1002/2014JB011547

Watkins, M. M., Wiese, D. N., Yuan, D.-N., Boening, C., & Landerer, F. W. (2015b). 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. doi: 10.1002/2014JB011547

Wiese, D., Landerer, F. W., & Watkins, M. M. (2016). Quantifying and reducing leakage errors in the JPL RL05M GRACE mascon solution. *Water Resources Research*, *52*(9), 7490–7502. doi: 10.1002/2016WR019344

Wiese, D., Yuan, D., Boening, C., Landerer, F., & Watkins, M. (2018). JPL GRACE mascon ocean, ice, and hydrology equivalent water height release 06 coastal resolution improvement (CRI) filtered version 1.0. *DAAC: Pasadena, CA, USA*. (Last accessed: 1 February 2021) doi: 10.5067/TEMSC-3MJC6

Wiese, D. N., Landerer, F. W., & Watkins, M. M. (2016). Quantifying and reducing leakage errors in the JPL RL05M GRACE mascon solution. *Water Resources Research*, *52*(9), 7490–7502. doi: 10.1002/2016WR019344