AGU ADVANCING EARTH AND SPACE SCIENCES

Water Resources Research

RESEARCH ARTICLE
10.1029/2025WR040389

Global Groundwater Drought Assessment Revisited: A Holistic Re-Evaluation of the GRACE-Groundwater Drought Index Across Major Aquifers

Mohamed Akl[1,2], **Brian F. Thomas**[3], **and Peter J. Clarke**[1]

[1]Civil and Geospatial Engineering, School of Engineering, Newcastle University, Newcastle upon Tyne, UK, [2]Public Works Engineering Department, Faculty of Engineering, Tanta University, Tanta, Egypt, [3]Department of Earth Sciences, University College London, London, UK

Key Points:

- Multi-model Gravity Recovery and Climate Experiment (GRACE)-groundwater assessments reveal substantial uncertainty, undermining groundwater drought assessment
- Variations in aquifer memory and groundwater drought indicators reveal how memory shapes perceived groundwater drought
- Aquifer memory corresponds to fewer drought occurrences but leads to more prolonged and intense events within multi-model GRACE-groundwater assessments

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

M. Akl,
m.a.a.akl2@newcastle.ac.uk

Citation:

Akl, M., Thomas, B. F., & Clarke, P. J. (2025). Global groundwater drought assessment revisited: A holistic re-evaluation of the GRACE-groundwater drought index across major aquifers. *Water Resources Research*, *61*, e2025WR040389. https://doi.org/10.1029/2025WR040389

Received 3 MAR 2025
Accepted 30 NOV 2025

Abstract The Gravity Recovery and Climate Experiment (GRACE) and GRACE Follow-On missions have enriched global groundwater monitoring, forming the basis for tools that detect groundwater drought, including the GRACE-Groundwater Drought Index (GGDI). The reliability of GGDI is fundamentally tied to the accurate isolation of a representative groundwater storage anomaly (GRACE-GWA) signal from GRACE observations, a challenge heightened by the scarcity of direct water budget measurements and the diverse methodologies applied in GRACE data processing. In this global assessment, we integrate multi-model GRACE-GWA estimates into the GGDI framework to examine how variability among these estimates influences groundwater drought interpretation across 37 study aquifers. Results reveal substantial sensitivity of key drought indicators to input uncertainty, with maximum observed intra-basin discrepancies reaching 11 events, 122 months in maximum duration, 63.33 months in average duration, 24.47 in severity, and 5.4 in intensity. Aquifer memory, inferred from GGDI autocorrelation, reveals pronounced variability, most notably in the Nubian Basin where memory estimates range from 3 to 61 months amongst multi-model realizations. Aquifers with higher memory tended to experience fewer drought events, yet those droughts were typically longer and more intense. Our findings underscore that even modest discrepancies in GRACE-GWA methodologies can translate into considerable uncertainties in both drought indicators and aquifer memory, thereby compromising the reliability of groundwater drought assessments.

1. Introduction

Drought, characterized as an extended period of water scarcity, disrupts ecosystems, economies, and societies by virtue of an anomalous reduction in precipitation (Dracup et al., 1980; Mishra & Singh, 2010; Wilhite & Glantz, 1985; Wilhite et al., 2014). Although drought commences with a precipitation deficit, incoming water shortages propagate through the hydrological system, leading to diminished groundwater storage (Mishra & Singh, 2010). Groundwater drought presents a complex and evolving challenge, shaped by the interplay of climatic variability and anthropogenic pressures. Storage deficits can result from reduced recharge due to prolonged precipitation deficits and limited infiltration (Goodarzi et al., 2016) and excessive groundwater abstraction that exceeds natural replenishment rates (Bloomfield & Marchant, 2013; Bloomfield et al., 2015; Mishra & Singh, 2010). Intensifying climate extremes coupled with escalating human interventions (Famiglietti, 2014; Taylor et al., 2013; Thomas & Famiglietti, 2019) have magnified the frequency and severity of groundwater drought (Panda et al., 2007; Tallaksen & Van Lanen, 2023). Often seen as a natural buffer during hydrological (i.e., reduced streamflow) or agricultural (i.e., depleted soil moisture) droughts (Famiglietti, 2014), groundwater has reliably supported human consumption and agricultural production (Castle et al., 2014; Hughes et al., 2012; Scanlon et al., 2012; Siebert et al., 2010). However, unsustainable pumping practices undermine this resilience, compounding water scarcity and intensifying drought impacts (Castle et al., 2014; Famiglietti et al., 2011; Famiglietti & Rodell, 2013). Operational groundwater drought monitoring is hampered by the challenges of directly measuring groundwater storage, particularly in transboundary or remote aquifers where in situ monitoring is sparse or absent (Condon et al., 2021; Famiglietti et al., 2011; Lall et al., 2020; Giordano, 2009; Jasechko et al., 2024).

Thomas, Famiglietti, et al. (2017) introduced the Gravity Recovery and Climate Experiment (GRACE) groundwater drought index (GGDI), a normalized groundwater storage-based indicator to enumerate groundwater drought. By leveraging GRACE-derived groundwater storage anomalies (GRACE-GWA), the GGDI

captures both surpluses and deficits, offering a comprehensive evaluation of groundwater drought conditions (Thomas, Famiglietti, et al., 2017). Since its development, GGDI has been widely applied to analyze groundwater drought and characterize its spatiotemporal variability across diverse aquifer systems (e.g., Ali et al., 2022; Aon et al., 2024; Nikraftar et al., 2024; Song et al., 2024; Wang et al., 2020; Wang et al., 2022; Zhang et al., 2024). Its broad applicability has enabled insight into groundwater system responses across a range of hydroclimatic settings, facilitating the monitoring of drought onset, duration, intensity, and recovery, while also illuminating the influence of both natural variability and anthropogenic pressures. However, despite the growing application of GGDI, its interpretation hinges on the variability in GRACE-GWA time series (Thomas, Famiglietti, et al., 2017). This variability is driven by differences in GRACE processing methods and uncertainties in other water budget components (e.g., soil moisture, snow, surface water) (Akl et al., 2025; Akl & Thomas, 2023, 2024; Saxe et al., 2021). Many GGDI applications since Thomas, Famiglietti, et al. (2017) have overlooked this inherent variability, raising questions about the reliability of groundwater drought interpretations. For instance, Wang et al. (2020) identified a severe drought event in the North China Plain (GGDI $= -1.36$) between August 2013 and September 2014, while Ali et al. (2022) cited an extreme event in the Indus Basin in August 2010 (GGDI $= -1.81$), based on study periods of 2003–2015 and 2003–2016, respectively. Both studies relied on a single land surface model (LSM) to disaggregate GRACE-TWSA into GRACE-GWA, thereby disregarding structural model biases (Clark et al., 2008; Konapala et al., 2020; Ruddell et al., 2019) and possibly conflating groundwater signals with unaccounted surface water components (Akl et al., 2022; Akl & Thomas, 2024; Castle et al., 2014; Thomas & Nanteza, 2023).

Our evolving knowledge of the challenges in applying the water budget framework for estimating GRACE-GWA (Akl et al., 2025; Akl & Thomas, 2024) translates to a compelling need to re-examine the GGDI framework. Central to this re-examination is elucidating how variability amongst multi-model GRACE-GWA estimates impacts GGDI interpretation. This can be achieved by quantifying and characterizing variability in key groundwater drought indicators derived from the GGDI framework. In our study, we seek to (a) assess the influence of multi-model GRACE-GWA realizations on key groundwater drought indicators and (b) assess aquifer memory variability for perceived groundwater drought. Our analysis centers on the world's largest aquifers, which collectively store the majority of global groundwater and are vital to water and food security (Gleeson et al., 2012; Margat, 2008; Margat & Van der Gun, 2013; Van der Gun, 2022).

To achieve our objectives, this study employs an integrated water balance framework that synthesizes multiple GRACE solutions with a range of auxiliary data sets, resulting in up to 180 multi-model GRACE-GWA realizations across 37 aquifer systems. This framework enables a systematic evaluation of how uncertainties in GRACE-GWA propagate through the GRACE-Groundwater Drought Index (GGDI), influencing the robustness and reliability of groundwater drought characterization. It also incorporates aquifer memory estimation derived from GGDI autocorrelation, providing new insights into the persistence and recovery dynamics of aquifers under varying hydroclimatic conditions.

Rather than prescribing a singular or deterministic drought narrative, this study foregrounds the interpretive consequences of input divergence in GRACE-based drought assessment. By explicitly accounting for uncertainties introduced through variations in GRACE-TWSA solutions and water budget components, our multi-model framework provides a transparent and defensible approach to groundwater drought characterization. Our framework is particularly salient in data-sparse regions, where GRACE-based indicators may represent the sole line of evidence of drought onset and recovery. Ultimately, this study seeks to enhance the interpretability and credibility of GRACE-based drought assessments, supporting more context-aware and evidence-driven groundwater management decisions.

2. Data and Methods

This study focuses on 37 large aquifer systems as defined by the Worldwide Hydrogeological Mapping and Assessment Program (WHYMAP; Margat & Van der Gun, 2013) for the period spanning April 2002 to December 2022 (Figure 1 and Table S1 in Supporting Information S1). These aquifers are among the world's most productive, holding a substantial share of accessible global groundwater resources (Gleeson et al., 2012; Margat, 2007; Richey et al., 2015; Thomas, Caineta, & Nanteza, 2017; Van der Gun, 2022). In addition, each aquifer exceeds the minimum 100,000 km^2 spatial threshold necessary for reliable GRACE-based assessments (Scanlon et al., 2016).

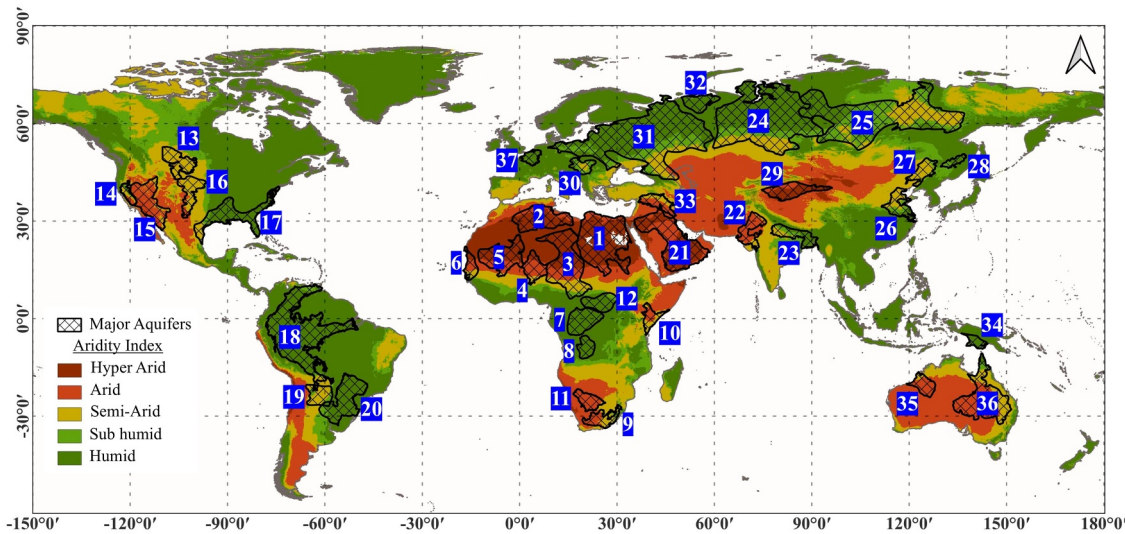

Figure 1. Map of the 37 study aquifers, sourced from the World-wide Hydrogeological Mapping and Assessment Program (WHYMAP; Margat & Van der Gun, 2013). Numeric labels refer to aquifer names listed in Table S1 in Supporting Information S1. The color scale represents the aridity index, sourced from the CGIAR-CSI Global Aridity Index and Potential Evapotranspiration Climate Database V3 (Global-AI_PET; Zomer et al., 2022), where Hyper-Arid regions are shown in dark brown, Arid in medium brown, Semi-Arid in light brown, Sub-Humid in medium green, and Humid in dark green.

These aquifers span a wide range of climatic zones, from hyper-arid to humid conditions, offering diverse contexts for investigating hydrological processes (Figure 1).

2.1. GRACE-Terrestrial Water Storage Anomalies (GRACE-TWSA)

GRACE-Terrestrial Water Storage Anomalies (GRACE-TWSA) were obtained from five GRACE-based solutions covering the interval from April 2002 to December 2022, including both spherical harmonic (SH; Landerer & Swenson, 2012; Landerer, 2021) and mass concentration (mascon; Landerer et al., 2020; Save, 2020; Save et al., 2016; Watkins et al., 2015; Wiese et al., 2016; Wiese et al., 2018) approaches. The SH solutions were provided on 1° × 1° global grids by three data processing centres: the Centre for Space Research (CSR-SH; RL06 V04), Jet Propulsion Laboratory (JPL-SH; RL06 V04), and GeoForschungsZentrum (GFZ-SH; RL06 V04) (https://grace.jpl.nasa.gov/; last access: 3 December 2024). Meanwhile, mascon solutions were supplied by CSR (CSR-M; RL06 V02) on a 0.25° × 0.25° global grid (https://www2.csr.utexas.edu/grace/; last access: 3 December 2024) and by JPL (JPL-M; RL06 V04) on a 0.5° × 0.5° global grid (https://grace.jpl.nasa.gov/; last access: 3 December 2024).

Data gaps between the GRACE and GRACE-FO missions (July 2017–May 2018) were omitted, and shorter missing intervals (one to two months) were filled using cubic spline interpolation (Aon et al., 2024; Wei et al., 2021). This study avoided JPL scale factors derived from the Community Land Model (CLM; Lawrence et al., 2011), due to concerns regarding their reliability in regions where key hydrological processes are either oversimplified or entirely absent (Wiese et al., 2016). For example, CLM has well-documented limitations in representing snow accumulation and melt dynamics, as well as anthropogenic groundwater withdrawals (Wiese et al., 2016), processes that are critical for accurately resolving groundwater storage changes in many of the aquifer systems investigated.

2.2. Water Budget Components

2.2.1. Snow Water Equivalent Anomalies (SWEA)

Monthly SWEA were derived from six modeling and reanalysis systems, each offering unique spatial resolutions and physical representations. Three data sets were obtained from LSMs provided by the NASA Global Land Data Assimilation System (GLDAS; V2.1; last access: 9 December 2024; Rodell et al., 2004), including NOAH at 1° × 1°, Variable Infiltration Capacity (VIC) at 1° × 1°, and Catchment Land Surface Model (CLSM) at 1° × 1°. High-resolution SWEA data were also sourced from NOAH at 0.10° × 0.10° via the Famine Early Warning

System Network (FEWS NET) Land Data Assimilation System (FLDAS; V001; last access: 9 December 2024; McNally et al., 2017). Additional data sets included the WaterGAP Global Hydrology Model v2.2e (WGHM; last access: 9 December 2024; Müller Schmied et al., 2024) at 0.5° × 0.5° resolution and the ERA5-Land data set at 0.10° × 0.10° resolution (Copernicus Climate Change Service, 2022; last access: 9 December 2024; Muñoz-Sabater et al., 2021).

2.2.2. Soil Moisture Anomalies (SMA)

Soil moisture anomalies (SMA) were derived from the same modeling systems and reanalysis used to estimate SWEA (Section 2.2.1). These model-based SMA data sets are widely used in GRACE-based groundwater assessments due to their long-term availability and spatial completeness. In contrast, remote sensing products, such as SMAP (Entekhabi et al., 2010), were deliberately excluded due to limitations including elevated noise, temporal discontinuities and limited sensing depth (Beck et al., 2021).

The employed model output represents soil moisture using distinct structural formulations and vertical discretizations. The NOAH model (at 1° and 0.10° resolutions), simulates soil moisture through a multi-layer soil profile divided into four depths: 0–10 cm, 10–40 cm, 40–100 cm, 100–200 cm (Rodell et al., 2004). VIC represents soil moisture within three layers: a surface layer (0–30 cm), a variable-depth second layer, and a bottom layer. CLSM conceptualizes soil moisture within three reservoirs: the Surface layer (0–2 cm), Root Zone (0–100 cm), and the Profile reservoir (varies grid-by-grid). WGHM simulates monthly soil moisture within the effective root zone using a single-layer soil water storage compartment that integrates land cover, soil-specific maximum storage capacity, and soil texture (Müller Schmied et al., 2024). ERA5-Land provides volumetric soil moisture estimates across four layers: 0–7 cm, 7–28 cm, 28–100 cm, and 100–289 cm (Muñoz-Sabater et al., 2021). To avoid inclusion of saturated or groundwater-affected zones, especially across humid basins (Fan et al., 2013), only the upper 0–1 m of the modeled soil column was considered. This depth range captures the dynamic portion of the root zone and is widely used in GRACE-based studies to represent the soil moisture component of the terrestrial water balance (Akl & Thomas, 2024; Thomas et al., Thomas, Famiglietti, et al., 2017).

However, soil thickness remains a critical and often underappreciated source of structural uncertainty in water balance modeling. Numerous studies have shown that the depth of modeled soil layers can significantly influence soil moisture dynamics, particularly during drought conditions. Shallow soil schemes often produce unrealistically persistent moisture signals due to limited water-holding capacity and weak coupling with deeper subsurface layers (Houborg et al., 2012; Swenson & Lawrence, 2015; Thomas, Famiglietti, et al., 2017; Thomas, Caineta, & Nanteza, 2017). Most LSMs and hydrological models impose fixed saturation thresholds and typically do not simulate dynamic groundwater–soil moisture interactions (Rodell & Famiglietti, 2001). In humid regions, where the simulated water table may intersect the upper soil layers, this can lead to inadvertent inclusion of shallow groundwater in modeled SMA, causing underestimation of GRACE-GWA (Fan et al., 2013). In contrast, in arid and semi-arid regions characterized with deep unsaturated zones, modeled soil moisture may underrepresent actual storage changes, resulting in overestimation of GRACE-GWA (Fan et al., 2013; Scanlon et al., 2009; Shamsudduha & Taylor, 2020). These discrepancies highlight the critical importance of accounting for soil depth assumptions and their hydrological implications when incorporating SMA into the GRACE-based groundwater framework.

2.2.3. Surface Water Storage Anomalies (SWA)

Surface water storage anomalies (SWA) primarily capture variations in lakes, reservoirs, rivers and floodplains, which play a pivotal role in modulating basin-scale hydrological regimes. Over half of the world's major river systems are influenced by reservoir operations (Nilsson et al., 2005), and this influence is expected to grow with the ongoing expansion of global storage infrastructure (Duan & Bastiaanssen, 2013). Despite their significance, consistent and spatially resolved in situ observations of surface water dynamics remain limited. Where such records exist, they are often fragmented, inconsistently reported, or inaccessible due to institutional, political, or logistical constraints (Alsdorf et al., 2007; Busker et al., 2019; Tortini et al., 2020).

In the absence of systematic monitoring networks, SWA has traditionally been approximated using river-routing components embedded in large-scale hydrological models (e.g., Han et al., 2009; Kim et al., 2009), land surface models such as GLDAS (e.g., Shamsudduha & Taylor, 2020; Thomas, Caineta, & Nanteza, 2017), or reanalysis products like ERA5-Land (e.g., Amazirh et al., 2024). Although these data sets provide long-term continuity and

global coverage, they remain constrained by limited calibration and sparse validation, particularly across data-scarce regions, resulting in substantial uncertainty in simulated surface water storage (Gao et al., 2012).

To address SWA accounting challenges (Alsdorf et al., 2007; Busker et al., 2019; Tortini et al., 2020), this study assumed that reservoir and lake storage represents SWA and applied integrated satellite-based data sets from GloLakes (Hou et al., 2024; last access: 08 June 2023) and the Copernicus Climate Change Service (Copernicus Climate Change Service, 2020; https://cds.climate.copernicus.eu/; V5; last access: 15 December 2024). The Hydrological Data and Maps Based on Shuttle Elevation Derivatives at Multiple Scales data set (HydroSHEDS; Lehner et al., 2008; V10; last access: 01 June 2023) was used to delineate lake and reservoir extents (Figure S1 in Supporting Information S1). While GloLakes offered lake-storage data, Copernicus supplied lake water-level observations that required integration with time-varying lake area. This integration was achieved using the global lake bathymetry data set (GLOBathy; Khazaei et al., 2022), enabling a reliable representation of surface water storage changes (Akl & Thomas, 2024).

A quality-control algorithm excluded lakes with more than one missing seasonal reading in any calendar year, producing a final data set of 429 lakes (415 from GloLakes and 14 from Copernicus; Table S2 and Figure S2 in Supporting Information S1). The data were aggregated to produce consistent monthly intervals, with any remaining gaps filled via cubic spline interpolation.

2.3. Disaggregating GRACE-TWSA Into GRACE-GWA

GRACE-TWSA offers an integrated estimate of terrestrial water storage changes, encompassing soil moisture anomalies (SMA), snow water equivalent anomalies (SWEA), surface water anomalies (SWA), vegetation canopy anomalies (VCA), and groundwater storage anomalies (GRACE-GWA). In this study, VCA was excluded from the water balance equation due to its negligible influence on GRACE-TWSA, consistent with prior work across a range of hydroclimatic settings (Nanteza et al., 2016; Richey et al., 2015; Thomas, Caineta, & Nanteza, 2017). Isolating GRACE-GWA from GRACE-TWSA relies on the water balance equation (Equation 1), which assumes that accurate accounting for water budget components (i.e., SMA, SWEA, SWA) allows the extraction of a representative GRACE-GWA signal (Rodell & Famiglietti, 1999).

$$\text{GRACE-TWSA}_t = \text{SMA}_t + \text{SWEA}_t + \text{SWA}_t + \text{GRACE-}GWA_t \quad (1)$$

where the subscribe t denotes time.

To ensure consistency with GRACE processing methodologies, all water budget components (soil moisture, snow, and surface water) were processed in accordance with GRACE standards (Nanteza et al., 2016). For spherical harmonic (SH) solutions, with significant leakage potential (Awange et al., 2014; Fatolazadeh et al., 2016; Guo et al., 2016; Klees et al., 2007; Landerer & Swenson, 2012; Lenk, 2013; Longuevergne et al., 2013; Nanteza et al., 2016; Swenson & Wahr, 2002), water budget data were projected onto GRACE grids and underwent truncation at harmonic degree 60, destriping with a decorrelation filter (Swenson & Wahr, 2006) and smoothing using a 300 km radius Gaussian filter (Landerer, 2021; Landerer & Swenson, 2012). Subsequently, these processed components were converted to anomalies by removing the mean of the time series over the period from January 2004 to December 2009, before being subtracted from GRACE-TWSA grids. In contrast, for mascon solutions, potential leakage across GRACE grids was presumed to be effectively addressed during processing (Bhanja et al., 2016; Neves et al., 2020; Richey et al., 2015; Rodell et al., 2018; Save et al., 2012, 2016; Scanlon et al., 2016, 2021, 2022; Thomas, Caineta, & Nanteza, 2017; Thomas, Famiglietti, et al., 2017; Thomas & Famiglietti, 2019; Watkins et al., 2015; Wang et al., 2020; Wiese et al., 2016). Water budget components were converted to anomalies by subtracting the mean over January 2004 to December 2009 and further subtracted from the respective GRACE grids prior to averaging at the basin scale. The original spatial resolutions of both SH and mascon grids were preserved, with data clipped to basin boundaries to maintain spatial fidelity and integrity (Akl & Thomas, 2024).

This analysis utilized a structured multi-model framework that integrated five GRACE-TWSA solutions, six SMA data sets, six SWEA data sets, and a SWA estimate within the water balance equation (Equation 1) to derive multiple GRACE-GWA realizations across 37 study aquifers. These aquifers span a diverse range of climatic and hydrological conditions, reflecting substantial variability in the composition of water budget components incorporated into the water balance equation (Equation 1; Figure 1). Rather than aggregating data sets, each was

treated as a discrete input to explicitly represent structural uncertainty arising from input selection. For aquifers where all four components (GRACE-TWSA, SMA, SWEA, and SWA) were available, this configuration yielded 180 possible GRACE-GWA realizations (5 GRACE-TWSA × 6 SMA × 6 SWEA × 1 SWA). Where one or more components were unavailable (i.e., SWEA and SWA), the number of realizations was adjusted accordingly (Table S3 in Supporting Information S1).

This multi-model framework was not intended to estimate internal variability or measurement error within a given configuration, but instead to characterize the methodological divergence that emerges from the use of multiple, plausible input data sets (Akl & Thomas, 2024). While this cross-combination may not preserve strict physical consistency between individual water budget components (e.g., pairing SWEA from one model with SMA from another), the intention was not to maintain such consistency but to quantify the methodological divergence that arises from multiple, independently derived yet plausible data sets. This approach aligns with recent multi-model studies designed to assess input-driven uncertainty in GRACE-based groundwater storage estimates (e.g., Akl et al., 2025; Akl & Thomas, 2024; Li et al., 2023; Rateb et al., 2020).

2.4. GRACE-Groundwater Drought Index (GGDI)

The GRACE-Groundwater Drought Index (GGDI; Thomas, Famiglietti, et al., 2017) is applied to detect and quantify groundwater drought across study aquifers. As a normalized measure of groundwater storage deviation, GGDI captures both natural variability and anthropogenic pressures on subsurface storage, thereby offering an integrated perspective on drought conditions (Thomas, Famiglietti, et al., 2017). This interpretation aligns with the conceptual framework proposed by Van Loon et al. (2016), which highlights the importance of accounting for both natural and human-induced factors in drought characterization. GGDI is derived from GRACE-GWA and reflects departures from expected seasonal conditions. To remove the influence of seasonality, a monthly climatology (C_i) was calculated for each calendar month ($i = 1, 2,\ldots,12$) given as:

$$C_i = \frac{\sum_1^{n_i} \text{GRACE-GWA}_i}{n_i} \tag{2}$$

where n_i represents the number of years of data available for month i (Thomas et al., 2014; Thomas, Famiglietti, et al., 2017). This climatology defines the expected seasonal pattern of groundwater storage and serves as the reference baseline for anomaly detection. The climatology (C_i) is removed from GRACE-GWA to compute the Groundwater Storage Deviation (GSD), which represents net seasonal variation in groundwater storage (Thomas, Famiglietti, et al., 2017). The GSD was further standardized to derive the GGDI using the formula:

$$\text{GGDI}_t = \frac{\text{GSD}_t - \overline{\text{X}}_{\text{GSD}}}{\sigma_{\text{GSD}}} \tag{3}$$

where t accounts for time and $\overline{\text{X}}_{\text{GSD}}$ and σ_{GSD} are the mean and standard deviation of the GSD time series, respectively (Thomas, Famiglietti, et al., 2017). Following Thomas, Famiglietti, et al. (2017), a groundwater drought event is defined as a period beginning with at least three consecutive months of GGDI < 0, and it terminates after three consecutive months of GGDI ≥ 0.

2.5. Groundwater Drought Indicators

GGDI indicators were applied to characterize groundwater drought: *the number of drought events*, representing the total count of distinct drought occurrences; *the maximum drought duration*, defined as the total number of months characterizing the longest drought over the observational period; *the average drought duration*, capturing the mean length of all drought events occurring during the observational period; *the drought severity*, quantified as the cumulative sum of GGDI values during drought periods; and *drought maximum intensity*, representing the minimum GGDI value recorded across all drought events, captures the most extreme intensity observed among all identified drought periods.

2.6. Aquifer Memory

Aquifer memory refers to the temporal persistence of groundwater storage anomalies, capturing how long the hydrological impacts of a drought event continue to influence subsequent conditions. It reflects the intrinsic response time of an aquifer system and serves as a proxy for its resilience to climatic variability and anthropogenic stress (Neves, 2024; Opie et al., 2020). In this study, aquifer memory was quantified by examining the autocorrelation structure of the GGDI time series. Specifically, memory was defined as the number of months required for the autocorrelation coefficient to decline below 0.1, a threshold commonly used to indicate the point at which the influence of antecedent conditions becomes statistically insignificant (Bloomfield & Marchant, 2013; Dubois & Larocque, 2024).

Aquifer memory provides an informative characterization of groundwater persistence, enabling cross-aquifer comparison of recovery dynamics. Longer memory values indicate sustained drought impacts and delayed hydrological recovery, while shorter memory signals a more responsive system. To examine how memory correlated with drought behavior, estimated memory were systematically compared with key groundwater drought indicators, including event frequency, maximum and average duration, severity, and maximum intensity.

3. Results

3.1. Propagation of Uncertainty Through GRACE-Groundwater Drought Index (GGDI)

Figure 2 depicts the variability in multi-model GRACE-GWA for the Northern Great Plains Aquifer–Judith River (Basin 13), emphasizing how discrepancies in GRACE processing methods and water budget accounting propagate through each sequential stage of GGDI derivation. The figure is arranged in four panels—GRACE-GWA (top), Climatology (second), GSD (third), and GGDI (bottom)—with each panel displaying the multi-model realizations (gray lines/shading) and their corresponding mean (green line). In the top panel, the GRACE-GWA time series is anchored to a January 2004–December 2009 baseline, rendering anomalies within this reference window comparably subdued. Post-2009, however, the amplitude volatility increases, particularly between 2010 and 2012 when the aquifer recharged due to reported episodes of rapid snowmelt and intense precipitation (Coles et al., 2017; Dumanski et al., 2015). The divergence among realizations in both the magnitude and timing reflects the influence of model structural bias when disaggregating GRACE-TWSA into GRACE-GWA (Akl & Thomas, 2024).

In the second panel (Climatology), the derived monthly mean captures the aquifer's expected seasonal cycle of storage increases and decreases, exposing significant discrepancies among estimates driven by divergent model assumptions. These discrepancies are particularly pronounced during seasonal peaks and troughs, where impediments in water partitioning led to pronounced variations in climatology amplitudes. In the third panel, the Groundwater Storage Deficit (GSD) quantifies deviation from the seasonal climatology, delineating intervals of groundwater surplus (positive GSD) or deficits (negative GSD). Notably, abrupt shifts—such as those noted in early 2010—coincide with notable expansions in the GSD realization spread, signaling heightened ambiguity in model water budget allocation.

Finally, in the bottom panel (GGDI), GSD values are standardized to allow comparisons of groundwater drought intensity. By converting absolute storage deficits to a z-score measure, GGDI is thought to reduce amplitude-driven uncertainty. Variability in GGDI, attributed to GRACE processing methodologies and water budget component selection, increases notably during abrupt hydrologic transitions, indicating heightened model water allocation discord under rapid change. For example, GGDI values range from 0.25 to −2.80 in January 2003, 0.15 to −2.75 in February 2010, 2.40 to −1.50 in March 2011, and 1.00 to −2.15 in June 2022. When the realization means hovers near zero (e.g., in 2010–2012 or 2019–2020), small deviations can shift apparent conditions between mild drought and near-normal, fundamentally altering drought classification.

3.2. Groundwater Drought Indicators Across Study Basins

Figure 3 presents five key groundwater drought indicators for the 37 aquifer systems. The number of drought events exhibits substantial inter-basin variability, with ranges extending up to 11 events (Figure 3 and Table 1). For instance, Basin 25 (Yakut; Sub-Humid) displayed the widest range—from 1 to 12 events—while Basin 22 (Indus; Arid) showed the narrowest, with counts between 3 and 5. In general, humid basins tend to register both a higher number of drought events and greater variability, whereas arid basins typically record fewer events with

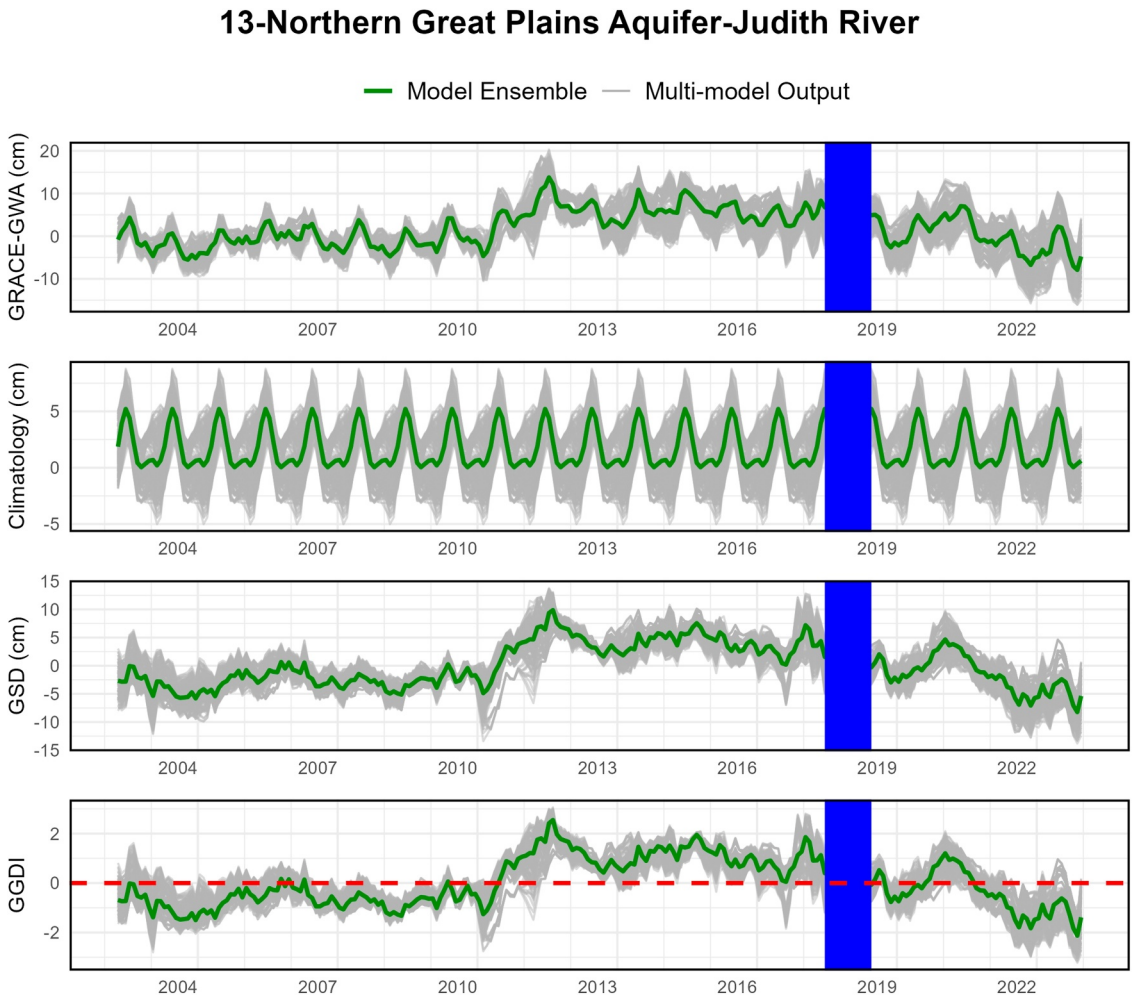

Figure 2. Time series of GRACE-GWA and associated climatology, groundwater storage defects (GSD), and GRACE-groundwater drought index (GGDI) across study basin number 13 (Northern Great Plains Aquifer-Judith River) from April 2002 to December 2022. The blue-shaded area marks the gap period between the GRACE and GRACE-FO missions, which was excluded from the analysis. For the time series of all 37 study basins, please refer to Figures S3:S39 in Supporting Information S1.

less variability. Nonetheless, notable outliers exist among arid basins: Basin 1 (Nubian) recorded a range of 1–10 events, Basin 5 (Taoudeni-Tanezrouft) ranged from 3 to 11, and Basin 29 (Tarim) also varied between 3 and 11 events.

Variability in maximum drought duration also differed markedly across basins, with ranges spanning from 0 to 122 months (Figure 3 and Table 1). For example, Basin 1 (Nubian; Hyper-Arid) exhibited the greatest variability, with maximum drought durations ranging from 30 to 152 months, whereas Basin 22 (Indus; Arid) demonstrated no variability, with a consistently recorded maximum drought duration of 55 months. Average drought duration followed a similar pattern, varying between 6 and 63.5 months. Notably, Basin 25 (Yakut; Sub-Humid) recorded the highest variability in average duration—ranging from 11 to 74.5 months—while Basin 22 (Indus; Arid) again exhibited the lowest variability, with average durations confined to a narrow band between 32.5 and 38.5 months. The overall findings suggest that arid basins tend to exhibit longer drought duration and greater variability in both maximum and average drought durations, in contrast to shorter drought duration and reduced variability observed in humid basins.

Drought severity revealed pronounced inter-basin differences (Figure 3 and Table 1). Basin 1 (Nubian; Hyper-Arid) exhibited the highest variability, with severity values ranging from −101 to −76.5 (a range of 24.5), whereas Basin 18 (Amazonas; Humid) showed the lowest variability, with values spanning from −102.5 to −97.5 (a range of 5.0). In general, arid basins tended to display high variability in drought severity—except for Basins 4

10.1029/2025WR040389

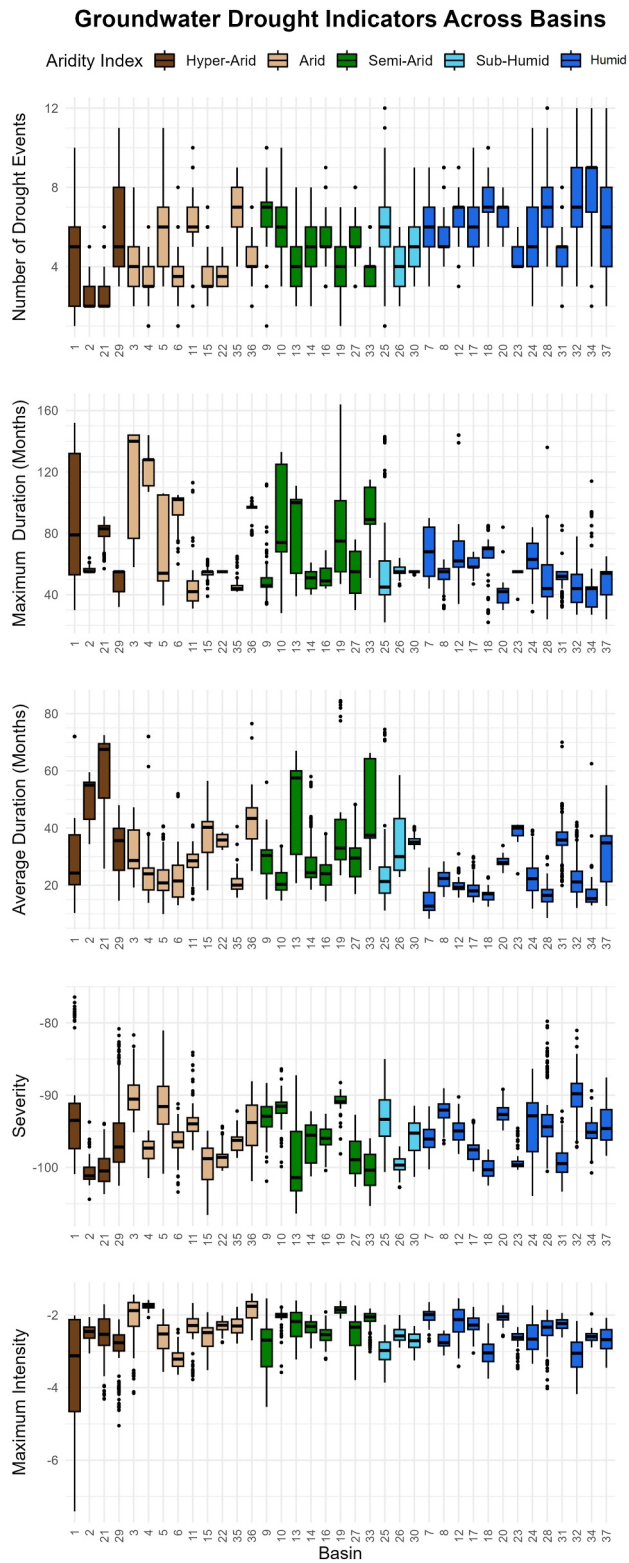

Figure 3. Groundwater drought indicators derived from GGDI time series for 37 study aquifers spanning April 2002 to December 2022. Five indicators are displayed across panels: the number of drought events, maximum duration, average duration, severity, and maximum intensity. Aquifers are ordered by aridity class to facilitate climate-based comparisons, where Hyper-Arid regions are shown in dark brown, Arid in light brown (tan), Semi-Arid in green, Sub-Humid in sky blue, and Humid in medium blue (royal blue). The gap period between the GRACE and GRACE-FO missions was excluded from this analysis.

Table 1

Climate Zones Across Study Basins and the Variability Ranges of Groundwater Drought Indicators, Defined as the Difference Between the Maximum and Minimum Values, Including the Number of Drought Events, Maximum Duration (Months), Average Duration (Months), Severity, and Maximum Intensity

		Range of variability (maximum—minimum)				
Basin	Climate zone	Number of drought events	Maximum duration (m)	Average duration (m)	Severity	Max intensity
1	Hyper Arid	9	122	61.69	24.47	5.39
2	Hyper Arid	3	10	25.13	10.68	1.00
21	Hyper Arid	4	34	46.67	9.76	2.61
29	Hyper Arid	8	23	33.42	21.71	2.93
3	Arid	6	86	28.04	13.46	2.72
4	Arid	5	37	58.17	6.59	0.48
5	Arid	8	73	30.73	19.81	1.74
6	Arid	7	45	39.00	12.19	1.25
11	Arid	8	82	25.90	13.53	2.11
15	Arid	5	24	38.25	11.66	1.60
22	Arid	2	0	6.17	6.14	0.74
35	Arid	5	23	24.83	6.47	1.01
36	Arid	5	24	51.50	13.81	1.29
9	Semi-Arid	9	78	40.98	13.59	2.99
10	Semi-Arid	7	105	19.04	13.49	1.80
13	Semi-Arid	6	72	46.25	19.12	1.62
14	Semi-Arid	6	21	39.57	9.01	0.93
16	Semi-Arid	6	25	23.58	7.85	1.30
19	Semi-Arid	6	117	61.00	9.85	0.49
27	Semi-Arid	5	46	31.33	9.95	2.06
33	Semi-Arid	3	64	40.92	9.37	1.19
25	Sub humid	11	121	63.33	15.63	1.59
26	Sub humid	4	18	35.63	5.66	0.90
30	Sub humid	6	2	8.00	9.83	0.94
7	Humid	6	46	17.85	8.71	1.08
8	Humid	5	32	12.42	7.66	0.69
12	Humid	6	110	15.30	7.89	1.88
17	Humid	6	21	16.00	7.06	1.27
18	Humid	5	63	10.38	5.00	1.52
20	Humid	3	54	9.63	5.70	0.86
23	Humid	2	18	17.17	5.74	1.15
24	Humid	9	55	27.42	17.59	1.61
28	Humid	8	112	21.44	20.76	2.17
31	Humid	6	53	50.08	11.12	0.67
32	Humid	9	51	29.94	16.15	2.01
34	Humid	10	87	49.56	11.36	0.92
37	Humid	10	41	42.23	10.82	1.37

Note. For comprehensive statistical details, please refer to Tables S4:S9 in Supporting Information S1.

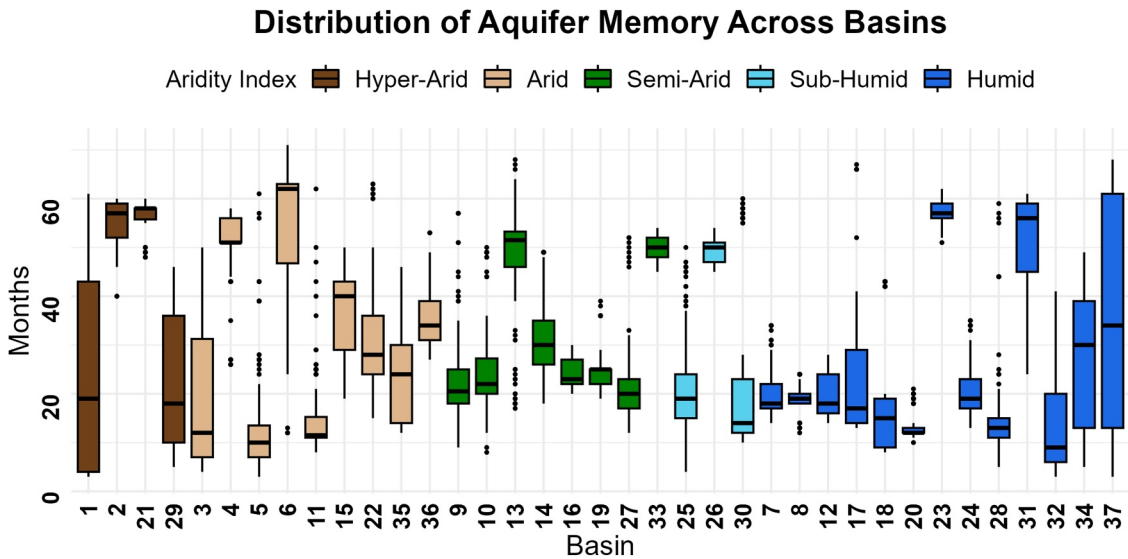

Figure 4. Box blots of aquifer memory (in months) across 37 study aquifers, ordered by aridity class, where Hyper-Arid regions are shown in dark brown, Arid in light brown (tan), Semi-Arid in green, Sub-Humid in sky blue, and Humid in medium blue (royal blue).

(Iulluemeden-Irhazer; Arid), 16 (Ogallala; Semi-Arid), 22 (Indus; Arid), and 35 (Canning; Arid), which exhibited narrower ranges of 6.5, 7.8, 6.1, and 6.5, respectively. Conversely, while humid basins typically exhibited lower variability, outliers were noted in Basins 24 (West Siberian), 28 (Middle Heilongjiang-Amur River), and 32 (Pechora), with ranges of 17.5, 20.8, and 16.1, respectively.

Maximum intensity revealed significant heterogeneity across the study basins (Figure 3 and Table 1). Notably, Basin 1 (Nubian; Hyper-Arid) exhibited the highest variability, with GGDI ranging from −2.0 to −7.5—a range of approximately 5.5 points. Following this apex of variability, other arid basins—specifically Basins 3 (Chad), 9 (Karoo), 21 (Arabian) and 29 (Tarim)—displayed variability with ranges around 3 points. Meanwhile, Basins 4 (Iulluemeden-Irhazer; Arid), 8 (Coango; Humid), 19 (Yrendá-Toba-Tarijeño; Semi-Arid), and 31 (Russian Platform; Humid) exhibited minimal variability, confined to roughly 0.50 points. The findings underscore that arid basins are distinguished by both elevated maximum intensity values and markedly greater variability, whereas humid basins consistently register lower maximum intensity values accompanied by reduced variability.

In humid basins, drought events occur more frequently—with a broad range of variability—yet these events tend to have relatively short and consistent durations and severity. In contrast, arid basins, while recording fewer events with a narrower variability range, often exhibit prolonged durations and more extreme intensity values, accompanied by high variability. Moreover, several outlier basins underscore the complex influence of multi-model GRACE-GWA variability on drought characterization.

3.3. Aquifer Memory Across Study Basins

Figure 4 displays aquifer memory estimates for the 37 study basins. These memory estimates are derived from a series of GGDI calculations based on multi-model GRACE-GWA estimates. Aquifer memory demonstrated striking inter-basin variability, ranging from 9 to 65 months (Figure 4 and Table 2). For example, Basin 37 (Paris; Humid) exhibited the broadest range, with estimates spanning from 3 to 68 months, whereas Basins 26 (North China Plains; Sub-Humid) and 33 (Taurus-Zagros; Semi-Arid) showed much narrower ranges, varying from 45 to 54 months.

In general, hyper-arid and arid basins tended to display both higher aquifer memory values and greater variability. For instance, Basin 1 (Nubian) exhibited a variability range of 58 months (3–61 months), and Basin 6 (Senegalo-Mauretanian) recorded some of the highest values, ranging from 12 to 71 months. An intriguing outlier was Basin 21 (Arabian), which, despite having relatively high memory values (48–60 months), showed only a 12-month variability range, possibly reflecting the combined effects of minimal precipitation and intensive groundwater

Table 2
Climate Zones Across Study Basins and the Aquifer Memory Statistics, Including Minimum Values (Min), Lower Quartile (Q1), Median, Upper Quartile (Q3), and Maximum Values (Max)

		Aquifer memory statistics				
Basin	Climate zone	Min	Q1	Median	Q3	Max
1	Hyper Arid	3	4	19	43	61
2	Hyper Arid	40	52	57	59	60
21	Hyper Arid	48	55.75	58	58	60
29	Hyper Arid	5	10	18	36	46
3	Arid	4	7	12	31.25	50
4	Arid	26	51	51	56	58
5	Arid	3	7	10	13.5	61
6	Arid	12	46.75	62	63	71
11	Arid	8	11	11.5	15.25	62
15	Arid	19	29	40	43	50
22	Arid	15	24	28	36	63
35	Arid	12	14	24	30	46
36	Arid	27	31	34	39	53
9	Semi-Arid	9	18	20.5	25	57
10	Semi-Arid	8	20	22	27.25	50
13	Semi-Arid	17	46	51.5	53.25	68
14	Semi-Arid	18	26	30	35	49
16	Semi-Arid	20	22	23	27	30
19	Semi-Arid	19	22	25	25	39
27	Semi-Arid	12	17	20	23	52
33	Semi-Arid	45	48	50	52	54
25	Sub humid	4	15	19	24	50
26	Sub humid	45	47	50	51	54
30	Sub humid	10	12	14	23	60
7	Humid	14	17	18	22	34
8	Humid	12	18	19	20	24
12	Humid	14	16	18	24	28
17	Humid	13	14	17	29	67
18	Humid	8	9	15	19	43
20	Humid	10	12	12	13	21
23	Humid	51	56	57	59	62
24	Humid	13	17	19	23	35
28	Humid	5	11	13	15	59
31	Humid	24	45	56	59	61
32	Humid	3	6	9	20	41
34	Humid	5	13	30	39	49
37	Humid	3	13	34	61	68

extraction for agricultural purposes (Siebert et al., 2010). In contrast, humid and sub-humid basins generally exhibited lower aquifer memory values with minimal variability. For example, Basin 18 (Amazonas) recorded memory estimates ranging from 14 to 28 months, a narrow variability of just 14 months. Yet, outliers such as Basin 23 (Ganges-Brahmaputra) and Basin 26 (North China Plains), despite maintaining relatively narrow ranges, reported considerably higher memory values (51–62 and 45–54 months, respectively), likely reflecting the impact of dense populations and intense irrigation demand (Richey et al., 2015). Notably, Basin 37 (Paris) emerges as an extreme case, demonstrating the most pronounced variability with a range spanning 65 months.

Figure 5 compares aquifer memory estimates derived from GRACE-GWA and GGDI time series across the 37 study basins. In both cases, memory is defined as the number of months required for the lagged autocorrelation of the time series to fall below 0.1, following a consistent methodological framework. This comparison isolates the influence of climatology removal, applied during GGDI derivation, on the temporal persistence of groundwater anomalies.

The removal of seasonal climatology exerts a substantial influence on aquifer memory, altering both the magnitude of estimates and the range of inter-realization variability across the 37 study basins (Figure 5). While aquifer memory generally increases following climatology removal, the extent and nature of this change vary markedly across hydroclimatic regimes. In humid basins, where groundwater storage is regulated by recurring seasonal recharge (Jasechko et al., 2014), removing the climatological signal exposes slower, interannual fluctuations—thereby increasing both aquifer memory and variability in aquifer memory magnitudes. For example, the range of aquifer memory widened in Basin 18 (Amazonas) from 1 to 35 months, in Basin 12 (Karoo-Carbonate) from 1 to 14 months, and in Basin 8 (Coango) from 1 to 12 months.

Conversely, arid and semi-arid basins, where groundwater dynamics are shaped primarily by episodic recharge and long-term storage trends, tended to show reduced aquifer memory variability was noted following climatology removal. Notable examples include Basin 33 (Taurus-Zagros), where range narrowed from 47 to 9 months; Basin 21 (Arabian), from 27 to 12 months; Basin 16 (Ogallala), from 23 to 10 months; and Basin 2 (Northwest Sahara), from 43 to 20 months. In a few basins, the effect was minimal, as observed in Basins 30 (North Caucasus; Sub-Humid) and 36 (Great Artesian; Arid), where aquifer memory was little changed, from 49 to 50 and from 25 to 26 months, respectively. In certain cases, the influence of climatology removal was especially pronounced. For instance, in Basin 23 (Ganges-Brahmaputra), aquifer memory increased considerably from 5–30 months to 51–62 months.

Notable changes in aquifer memory underscore the hydrological significance of climatology removal in shaping aquifer memory diagnostics. In humid regions, where seasonal recharge dominates, eliminating the recurring signal unveils deeper persistence patterns that substantially influence aquifer memory. Conversely, in arid environments, where groundwater variability is driven more by episodic or decadal-scale events (Scanlon et al., 2022; Shamsudduha & Taylor, 2020), the effect of climatology removal is comparatively muted. This contrast highlights the methodological sensitivity of aquifer memory to signal decomposition and reinforces the need to interpret persistence metrics through the lens of basin-specific hydroclimatic context.

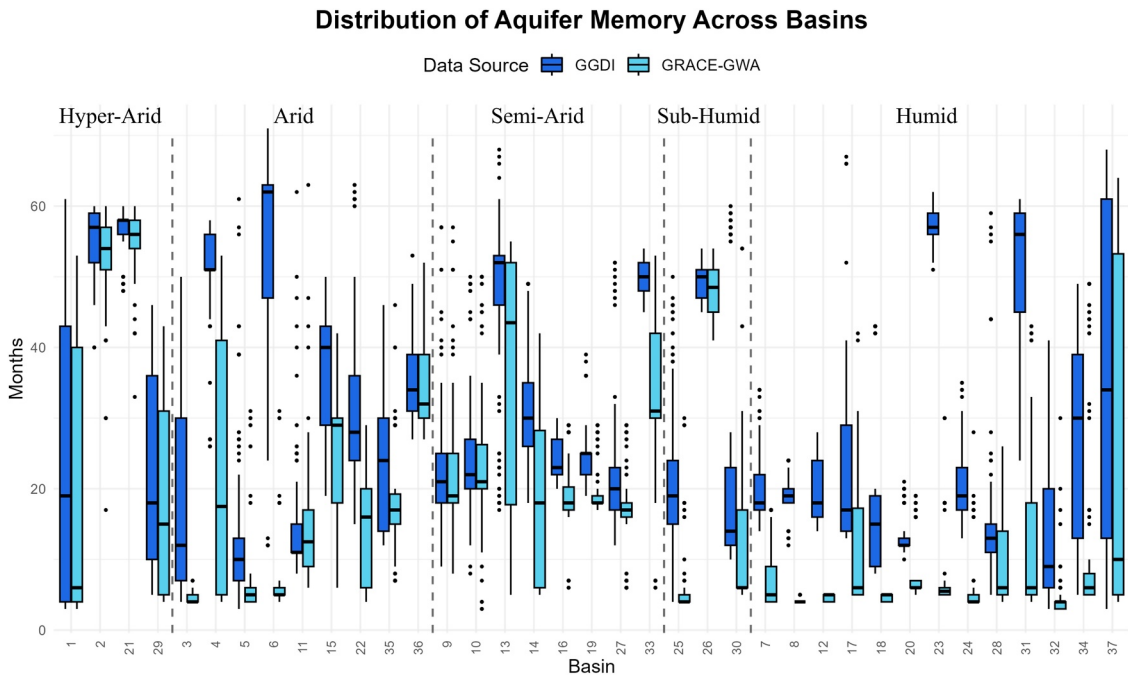

Figure 5. A comparison of aquifer memory (in months), calculated from the GRACE-Groundwater Drought Index (GGDI; medium blue (royal blue)) and GRACE-groundwater storage anomalies (GRACE-GWA; sky blue), across 37 study aquifers.

3.4. Correlation Analysis of Aquifer Memory and Drought Indicators

Spearman correlation analysis was conducted to document the relation between aquifer memory and groundwater drought indicators across study basins (Figure 6). When comparing aquifer memory and drought event frequency, average Spearman correlation coefficients (ρ) were −0.77 for hyper-arid, −0.64 for arid, −0.56 for semi-arid, −0.64 for sub-humid, and −0.54 for humid basins. These negative correlations indicate that aquifer memory is generally linked to reduced drought event frequency, suggesting that aquifer memory may serve as an effective buffer against frequent drought occurrences. However, certain basins—such as Basin 8 (Coango; Humid), Basin 17 (Atlantic and Gulf Coastal Plains; Humid), and Basin 33 (Taurus–Zagros; Semi-Arid)—display positive correlations (0.26, 0.63, and 0.33, respectively).

In contrast, the relationship between aquifer memory and maximum drought duration generally shows a positive correspondence (Figure 6). Average correlations are $\rho = 0.50$ for hyper-arid, 0.49 for arid, 0.50 for semi-arid, 0.36 for sub-humid, and 0.30 for humid basins, indicating that while aquifer memory may hinder frequent droughts, it prolongs the maximum duration of drought events when they do occur. However, notable deviations emerge in specific cases— Basin 20 (Guarani; Humid), Basin 21 (Arabian; Hyper-Arid), and Basin 33 (Taurus–Zagros; Semi-Arid) exhibit negative correlations (−0.07, −0.41, and −0.57, respectively). A similar positive association is observed when comparing aquifer memory with mean drought duration, with average correlations of $\rho = 0.77$ for hyper-arid, 0.15 for arid, 0.50 for semi-arid, 0.21 for sub-humid, and 0.54 for humid basins. Yet again, significant outliers are evident: Basins 8 (Coango; Humid), 9 (Karoo; Semi-Arid), 20 (Guarani; Humid), 21 (Arabian; Hyper-Arid), 23 (Ganges–Brahmaputra; Humid), 33 (Taurus -Zagros), and 36 (Great Artesian) display negative correlations ranging from −0.05 to −0.41.

Drought severity generally exhibits a negative relationship with aquifer memory (Figure 6). Hyper-arid basins exhibit an average correlation of $\rho = -0.80$, arid basins $\rho = -0.53$, semi-arid basins $\rho = -0.52$, sub-humid basins $\rho = -0.73$, and humid basins $\rho = -0.51$. This pattern suggests that aquifer memory is capable of mitigating the severity of drought conditions. However, variability persists, as evidenced by basins such as Basin 8 (Coango; Humid), Basin 11 (Stampriet-Kalahari; Arid), Basin 17 (Atlantic and Gulf Coastal Plains; Humid), Basin 20 (Guarani; Humid), Basin 25 (Yakut; Sub-humid), Basin 27 (Song-Liao; Semi-Arid), Basin 33 (Taurus–Zagros; Semi-Arid), and Basin 35 (Canning; Arid), which exhibit correlations ranging from 0.06 to 0.45.

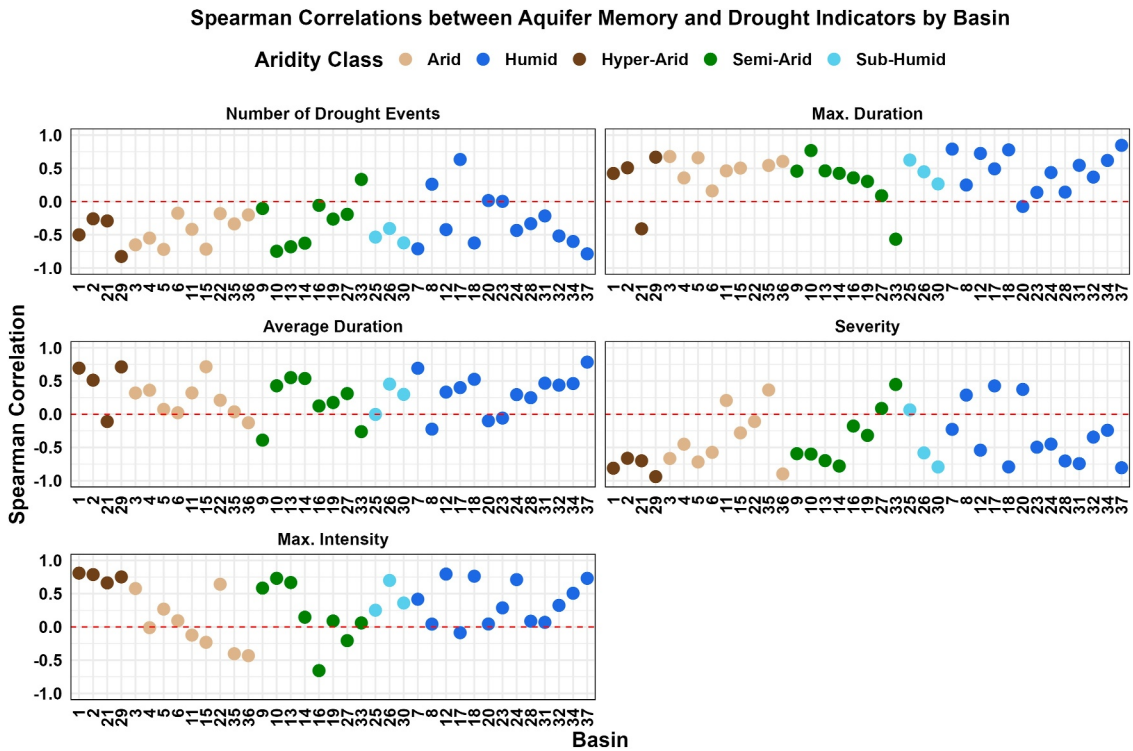

Figure 6. Spearman correlations depicting the relationship between aquifer memory and five drought indicators, including number of drought events, maximum duration, average duration, severity, and maximum intensity, across 37 study basins. Aquifers are ordered by aridity class, where Hyper-Arid regions are shown in dark brown, Arid in light brown (tan), Semi-Arid in green, Sub-Humid in sky blue, and Humid in medium blue (royal blue).

The analysis of maximum drought intensity reveals an overall positive association with aquifer memory (Figure 6). In hyper-arid basins, the average correlation is $\rho = 0.64$, while in arid, semi-arid, sub-humid, and humid basins, the average correlations are $\rho = 0.048$, $\rho = 0.27$, $\rho = 0.54$, and $\rho = 0.21$, respectively. This suggests that, in general, aquifer memory may contribute to higher peak drought intensities, possibly due to a delayed groundwater response that allows more severe water deficits to develop. Yet, this trend is not universal; significant outliers exist. For example, several arid basins—such as Basin 4 (Iulluemeden-Irhazer), Basin 11 (Stampriet-Kalahari), Basin 15 (Basin and Range Basin-fill), Basin 35 (Canning), and Basin 36 (Great Artesian)—along with Basin 16 (Ogallala; Semi-Arid), Basin 17 (Atlantic and Gulf Coastal Plains; Humid), and Basin 27 (Song-Liao; Semi-Arid), exhibit negative correlations ranging from −0.08 to −0.66, indicating that in these cases, aquifer memory may be associated with reduced maximum drought intensity.

4. Discussion

This study demonstrates that biases and variability in multi-model GRACE-GWA realizations, arising from differences among GRACE solutions and water budget assumptions, introduce uncertainty into GGDI interpretation (Figures 2–4). These findings reinforce the concerns raised by Akl and Thomas (2024), highlighting the necessity of a comprehensive evaluation of water budget components to enhance the accuracy and credibility of GRACE-groundwater based studies. Although normalization within the GGDI framework mitigates some amplitude-related discrepancies, it cannot fully resolve model divergences, particularly under extreme hydrological events or near-threshold conditions where even minor shifts in realization spreads can alter drought classification (Figure 2). The broad uncertainty ranges observed in groundwater drought indicators (Figure 3) and aquifer memory (Figure 4) across study basins highlight critical deficiencies in our current knowledge of groundwater drought dynamics.

Previous GGDI studies (Ali et al., 2022; Aon et al., 2024; Huang et al., 2023; Liu et al., 2022, 2025; Nandi & Biswas, 2024; Neves, 2024; Nigatu et al., 2024; Nikraftar et al., 2024; Paredes-Trejo et al., 2021; Satish Kumar et al., 2021; Song et al., 2024; Wang et al., 2020, 2022; Zhang et al., 2024; Zheng et al., 2024) relied on a single-

model approach to disaggregate GRACE-TWSA into GRACE-GWA. This strategy overlooks inherent model structure biases (Clark et al., 2008; Konapala et al., 2020; Ruddell et al., 2019) and underestimates the influence of surface water storage anomalies within the water balance equation (Han et al., 2009; Kim et al., 2009). In contrast, our multi-model analysis reveals that discrepancies in amplitude and timing across different, yet equally plausible, GRACE and water budget components data sets introduce substantial uncertainty into GGDI outcomes (Figures 3 and 4). Consequently, reliance on a single-model approach for deriving GRACE-GWA embeds critical biases into groundwater drought assessments, thereby compromising their overall robustness and reliability.

Rather than identifying a single "best" configuration, our analysis demonstrates that realization mean can mask critical variability and bias within groundwater drought indicators. For example, in Basin 1 (Nubian), the realization mean suggests six drought events, a maximum duration of 70 months, an average duration of 14 months, a severity of −92.50, a maximum intensity of −2.90, and an aquifer memory of 4 months. However, when the full range of realizations is considered, these values vary dramatically ranging from 1 to 10 events, 30 to 152 months in maximum duration, 10 to 72 months in average duration, −76.50 to −101 in severity, −2.00 to −7.40 in maximum intensity, and 3 to 61 months in aquifer memory. Such discrepancies highlight that meaningful hydrological information may be lost through averaging. By preserving the full realization spread, this framework provides a more transparent and probabilistic interpretation of groundwater drought, enhancing the reliability of GRACE-based drought assessments. Notably, our multi-model approach does not seek to select an optimal GRACE-GWA time series for GGDI analysis but takes the position that a multiple realization approach derives a comprehensive understanding of the groundwater system.

The considerable uncertainty in GGDI indicators observed across study basins highlights the challenge of accurately characterizing groundwater drought conditions (Figure 3). Notably, Basin 25 (Yakut) exhibited a variation of 11 drought events and a 63-month spread in average drought duration across realizations. In Basin 1 (Nubian), maximum drought duration estimates diverged by as much as 122 months, while drought severity and peak intensity varied by 24.47 and 5.39 units, respectively. Such discrepancies underscore the strong sensitivity of groundwater drought indicators to input variability, revealing the difficulty of achieving consistent and robust drought classifications.

Groundwater drought indicators are essential for informing resource allocation, aquifer recharge strategies, usage regulations, and long-term sustainability planning (e.g., Liu et al., 2025; Thomas, Caineta, & Nanteza, 2017; Zhang et al., 2024). However, failure to account for variability in groundwater time series increases the risk of misjudging groundwater deficits, potentially leading to either resource depletion due to underestimated drought severity or overly restrictive management actions that unnecessarily constrain water use. By explicitly capturing the spread among GRACE-based realizations, the multi-model framework transforms GGDI indicators into diagnostic tools that expose where assessments are robust and where uncertainty remains high. This diagnostic insight enhances policy relevance, enabling managers to gauge confidence in GRACE-derived indicators, identify vulnerable aquifers, and align intervention or monitoring strategies with the reliability of the supporting data. In doing so, the framework promotes transparent, evidence-informed groundwater governance and fosters more adaptive, resilient responses to drought.

The wide range of variability in aquifer memory estimates across study basins hampers our ability to anticipate groundwater response and implement efficient sustainable strategies (Figure 4). For example, Basin 1 (Nubian), Basin 3 (Lake Chad), Basin 6 (Senegalo-Mauretanian), and Basin 37 (Paris) exhibit aquifer memory ranges spanning from 3 to 61 months, 4 to 50 months, 12 to 71 months, and 7 to 68 months, respectively. Such variability makes it difficult to predict when an aquifer will recover from drought, complicating long-term planning. If aquifer memory is highly uncertain, it is unclear whether prolonged dry spells will quickly subside once wetter conditions return or persist for extended periods. This ambiguity can lead to either overconfidence or excessive caution in water-use restrictions.

Our findings suggest a surprisingly consistent relationship between aquifer memory and various groundwater drought indicators across study basins despite variability attributed to GRACE-GWA extraction (Figure 6). Generally, aquifer memory corresponds to fewer drought events and lower severity, while also correlating with extended drought durations and more intense peaks. However, certain basins depart from this pattern. For instance, Basin 33 (Taurus-Zagros; Semi-Arid) shows a contrasting trend, with a positive correlation between aquifer memory and drought frequency ($\rho = 0.33$), negative correlations with maximum ($\rho = -0.57$) and average ($\rho = -0.26$) drought duration, a positive correlation with drought severity ($\rho = 0.45$), and virtually no relationship

with maximum drought intensity ($\rho = 0.06$). Likewise, Basin 21 (Arabian; Hyper-Arid) presents negative correlations between aquifer memory and both maximum ($\rho = -0.42$) and average ($\rho = -0.11$) drought duration. Basin 33 is predominantly populated rangeland, while Basin 21 comprises barren, remote rangeland interspersed with populated areas (Richey et al., 2015). Both basins experience limited recharge and considerable groundwater pumping for agriculture, conditions that likely contribute to their atypical drought behavior.

Validation of large-scale GRACE-GWA is inherently constrained by the limited spatial and temporal coverage of in situ groundwater observations (Condon et al., 2021; Famiglietti et al., 2011; Giordano, 2009; Jasechko et al., 2024; Lall et al., 2020). Consequently, the practical value of the proposed framework lies in its ability to quantify and interpret uncertainty where direct validation is not feasible. Moreover, even in regions where in situ data exist, the framework enhances interpretability by revealing the full range of plausible groundwater responses rather than obscuring meaningful variability through realization averaging or model weighting. By maintaining the independence of each GRACE-TWSA and water balance component, the framework explicitly propagates methodological divergence through the GRACE-GWA estimation process, offering a transparent and defensible basis for policy-relevant groundwater drought assessment.

In this study, we retained the 2004–2009 baseline for computing anomalies in GRACE-TWSA and the associated water budget components, aligning with the official reference period established by major GRACE processing centers (Cooley & Landerer, 2019). Given that GRACE-GWA reflects deviations from the selected baseline rather than absolute groundwater storage, the baseline plays a pivotal role in shaping both the magnitude and interpretability of the resulting anomalies. This baseline choice introduces a structural artifact: an apparent agreement among GRACE-GWA realizations during the baseline years, not due to alignment in model behavior, but rather to the shared constraint of normalization. It is outside this reference window, before 2004 and after 2009, where model-specific assumptions, structural differences, and input data combinations exert greater influence, revealing the true extent of divergence across GRACE-GWA realizations. Although extending the anomaly baseline to is feasible, we intentionally applied the 2004–2009 baseline to allow model divergence to manifest more clearly across the remaining time series. This decision enhances the transparency of uncertainty propagation within the GGDI framework and enables a robust evaluation of how discrepancies in GRACE-TWSA and water budget components influence drought identification and classification.

One limitation of this study arises from the spatial resolution mismatch between the employed water budget components and GRACE products, requiring interpolation to align their grid resolution. While necessary, this adjustment introduces an additional layer of uncertainty. Furthermore, to minimize uncertainties, we excluded the data gap between the GRACE and GRACE-FO missions; however, this omission may inadvertently obscure critical information essential for a more comprehensive evaluation of groundwater drought indicators. Notably, the GRACE observation period was sufficient for aquifer memory calculations, ensuring that this gap did not affect the robustness of memory estimates. Although this study leveraged recent advancements in remote sensing to account for surface water across the study basins, data limitations persist in many regions where surface water bodies remain underrepresented (Figures S1 and S2 and Table S2 in Supporting Information S1). This issue is not unique to the present analysis but represents a broader, systemic challenge in global hydrological monitoring (Alsdorf et al., 2007; Busker et al., 2019; Tortini et al., 2020). As a result, GRACE-GWA estimates may still carry residual signals from unresolved surface water variability (Castle et al., 2014; Thomas and Nanteza, 2023), complicating efforts to isolate pure groundwater storage changes from mixed terrestrial water components (Akl et al., 2025; Akl & Thomas, 2024). Furthermore, as with any normalized index, GGDI results are inherently dependent on the chosen study period. Any modification to this period could alter GGDI calculations, thereby impacting the estimated range of uncertainty.

5. Conclusions

Reliable groundwater drought assessment is paramount, given groundwater's critical role during periods of water scarcity (Castle et al., 2014; Famiglietti, 2014; Hughes et al., 2012). In this study, we integrate multi-model GRACE-GWA, sourced from diverse GRACE-TWSA solutions and water budget data, within the GGDI framework across 37 study aquifers. By evaluating key five groundwater drought indicators, that is number of drought events, maximum duration, average duration, severity, and maximum intensity, our findings demonstrate that variability among multi-model GRACE-GWA introduces significant uncertainty to GGDI, thereby constraining the reliability of groundwater drought assessments. Furthermore, our exploration of aquifer memory,

quantified via the autocorrelation in GGDI time series, reveals substantial uncertainty, limiting our ability to accurately capture aquifer response and recovery dynamics, complicating the development of long-term water sustainability strategies. Additionally, this study uncovered the relation between aquifer memory and groundwater drought indicators documenting that aquifer memory correlates with fewer drought events yet results in longer and more intense drought episodes when they occur. These findings underscore the inherent uncertainty in large-scale groundwater drought assessments and reinforce the necessity of employing a multi-model GRACE-GWA approach within the GGDI framework. By capturing variability, this approach enhances uncertainty awareness and improves the reliability of groundwater drought assessments.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

CSR-SH (Landerer, 2021), JPL-SH (Landerer, 2021), GFZ-SH (Landerer, 2021), and JPL-M (Wiese et al., 2018) are accessible through JPL website at (https://grace.jpl.nasa.gov/data/get-data/monthly-mass-grids-land/; Last access: 03 December 2024). For the CSR-M (Save, 2020), visit CSR website (https://www2.csr.utexas.edu/grace/RL06_mascons.html; Last access: 03 December 2024). GLDAS (Rodell et al., 2004) and FLDAS (McNally et al., 2017) data sets are available from The NASA Goddard Earth Sciences Data and Information Services Centre (https://ldas.gsfc.nasa.gov/gldas/gldas-get-data; Last access: 09 December 2024). WaterGAP v2.2e (Müller Schmied et al., 2024) outputs can be retrieved from the Goethe University Data Repository (GUDe) at (https://gude.uni-frankfurt.de/entities/researchdata/7d500a84-9a36-4cd2-9f76-a4081f9ee94e/details; Last access: 09 December 2024). ERA5-Land data sets (Copernicus Climate Change Service, 2022) can be found on the Copernicus website at (https://cds.climate.copernicus.eu/datasets/reanalysis-era5-land-monthly-means?tab=overview; Last access: 09 December 2024). Lake/reservoir water storage data are downloadable from GloLakes (Hou et al., 2024) at (https://doi.org/10.25914/K8ZF-6G46; Last access: 08 June 2023) and from Copernicus (Copernicus Climate Change Service, 2020) at (https://cds.climate.copernicus.eu/datasets/satellite-lake-water-level?tab=overview; Last access: 15 December 2024).

Acknowledgments
This study was supported by a PhD scholarship (Grant ID: MM59/19) awarded to the first author by the Egyptian Ministry of Higher Education and Scientific Research, represented by the Egyptian Bureau for Cultural and Educational Affairs in London.

References

Akl, M., Thomas, B., & Mills, J. (2022). GRACE-Derived groundwater storage estimation: Lake/reservoir storage controls across Canada. In *EGU general assembly conference abstracts*. https://doi.org/10.5194/egusphere-egu22-2495.EGU22-2495

Akl, M., & Thomas, B. F. (2023). Evaluating the use of "Goodness-of-Fit" metrics in GRACE validation: GRACE accuracy for monitoring groundwater dynamics. *Authorea Preprints*. https://doi.org/10.22541/essoar.167388164.42151735/v1

Akl, M., & Thomas, B. F. (2024). Challenges in applying water budget framework for estimating groundwater storage changes from GRACE observations. *Journal of Hydrology*, *639*, 131600. https://doi.org/10.1016/j.jhydrol.2024.131600

Akl, M., Thomas, B. F., & Clarke, P. J. (2025). A multi-objective comparative framework for enhanced GRACE-groundwater comparative analysis. *Journal of Hydrology*, *664*, 134403. https://doi.org/10.1016/j.jhydrol.2025.134403

Ali, S., Liu, D., Fu, Q., Cheema, M. J. M., Pal, S. C., Arshad, A., et al. (2022). Constructing high-resolution groundwater drought at spatio-temporal scale using GRACE satellite data based on machine learning in the Indus Basin. *Journal of Hydrology*, *612*, 128295. https://doi.org/10.1016/j.jhydrol.2022.128295

Alsdorf, D. E., Rodríguez, E., & Lettenmaier, D. P. (2007). Measuring surface water from space. *Reviews of Geophysics*, *45*(2). https://doi.org/10.1029/2006RG000197

Amazirh, A., Ouassanouan, Y., Bouimouass, H., Baba, M. W., Bouras, E. H., Rafik, A., et al. (2024). Remote sensing-based multiscale analysis of total and groundwater storage dynamics over semi-arid North African basins. *Remote Sensing*, *16*(19), 3698. https://doi.org/10.3390/rs16193698

Aon, S., Nandi, S., Sen, S., & Biswas, S. (2024). GRACE based groundwater drought evaluation of Ganga Basin and analysis of drought propagation using wavelet based quantitative approach. *The Science of the Total Environment*, *951*, 175666. https://doi.org/10.1016/j.scitotenv.2024.175666

Awange, J. L., Forootan, E., Fleming, K., & Odhiambo, G. (2014). Dominant patterns of water storage changes in the Nile Basin during 2003–2013. *Remote Sensing of the Terrestrial Water Cycle*, 367–381. https://doi.org/10.1002/9781118872086.ch22

Beck, H. E., Pan, M., Miralles, D. G., Reichle, R. H., Dorigo, W. A., Hahn, S., et al. (2021). Evaluation of 18 satellite-and model-based soil moisture products using in situ measurements from 826 sensors. *Hydrology and Earth System Sciences*, *25*(1), 17–40. https://doi.org/10.5194/hess-25-17-2021

Bhanja, S. N., Mukherjee, A., Saha, D., Velicogna, I., & Famiglietti, J. S. (2016). Validation of GRACE based groundwater storage anomaly using in-situ groundwater level measurements in India. *Journal of Hydrology*, *543*, 729–738. https://doi.org/10.1016/j.jhydrol.2016.10.042

Bloomfield, J. P., & Marchant, B. P. (2013). Analysis of groundwater drought building on the standardised precipitation index approach. *Hydrology and Earth System Sciences*, *17*(12), 4769–4787. https://doi.org/10.5194/hess-17-4769-2013

Bloomfield, J. P., Marchant, B. P., Bricker, S. H., & Morgan, R. B. (2015). Regional analysis of groundwater droughts using hydrograph classification. *Hydrology and Earth System Sciences*, *19*(10), 4327–4344. https://doi.org/10.5194/hess-19-4327-2015

Busker, T., de Roo, A., Gelati, E., Schwatke, C., Adamovic, M., Bisselink, B., et al. (2019). A global lake and reservoir volume analysis using a surface water dataset and satellite altimetry. *Hydrology and Earth System Sciences*, *23*(2), 669–690. https://doi.org/10.5194/hess-23-669-2019

Castle, S. L., Thomas, B. F., Reager, J. T., Rodell, M., Swenson, S. C., & Famiglietti, J. S. (2014). Groundwater depletion during drought threatens future water security of the Colorado River Basin. *Geophysical Research Letters*, *41*(16), 5904–5911. https://doi.org/10.1002/2014GL061055

Clark, M. P., Slater, A. G., Rupp, D. E., Woods, R. A., Vrugt, J. A., Gupta, H. V., et al. (2008). Framework for understanding structural errors (FUSE): A modular framework to diagnose differences between hydrological models. *Water Resources Research*, *44*(12). https://doi.org/10.1029/2007WR006735

Coles, A. E., McConkey, B. G., & McDonnell, J. J. (2017). Climate change impacts on hillslope runoff on the northern Great Plains, 1962–2013. *Journal of Hydrology*, *550*, 538–548. https://doi.org/10.1016/j.jhydrol.2017.05.023

Condon, L. E., Kollet, S., Bierkens, M. F., Fogg, G. E., Maxwell, R. M., Hill, M. C., et al. (2021). Global groundwater modeling and monitoring: Opportunities and challenges. *Water Resources Research*, *57*(12), e2020WR029500. https://doi.org/10.1029/2020WR029500

Cooley, S. S., & Landerer, F. W. (2019). *Gravity recovery and climate experiment Follow-on (GRACE-FO) Level-3 data product user handbook* (pp. 1–57). Jet Propulsion Laboratory California Institute of Technology: Pasadena, CA, USA. Retrieved from https://podaac.jpl.nasa.gov/gravity/grace-documentation

Copernicus Climate Change Service. (2020). Lake water levels from 1992 to present derived from satellite observation [Dataset]. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*. https://doi.org/10.24381/cds.5714c668

Copernicus Climate Change Service. (2022). ERA5-Land monthly averaged data from 1950 to present [Dataset]. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*. https://doi.org/10.24381/cds.68d2bb30

der Van Gun, J. (2022). Large aquifer systems around the world. *The Groundwater Project, Guelph, ON*. https://doi.org/10.21083/978-1-77470-020-4

Dracup, J. A., Lee, K. S., & Paulson Jr, E. G. (1980). On the statistical characteristics of drought events. *Water Resources Research*, *16*(2), 289–296. https://doi.org/10.1029/WR016i002p00289

Duan, Z., & Bastiaanssen, W. G. M. (2013). Estimating water volume variations in lakes and reservoirs from four operational satellite altimetry databases and satellite imagery data. *Remote Sensing of Environment*, *134*, 403–416. https://doi.org/10.1016/j.rse.2013.03.010

Dubois, E., & Larocque, M. (2024). Contribution of standardized indexes to understand groundwater level fluctuations in response to meteorological conditions in cold and humid climates. *Journal of Hydrology*, *634*, 131105. https://doi.org/10.1016/j.jhydrol.2024.131105

Dumanski, S., Pomeroy, J. W., & Westbrook, C. J. (2015). Hydrological regime changes in a Canadian Prairie basin. *Hydrological Processes*, *29*(18), 3893–3904. https://doi.org/10.1002/hyp.10567

Entekhabi, D., Njoku, E. G., O'neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., et al. (2010). The soil moisture active passive (SMAP) mission. *Proceedings of the IEEE*, *98*(5), 704–716. https://doi.org/10.1109/JPROC.2010.2043918

Famiglietti, J. S. (2014). The global groundwater crisis. *Nature Climate Change*, *4*(11), 945–948. https://doi.org/10.1038/nclimate2425

Famiglietti, J. S., Lo, M., Ho, S. L., Bethune, J., Anderson, K. J., Syed, T. H., et al. (2011). Satellites measure recent rates of groundwater depletion in California's Central Valley. *Geophysical Research Letters*, *38*(3). https://doi.org/10.1029/2010GL046442

Famiglietti, J. S., & Rodell, M. (2013). Water in the balance. *Science*, *340*(6138), 1300–1301. https://doi.org/10.1126/science.1236460

Fan, Y., Li, H., & Miguez-Macho, G. (2013). Global patterns of groundwater table depth. *Science*, *339*(6122), 940–943. https://doi.org/10.1126/science.1229881

Fatolazadeh, F., Voosoghi, B., & Naeeni, M. R. (2016). Wavelet and Gaussian approaches for estimation of groundwater variations using GRACE data. *Groundwater*, *54*(1), 74–81. https://doi.org/10.1111/gwat.12325

Gao, H., Birkett, C., & Lettenmaier, D. P. (2012). Global monitoring of large reservoir storage from satellite remote sensing. *Water Resources Research*, *48*(9). https://doi.org/10.1029/2012WR012063

Giordano, M. (2009). Global groundwater? Issues and solutions. *Annual Review of Environment and Resources*, *34*(1), 153–178. https://doi.org/10.1146/annurev.environ.030308.100251

Gleeson, T., Wada, Y., Bierkens, M. F., & Van Beek, L. P. (2012). Water balance of global aquifers revealed by groundwater footprint. *Nature*, *488*(7410), 197–200. https://doi.org/10.1038/nature11295

Goodarzi, M., Abedi-Koupai, J., Heidarpour, M., & Safavi, H. R. (2016). Development of a new drought index for groundwater and its application in sustainable groundwater extraction. *Journal of Water Resources Planning and Management*, *142*(9), 04016032. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000673

Guo, J., Mu, D., Liu, X., Yan, H., Sun, Z., & Guo, B. (2016). Water storage changes over the Tibetan Plateau revealed by GRACE mission. *Acta Geophysica*, *64*(2), 463–476. https://doi.org/10.1515/acgeo-2016-0003

Han, S. C., Kim, H., Yeo, I. Y., Yeh, P., Oki, T., Seo, K. W., et al. (2009). Dynamics of surface water storage in the Amazon inferred from measurements of inter-satellite distance change. *Geophysical Research Letters*, *36*(9). https://doi.org/10.1029/2009GL037910

Hou, J., Van Dijk, A. I., Renzullo, L. J., & Larraondo, P. R. (2024). GloLakes: Water storage dynamics for 27 000 lakes globally from 1984 to present derived from satellite altimetry and optical imaging [Dataset]. *Earth System Science Data*, *16*(1), 201–218. https://doi.org/10.5194/essd-16-201-2024

Houborg, R., Rodell, M., Li, B., Reichle, R., & Zaitchik, B. F. (2012). Drought indicators based on model-assimilated gravity recovery and climate experiment (GRACE) terrestrial water storage observations. *Water Resources Research*, *48*(7). https://doi.org/10.1029/2011WR011291

Huang, J., Cao, L., Wang, L., Liu, L., Yu, B., & Han, L. (2023). Identification and spatiotemporal migration analysis of groundwater drought events in the North China Plain. *Atmosphere*, *14*(6), 961. https://doi.org/10.3390/atmos14060961

Hughes, J. D., Petrone, K. C., & Silberstein, R. P. (2012). Drought, groundwater storage and stream flow decline in southwestern Australia. *Geophysical Research Letters*, *39*(3). https://doi.org/10.1029/2011GL050797

Jasechko, S., Birks, S. J., Gleeson, T., Wada, Y., Fawcett, P. J., Sharp, Z. D., et al. (2014). The pronounced seasonality of global groundwater recharge. *Water Resources Research*, *50*(11), 8845–8867. https://doi.org/10.1002/2014WR015809

Jasechko, S., Seybold, H., Perrone, D., Fan, Y., Shamsudduha, M., Taylor, R. G., et al. (2024). Rapid groundwater decline and some cases of recovery in aquifers globally. *Nature*, *625*(7996), 715–721. https://doi.org/10.1038/s41586-023-06879-8

Khazaei, B., Read, L. K., Casali, M., Sampson, K. M., & Yates, D. N. (2022). GLOBathy, the global lakes bathymetry dataset. *Scientific Data*, *9*(1), 36. https://doi.org/10.1038/s41597-022-01132-9

Kim, H., Yeh, P. J. F., Oki, T., & Kanae, S. (2009). Role of rivers in the seasonal variations of terrestrial water storage over global basins. *Geophysical Research Letters*, *36*(17). https://doi.org/10.1029/2009GL039006

Klees, R., Zapreeva, E. A., Winsemius, H. C., & Savenije, H. H. G. (2007). The bias in GRACE estimates of continental water storage variations. *Hydrology and Earth System Sciences*, *11*(4), 1227–1241. https://doi.org/10.5194/hess-11-1227-2007

Konapala, G., Kao, S. C., & Addor, N. (2020). Exploring hydrologic model process connectivity at the continental scale through an information theory approach. *Water Resources Research*, *56*(10), e2020WR027340. https://doi.org/10.1029/2020WR027340

Lall, U., Josset, L., & Russo, T. (2020). A snapshot of the world's groundwater challenges. *Annual Review of Environment and Resources*, *45*(1), 171–194. https://doi.org/10.1146/annurev-environ-102017-025800

Landerer, F. (2021). TELLUS_GRAC_L3_CSR_RL06_LND_v04 (Version RL06 v04) [Dataset]. *PO.DAAC*. https://doi.org/10.5067/TELND-3AC64

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

Landerer, F. W., & Swenson, S. C. (2012). Accuracy of scaled GRACE terrestrial water storage estimates. *Water Resources Research*, *48*(4). https://doi.org/10.1029/2011WR011453

Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., et al. (2011). Parameterization improvements and functional and structural advances in version 4 of the community land model. *Journal of Advances in Modeling Earth Systems*, *3*(1), M03001. https://doi.org/10.1029/2011ms000045

Lehner, B., Verdin, K., & Jarvis, A. (2008). New global hydrography derived from spaceborne elevation data. *Eos, Transactions American Geophysical Union*, *89*(10), 93–94. https://doi.org/10.1029/2008EO100001

Lenk, O. (2013). Satellite based estimates of terrestrial water storage variations in Turkey. *Journal of Geodynamics*, *67*, 106–110. https://doi.org/10.1016/j.jog.2012.04.010

Li, Q., Pan, Y., Zhang, C., & Gong, H. (2023). Quantifying multi-source uncertainties in GRACE-based estimates of groundwater storage changes in Mainland China. *Remote Sensing*, *15*(11), 2744. https://doi.org/10.3390/rs15112744

Liu, M., Pei, H., & Shen, Y. (2022). Evaluating dynamics of GRACE groundwater and its drought potential in Taihang Mountain Region, China. *Journal of Hydrology*, *612*, 128156. https://doi.org/10.1016/j.jhydrol.2022.128156

Liu, X., Wang, N., Wang, Y., Meng, N., Wang, Y., Qiao, B., et al. (2025). Research on groundwater drought and sustainability in Badain Jaran Desert and surrounding areas based on GRACE satellite. *Land*, *14*(1), 173. https://doi.org/10.3390/land14010173

Longuevergne, L., Wilson, C. R., Scanlon, B. R., & Crétaux, J. F. (2013). GRACE water storage estimates for the Middle East and other regions with significant reservoir and lake storage. *Hydrology and Earth System Sciences*, *17*(12), 4817–4830. https://doi.org/10.5194/hess-17-4817-2013

Margat, J. (2007). Which crises in the 21st century? *Making Peace with the Earth: What Future for the Human Species and the Planet*.

Margat, J. (2008). *Exploitation and use of groundwater in the world* (Co-edition). UNESCO and BRGM, 52.

Margat, J., & der Van Gun, J. (2013). *Groundwater around the world: A geographic synopsis*. Crc Press.

McNally, A., Arsenault, K., Kumar, S., Shukla, S., Peterson, P., Wang, S., et al. (2017). A land data assimilation system for Sub-Saharan Africa food and water security applications [Dataset]. *Scientific Data*, *4*(1), 1–19. https://doi.org/10.1038/sdata.2017.12

Mishra, A. K., & Singh, V. P. (2010). A review of drought concepts. *Journal of Hydrology*, *391*(1–2), 202–216. https://doi.org/10.1016/j.jhydrol.2010.07.012

Müller Schmied, H., Trautmann, T., Ackermann, S., Cáceres, D., Flörke, M., Gerdener, H., et al. (2024). The global water resources and use model WaterGAP v2. 2e: Description and evaluation of modifications and new features [Dataset]. *Geoscientific Model Development*, *17*(23), 8817–8852. https://doi.org/10.5194/gmd-17-8817-2024

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., et al. (2021). ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data*, *13*(9), 4349–4383. https://doi.org/10.5194/essd-13-4349-2021

Nandi, S., & Biswas, S. (2024). Spatiotemporal distribution of groundwater drought using GRACE-based satellite estimates: A case study of lower Gangetic Basin, India. *Environmental Monitoring and Assessment*, *196*(2), 151. https://doi.org/10.1007/s10661-024-12309-7

Nanteza, J., de Linage, C. R., Thomas, B. F., & Famiglietti, J. S. (2016). Monitoring groundwater storage changes in complex basement aquifers: An evaluation of the GRACE satellites over E ast a frica. *Water Resources Research*, *52*(12), 9542–9564. https://doi.org/10.1002/2016WR018846

Neves, M. C. (2024). Integrating standardized indices and performance indicators for better drought assessment in semi-arid coastal aquifers. *Groundwater for Sustainable Development*, *27*, 101341. https://doi.org/10.1016/j.gsd.2024.101341

Neves, M. C., Nunes, L. M., & Monteiro, J. P. (2020). Evaluation of GRACE data for water resource management in Iberia: A case study of groundwater storage monitoring in the Algarve region. *Journal of Hydrology: Regional Studies*, *32*, 100734. https://doi.org/10.1016/j.ejrh.2020.100734

Nigatu, Z. M., You, W., & Melesse, A. M. (2024). Drought dynamics in the Nile River Basin: Meteorological, agricultural, and groundwater drought propagation. *Remote Sensing*, *16*(5), 919. https://doi.org/10.3390/rs16050919

Nikraftar, Z., Parizi, E., Saber, M., Hosseini, S. M., Ataie-Ashtiani, B., & Simmons, C. T. (2024). Groundwater sustainability assessment in the Middle East using GRACE/GRACE-FO data. *Hydrogeology Journal*, *32*(1), 321–337. https://doi.org/10.1007/s10040-023-02717-3

Nilsson, C., Reidy, C. A., Dynesius, M., & Revenga, C. (2005). Fragmentation and flow regulation of the world's large river systems. *Science*, *308*(5720), 405–408. https://doi.org/10.1126/science.1107887

Opie, S., Taylor, R. G., Brierley, C. M., Shamsudduha, M., & Cuthbert, M. O. (2020). Climate–groundwater dynamics inferred from GRACE and the role of hydraulic memory. *Earth System Dynamics*, *11*(3), 775–791. https://doi.org/10.5194/esd-11-775-2020

Panda, D. K., Mishra, A., Jena, S. K., James, B. K., & Kumar, A. (2007). The influence of drought and anthropogenic effects on groundwater levels in Orissa, India. *Journal of Hydrology*, *343*(3–4), 140–153. https://doi.org/10.1016/j.jhydrol.2007.06.007

Paredes-Trejo, F., Barbosa, H. A., Giovannettone, J., Kumar, T. L., Thakur, M. K., Buriti, C. D. O., & Uzcátegui-Briceño, C. (2021). Drought assessment in the São Francisco River Basin using satellite-based and ground-based indices. *Remote Sensing*, *13*(19), 3921. https://doi.org/10.3390/rs13193921

Rateb, A., Scanlon, B. R., Pool, D. R., Sun, A., Zhang, Z., Chen, J., et al. (2020). Comparison of groundwater storage changes from GRACE satellites with monitoring and modeling of major US aquifers. *Water Resources Research*, *56*(12), e2020WR027556. https://doi.org/10.1029/2020WR027556

Richey, A. S., Thomas, B. F., Lo, M. H., Reager, J. T., Famiglietti, J. S., Voss, K., et al. (2015). Quantifying renewable groundwater stress with GRACE. *Water Resources Research*, *51*(7), 5217–5238. https://doi.org/10.1002/2015WR017349

Rodell, M., & Famiglietti, J. S. (1999). Detectability of variations in continental water storage from satellite observations of the time dependent gravity field. *Water Resources Research*, *35*(9), 2705–2723. https://doi.org/10.1029/1999WR900141

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

Rodell, M., Famiglietti, J. S., Wiese, D. N., Reager, J. T., Beaudoing, H. K., Landerer, F. W., & Lo, M. H. (2018). Emerging trends in global freshwater availability. *Nature*, *557*(7707), 651–659. https://doi.org/10.1038/s41586-018-0123-1

Rodell, M., Houser, P. R., Jambor, U. E. A., Gottschalck, J., Mitchell, K., Meng, C. J., et al. (2004). The global land data assimilation system [Dataset]. *Bulletin of the American Meteorological Society*, *85*(3), 381–394. https://doi.org/10.1175/BAMS-85-3-381

Ruddell, B. L., Drewry, D. T., & Nearing, G. S. (2019). Information theory for model diagnostics: Structural error is indicated by trade-off between functional and predictive performance. *Water Resources Research*, *55*(8), 6534–6554. https://doi.org/10.1029/2018WR023692

Satish Kumar, K., AnandRaj, P., Sreelatha, K., Bisht, D. S., & Sridhar, V. (2021). Monthly and seasonal drought characterization using grace-based groundwater drought index and its link to teleconnections across south Indian river basins. *Climate*, *9*(4), 56. https://doi.org/10.3390/cli9040056

Save, H. (2020). Csr grace and grace-fo rl06 mascon solutions v02 [Dataset]. *University of Texas at Austin*, *1*. http://www2.csr.utexas.edu/grace

Save, H., Bettadpur, S., & Tapley, B. D. (2012). Reducing errors in the GRACE gravity solutions using regularization. *Journal of Geodesy*, *86*(9), 695–711. https://doi.org/10.1007/s00190-012-0548-5

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

Saxe, S., Farmer, W., Driscoll, J., & Hogue, T. S. (2021). Implications of model selection: A comparison of publicly available, conterminous US-extent hydrologic component estimates. *Hydrology and Earth System Sciences*, *25*(3), 1529–1568. https://doi.org/10.5194/hess-25-1529-2021

Scanlon, B. R., Faunt, C. C., Longuevergne, L., Reedy, R. C., Alley, W. M., McGuire, V. L., & McMahon, P. B. (2012). Groundwater depletion and sustainability of irrigation in the US high plains and central Valley. *Proceedings of the National Academy of Sciences*, *109*(24), 9320–9325. https://doi.org/10.1073/pnas.1200311109

Scanlon, B. R., Rateb, A., Anyamba, A., Kebede, S., MacDonald, A. M., Shamsudduha, M., et al. (2022). Linkages between GRACE water storage, hydrologic extremes, and climate teleconnections in major African aquifers. *Environmental Research Letters*, *17*(1), 014046. https://doi.org/10.1088/1748-9326/ac3bfc

Scanlon, B. R., Rateb, A., Pool, D. R., Sanford, W., Save, H., Sun, A., et al. (2021). Effects of climate and irrigation on GRACE-based estimates of water storage changes in major US aquifers. *Environmental Research Letters*, *16*(9), 094009. https://doi.org/10.1088/1748-9326/ac16ff

Scanlon, B. R., Stonestrom, D. A., Reedy, R. C., Leaney, F. W., Gates, J., & Cresswell, R. G. (2009). Inventories and mobilization of unsaturated zone sulfate, fluoride, and chloride related to land use change in semiarid regions, southwestern United States and Australia. *Water Resources Research*, *45*(7). https://doi.org/10.1029/2008WR006963

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

Shamsudduha, M., & Taylor, R. G. (2020). Groundwater storage dynamics in the world's large aquifer systems from GRACE: Uncertainty and role of extreme precipitation. *Earth System Dynamics*, *11*(3), 755–774. https://doi.org/10.5194/esd-11-755-2020

Siebert, S., Burke, J., Faures, J. M., Frenken, K., Hoogeveen, J., Döll, P., & Portmann, F. T. (2010). Groundwater use for irrigation–a global inventory. *Hydrology and Earth System Sciences*, *14*(10), 1863–1880. https://doi.org/10.5194/hess-14-1863-2010

Song, X., Chen, H., Chen, T., Qin, Z., Chen, S., Yang, N., & Deng, S. (2024). GRACE-based groundwater drought in the Indochina Peninsula during 1979–2020: Changing properties and possible teleconnection mechanisms. *The Science of the Total Environment*, *908*, 168423. https://doi.org/10.1016/j.scitotenv.2023.168423

Swenson, S., & Wahr, J. (2002). Methods for inferring regional surface-mass anomalies from gravity recovery and climate experiment (GRACE) measurements of time-variable gravity. *Journal of Geophysical Research*, *107*(B9), ETG–3. https://doi.org/10.1029/2001JB000576

Swenson, S., & Wahr, J. (2006). Post-processing removal of correlated errors in GRACE data. *Geophysical Research Letters*, *33*(8). https://doi.org/10.1029/2005GL025285

Swenson, S. C., & Lawrence, D. M. (2015). A GRACE-based assessment of interannual groundwater dynamics in the Community L and Model. *Water Resources Research*, *51*(11), 8817–8833. https://doi.org/10.1002/2015WR017582

Tallaksen, L. M. & Van Lanen, H. A. (Eds.) (2023)., *Hydrological drought: Processes and estimation methods for streamflow and groundwater*.

Taylor, R. G., Scanlon, B., Döll, P., Rodell, M., Van Beek, R., Wada, Y., et al. (2013). Ground water and climate change. *Nature Climate Change*, *3*(4), 322–329. https://doi.org/10.1038/nclimate1744

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. https://doi.org/10.1002/2014GL059323

Thomas, B. F., Caineta, J., & Nanteza, J. (2017). Global assessment of groundwater sustainability based on storage anomalies. *Geophysical Research Letters*, *44*(22), 11–445. https://doi.org/10.1002/2017GL076005

Thomas, B. F., & Famiglietti, J. S. (2019). Identifying climate-induced groundwater depletion in GRACE observations. *Scientific Reports*, *9*(1), 4124. https://doi.org/10.1038/s41598-019-40155-y

Thomas, B. F., Famiglietti, J. S., Landerer, F. W., Wiese, D. N., Molotch, N. P., & Argus, D. F. (2017). GRACE groundwater drought index: Evaluation of California Central Valley groundwater drought. *Remote Sensing of Environment*, *198*, 384–392. https://doi.org/10.1016/j.rse.2017.06.026

Thomas, B. F., & Nanteza, J. (2023). Global assessment of the sensitivity of water storage to hydroclimatic variations. *Science of the Total Environment*, *879*, 162958. https://doi.org/10.1016/j.scitotenv.2023.162958

Tortini, R., Noujdina, N., Yeo, S., Ricko, M., Birkett, C. M., Khandelwal, A., et al. (2020). Satellite-based remote sensing data set of global surface water storage change from 1992 to 2018. *Earth System Science Data*, *12*(2), 1141–1151. https://doi.org/10.5194/essd-12-1141-2020

Van Loon, A. F., Stahl, K., Di Baldassarre, G., Clark, J., Rangecroft, S., Wanders, N., & Van Lanen, H. A. (2016). Drought in a human-modified world: Reframing drought definitions, understanding, and analysis approaches. *Hydrology and Earth System Sciences*, *20*(9), 3631–3650. https://doi.org/10.5194/nhess-24-3173-2024

Wang, F., Lai, H., Li, Y., Feng, K., Zhang, Z., Tian, Q., et al. (2022). Identifying the status of groundwater drought from a GRACE mascon model perspective across China during 2003–2018. *Agricultural Water Management*, *260*, 107251. https://doi.org/10.1016/j.agwat.2021.107251

Wang, F., Wang, Z., Yang, H., Di, D., Zhao, Y., & Liang, Q. (2020). Utilizing GRACE-based groundwater drought index for drought characterization and teleconnection factors analysis in the North China Plain. *Journal of Hydrology*, *585*, 124849. https://doi.org/10.1016/j.jhydrol.2020.124849

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

Wei, L., Jiang, S., Ren, L., Tan, H., Ta, W., Liu, Y., et al. (2021). Spatiotemporal changes of terrestrial water storage and possible causes in the closed Qaidam Basin, China using GRACE and GRACE Follow-On data. *Journal of Hydrology*, *598*, 126274. https://doi.org/10.1016/j.jhydrol.2021.126274

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. https://doi.org/10.1002/2016WR019344

Wiese, D. N., Yuan, D. N., Boening, C., Landerer, F. W., & Watkins, M. M. (2018). JPL GRACE mascon ocean, ice, and hydrology equivalent water height release 06 coastal resolution improvement (CRI) filtered version 1.0 [Dataset]. *DAAC: Pasadena, CA, USA*. https://doi.org/10.5067/TEMSC-3MJC6

Wilhite, D. A., & Glantz, M. H. (1985). Understanding: The drought phenomenon: The role of definitions. *Water International*, *10*(3), 111–120. https://doi.org/10.1080/02508068508686328

Wilhite, D. A., Sivakumar, M. V., & Pulwarty, R. (2014). Managing drought risk in a changing climate: The role of national drought policy. *Weather and Climate Extremes*, *3*, 4–13. https://doi.org/10.1016/j.wace.2014.01.002

Zhang, Q., Sun, J., Dai, C., Zhang, G., & Wu, Y. (2024). Sustainable development of groundwater resources under the large-scale conversion of dry land into rice fields. *Agricultural Water Management*, *298*, 108851. https://doi.org/10.1016/j.agwat.2024.108851

Zheng, W., Askari, K., Song, C., Shi, P., Ge, W., Shi, S., et al. (2024). Increasing vulnerability of vegetation to meteorological and groundwater drought: A case study in Argentina. *Journal of Hydrology: Regional Studies*, *55*, 101931. https://doi.org/10.1016/j.ejrh.2024.101931

Zomer, R. J., Xu, J., & Trabucco, A. (2022). Version 3 of the global aridity index and potential evapotranspiration database. *Scientific Data*, *9*(1), 409. https://doi.org/10.1038/s41597-022-01493-1