

1 **Reconstruction of Climate-Driven Global Terrestrial Water Storage**  
2 **Variations (2002–2021) Using a Four-Parameter Linear Recursive**  
3 **Model**

4 **Pu Xie, Shuang Yi\***

5 National Key Laboratory of Earth System Numerical Modeling and Application,  
6 College of Earth and Planetary Sciences, University of Chinese Academy of  
7 Sciences, Beijing, 101408, China

8 \* Corresponding email: [s.yi@ucas.ac.cn](mailto:s.yi@ucas.ac.cn)

9

## 10 **Abstract**

11 Terrestrial water storage anomalies (TWSA), jointly influenced by climatic variability and  
12 human activities, exhibits pronounced fluctuations across multiple temporal scales. A substantial  
13 portion of the fluctuations is attributed to climatic variability, like the El Niño–Southern Oscillation  
14 (ENSO). Empirical reconstruction of climate-driven water storage based on relationships between  
15 GRACE satellite gravity observations and meteorological forcing data has become a common  
16 approach; however, existing models often neglect the regulating role of temperature in the  
17 transformation of precipitation into water storage. In this study, we propose a linear, four-parameter  
18 coupled recursive model that explicitly incorporates temperature effects on both the conversion and  
19 dissipation efficiency of water storage. Using GRACE/GRACE-FO satellite observations and  
20 meteorological forcing data, we reconstructed climate-driven TWSA over the global land grid  
21 (excluding Antarctica) at a monthly temporal resolution and 0.5° spatial resolution for the period  
22 2002 to 2021. For 116 major global river basins, we further derived basin-scale TWSA  
23 reconstructions and quantitatively evaluated the fraction of precipitation converted into TWSA.  
24 Compared with existing statistical reconstruction products, the results indicate that: (1) the proposed  
25 method achieves substantially faster parameter convergence, improving computational efficiency  
26 by several tens of times during the TWSA reconstruction process; (2) the proposed model  
27 demonstrates superior performance in approximately 89% of river basins and 62% of global land  
28 grid cells. Additional comparisons with the physically based Catchment Land Surface Model  
29 (CLSM) product from NASA’s Global Land Data Assimilation System (GLDAS) show that the  
30 proposed method better captures the temporal variability of GRACE TWSA in most basins. At the  
31 daily scale, the reconstructed TWSA agrees well overall with the ITSG-Grace2018 daily solution  
32 and GLDAS-2.2. This study enhances the understanding of the mechanisms governing terrestrial  
33 water storage variations at both global and regional scales, provides a quantitative assessment of  
34 climate-driven water storage changes, and offers a solid foundation for disentangling the respective  
35 impacts of climatic variability and human activities on water resources.

36 **Keywords: GRACE, GRACE-FO, Reconstruction, TWSA, Climate variability**



## 38 **1 Introduction**

39 Terrestrial water storage (TWS) is the total amount of water stored in all components, including  
40 surface water (e.g., lakes and reservoirs), groundwater, soil moisture, snow and ice, and vegetation  
41 water, and it plays a vital role in the global hydrological cycle. Monitoring TWS is essential not only  
42 for understanding the climate system and its variability (Tapley et al., 2019), but also for supporting  
43 industrial development and ensuring global food security (Abbott et al., 2019; Rodell et al., 2009;  
44 Rodell et al., 2018). Temperature and precipitation are the primary climatic drivers of TWS  
45 variations. For instance, precipitation anomalies induced by ENSO are a dominant cause of  
46 interannual TWS variability (Liu et al., 2020; Ni et al., 2018).

47 Traditionally, changes in TWS have been estimated using two main approaches. The first is the  
48 water balance method, which requires precipitation, evapotranspiration, and runoff data.  
49 Precipitation is typically measured by rain gauges, evapotranspiration is monitored using eddy  
50 covariance systems (Baldocchi et al., 2001; Pastorello et al., 2020; Ma et al., 2024), and runoff is  
51 estimated from discharge records at hydrological gauging stations (Duan et al., 2024). The rate of  
52 TWS change ( $dTWS/dt$ ) is then inferred from the water balance equation. However, because  
53 ground-based observation stations are spatially discrete, this method can only provide localized,  
54 point-scale information. In regions constrained by limited financial resources or political barriers to  
55 data access, the sparse distribution of monitoring sites further hinders the establishment of  
56 continuous spatial coverage (Crochemore et al., 2020; Fekete et al., 2012; Laudon et al., 2017). The  
57 second is the model-based approach, in which land surface or hydrological models simulate  
58 individual storage components such as soil moisture and groundwater, which are then aggregated to  
59 estimate total TWS variations. However, both instrumental and measurement errors can propagate  
60 through the modeling process, introducing substantial uncertainties into the final TWS estimates  
61 (Long et al., 2015; Humphrey et al., 2023). Moreover, the uncertainty associated with hydrological  
62 model outputs remains difficult to quantify (Zheng et al., 2023).

63 Since its launch in 2002, the Gravity Recovery and Climate Experiment (GRACE) mission has  
64 provided precise measurements of temporal variations in Earth's gravity field and delivered monthly  
65 gravity field solutions (Tapley et al., 2019; Chen et al., 2022; Humphrey et al., 2023). Compared  
66 with traditional approaches that rely on sparse ground-based observations or hydrological models  
67 with large uncertainties, GRACE data have substantially improved the spatial continuity and  
68 observational reliability of Earth's gravity measurements, providing a crucial tool for monitoring  
69 global water storage changes. A major contributor to temporal variations in the terrestrial gravity  
70 field is the variations in TWS, which is jointly driven by natural climate variability, long-term

71 climate change, and human water use (Rodell et al., 2018; An et al., 2021). Distinguishing  
72 anthropogenic signals from natural variability in GRACE-observed TWS has therefore become an  
73 active research focus (An et al., 2021; Liu et al., 2021; Yi et al., 2016; Zhang et al., 2025).

74 Currently, reconstruction methods for climate-driven TWSA can be broadly classified into two  
75 categories. The first category is data-driven approaches, which establish nonlinear mappings  
76 between GRACE-based TWSA and climate–hydrological variables using machine learning or  
77 neural network models (Fan et al., 2021; Palazzoli et al., 2025; Li et al., 2021; Sun et al., 2020;  
78 Zhang et al., 2016; Long et al., 2014; Sun et al., 2021; Yin et al., 2023; Li et al., 2020). The advantage  
79 of such approaches lies in their ability to capture complex patterns without requiring prior  
80 assumptions about the underlying physical processes (Long et al., 2014) and without being  
81 constrained by the spatial resolution of the input drivers. For instance, Zhang et al. (2016)  
82 reconstructed TWSA in the Yangtze River Basin for the period from 1979 to 2012 using an artificial  
83 neural network driven by precipitation and ERA-Interim soil moisture, and subsequently estimated  
84 drought recovery times. Sun et al. (2021) introduced an automated machine learning framework for  
85 GRACE-based TWSA reconstruction, integrating GLDAS model outputs together with  
86 meteorological and climatic predictors to combine physical constraints with data-driven learning.  
87 Palazzoli et al. (2025) developed unidirectional and bidirectional Long Short-Term Memory  
88 networks to construct four climate-driven models, generating a continuous global TWSA record at  
89 0.5° resolution for the period from 1984 to 2021. Li et al. (2021) combined statistical analysis, time  
90 series decomposition, and machine learning techniques to reconstruct global (excluding Antarctica)  
91 gridded TWSA for the period from 1979 to 2020 using multiple predictors, such as sea surface  
92 temperature and climate indices. Yin et al. (2023) also adopted this approach and incorporated a  
93 broader range of predictors in addition to conventional climatic variables, including land use data  
94 and vegetation indices, to reconstruct a continuous global monthly TWSA dataset at 0.5° spatial  
95 resolution for the period 1940 to 2022.

96 The second category of methods involves data-driven statistical models for reconstructing  
97 TWSA (Zhong et al., 2019; Zhong et al., 2025; Humphrey and Gudmundsson, 2019; Xiao et al.,  
98 2025). Humphrey and Gudmundsson (2019) proposed a statistical reconstruction approach that uses  
99 only precipitation and temperature as predictors to estimate climate-driven TWS at both daily and  
100 monthly time scales over the past century. Their method not only successfully filled the ~1-year  
101 observational gap between the GRACE and GRACE-FO missions but also demonstrated  
102 performance comparable to or exceeding that of several global hydrological models. However, their  
103 model incorporates a mathematical fitting procedure to align the reconstructed TWSA with GRACE

104 observations. This adjustment improves amplitude consistency but does not provide a clear physical  
 105 interpretation. Building on this framework, Zhong et al. (2025) introduced a key modification by  
 106 incorporating the concept of the “mean daily fraction of precipitation converted to TWS”, thereby  
 107 adding a physically interpretable representation of the precipitation-to-storage transformation  
 108 process. Based on this enhancement, they re-estimated that approximately 64% of terrestrial  
 109 precipitation contributed to TWS across 121 global river basins during the period from 2002 to 2021,  
 110 with substantial spatial variability linked to climatic and geographic conditions. The specific  
 111 formulations and reconstruction workflows of both statistical models are detailed in Sections 3.1  
 112 and 3.2 of this study.

113 Expanding upon the frameworks of Humphrey and Gudmundsson (2019) and Zhong et al.  
 114 (2025), this study further advances parameter design. Whereas previous models relied on two  
 115 empirical parameters in a residence-time formulation that is only indirectly linked to individual  
 116 water-balance components, we propose a linear, four-parameter coupled recursive model at the daily  
 117 time scale. This model explicitly incorporates temperature modulation of both the immediate  
 118 conversion efficiency of precipitation and the retention rate of water storage. The model includes  
 119 four parameters. The four model parameters are designed to be more closely tied to these  
 120 hydrological processes. Detailed interpretations of the parameters are provided in Section 3.2.1. In  
 121 addition, the approach has been optimized for computational efficiency and numerical stability.  
 122 Using this method, we reconstructed a global (excluding Antarctica) gridded TWSA dataset at 0.5°  
 123 spatial resolution and monthly time step for the period from 2002 to 2021 (the product name is  
 124 provided in Section 4.1). The reconstructed TWSA closely matches GRACE observations and  
 125 achieves accuracy levels comparable to or exceeding those of existing reconstruction products, as  
 126 demonstrated in comprehensive benchmark evaluations.

127

## 128 **2 Data**

129 **Table 1 List of the data sets**

Dataset	Time	Spatial resolution	Temporal resolution
GSFC RL06v2.0 mascon solution	2002— 2021	1°equal area (provided on 0.5°)	monthly

JPL CRI Filtered RL06.1v03 mascon solution	2002— 2023	3.0°equal area (provided on 0.5°)	monthly
CSR RL06v2.0 mascon solution	2002— 2021	1°equal area (provided on 0.25°)	monthly
Precipitation (ERA5-land)	2000— 2023	0.1°	hourly
Precipitation (MSWEP v2.8)	2000— 2021	0.1°	daily
Temperature (ERA5-land)	2000— 2023	0.1°	hourly
Temperature (GLDAS-2.2)	2003— 2021	0.25°	daily
Evapotranspiration (ERA5-land)	2000— 2020	0.1°	hourly
TWS (GLDAS CLSM)	2003— 2021	1°	monthly
TWSA (GLDAS-2.2)	2003— 2021	0.25°	daily
GRACE-REC (Humphrey’s reconstruction)	2002— 2019	0.5°	monthly

130

## 131 2.1 GRACE data

132 This study employs three GRACE/GRACE-FO mascon solutions (Table 1): the JPL RL06 v03  
133 mascon product (JPLM) provided by the Jet Propulsion Laboratory (Wiese et al., 2019), the CSR  
134 RL06 v2.0 mascon product (CSRML) provided by the Center for Space Research (Save et al., 2016),  
135 and the GSFC RL06v2.0 mascon product (GSFCM) provided by the NASA Goddard Space Flight  
136 Center (Loomis et al., 2019). All datasets cover the period from 2002 to 2021. The native resolution  
137 of JPLM is  $3^\circ \times 3^\circ$  equal-area caps, and the data are distributed on a  $0.5^\circ$  latitude and longitude  
138 grid. The GSFCM product has a native resolution of  $1^\circ \times 1^\circ$  equal-area caps and is also provided  
139 on a  $0.5^\circ$  grid. Both mascon products have been preprocessed with geocenter correction, degree-2  
140 replacement, and GIA correction, and thus require no further post-processing.

## 141 **2.2 Precipitation and Temperature Data**

142 This study employs two precipitation products and two temperature products (Table 1). For  
143 precipitation, the following datasets are used: (1) The ERA5-Land dataset, developed by the  
144 European Centre for Medium-Range Weather Forecasts (ECMWF), is the land component of the  
145 ERA5 reanalysis. It is driven by atmospheric forcing variables from ERA5 and provides hourly  
146 precipitation data at a spatial resolution of  $0.1^\circ \times 0.1^\circ$  ( $\sim 9$  km), covering the period from January  
147 1950 to present (Muñoz-Sabater et al., 2021; Muñoz-Sabater, 2019); (2) The Multi-Source  
148 Weighted-Ensemble Precipitation (MSWEP) version 2.8 dataset (Beck et al., 2017; Beck et al., 2019)  
149 spans from 1979 to present and offers precipitation estimates at  $0.1^\circ \times 0.1^\circ$  spatial resolution and  
150 multiple temporal resolutions (3-hourly, daily, and monthly). In this study, the daily product is used.  
151 MSWEP combines gauge observations, satellite retrievals, and reanalysis data using a weighted  
152 ensemble approach to provide high-accuracy global precipitation estimates (Beck et al., 2017).

153 For temperature, the following datasets are used: (1) The ERA5-Land reanalysis dataset  
154 produced by ECMWF (Muñoz-Sabater, 2019); (2) the GLDAS-2.2 daily CLSM product (Li et al.,  
155 2019). This model offers a physically based framework suitable for assimilating GRACE-derived  
156 TWS observations. The GLDAS-2.2 daily product includes both a main version and an early product.  
157 This study uses the main version, which spans from February 2003 to December 2021 (as of  
158 manuscript preparation), with a spatial resolution of  $0.25^\circ \times 0.25^\circ$  and a daily temporal resolution.  
159 It should be noted that, due to the data usage agreement with the ECMWF, which prohibits  
160 redistribution of products from the Integrated Forecasting System (IFS), the GLDAS-2.2 daily  
161 product does not include the meteorological forcing fields. Instead, GLDAS-2.2 provides land  
162 surface variables simulated by the Catchment land surface model, which is forced by meteorological  
163 analysis fields from the operational ECMWF IFS.

164 Although both ERA5-Land and GLDAS-2.2 are ultimately based on meteorological  
165 information produced within the ECMWF framework, their temperature datasets exhibit a  
166 discrepancy of approximately 1–2 °C (Fig. S1) due to different data assimilating methods and  
167 models.

168 ERA5-Land precipitation and temperature are used as the primary meteorological inputs in this  
169 study. The MSWEP precipitation and GLDAS-2.2 temperature datasets are additionally used to  
170 compare standard deviations with ERA5-Land, in order to assess the uncertainty of ERA5-Land  
171 forcing data and support the subsequent parameter uncertainty analysis.

## 172 **2.3 Ancillary Data**

173 This study employs basin boundaries from the publicly available dataset “Major River Basins  
174 of the World” provided by the Global Runoff Data Centre (GRDC). This dataset has been widely  
175 used in large-scale hydrological studies (e.g., Burek and Smilovic (2023); Zhong et al. (2025); Gao  
176 et al. (2026)). We followed the strategy of Zhong et al. (2025) and ranked all basins according to  
177 drainage area to ensure spatial representativeness and avoid subjective selection. The 116 largest  
178 basins with drainage areas exceeding  $10^5 \text{ km}^2$  were selected for analysis.

179 The classification of basin climate types was based on the mean annual aridity index, derived  
180 from the Global Aridity Index and Potential Evapotranspiration Climate Database (Zomer et al.,  
181 2022; Trabucco and Zomer, 2019), which is available in GeoTIFF format. The aridity index is  
182 defined as the ratio of annual precipitation to mean annual potential evapotranspiration, where  
183 higher values indicate more humid conditions. Climate types were categorized according to the  
184 aridity index as follows: arid (aridity index  $< 0.2$ ), semi-arid ( $0.2 \leq$  aridity index  $< 0.5$ ), semi-  
185 humid ( $0.5 \leq$  aridity index  $< 0.65$ ), and humid (aridity index  $\geq 0.65$ ).

186

## 187 **3 Methods**

### 188 **3.1 Previous Statistical Models for TWSA Reconstruction**

#### 189 **3.1.1 Humphrey’s Reconstruction Model**

190 Precipitation is the primary input to terrestrial water storage. Statistical models that reconstruct  
191 TWSA from daily precipitation and temperature can be used to quantify the influence of climatic  
192 variability on terrestrial water storage (Humphrey and Gudmundsson, 2019). For comparison  
193 purposes, we first revisit the statistical reconstruction model proposed by Humphrey and  
194 Gudmundsson (2019) :

$$195 \quad TWS(t) = TWS(t-1) \cdot e^{-\frac{1}{\tau(t)}} + P(t). \quad (1)$$

196 Here  $t$  is the daily time index,  $TWS(t)$  is the water storage on day  $t$ ,  $P(t)$  is the  
197 precipitation input on day  $t$ , and  $\tau(t)$  is the water residence time. A larger  $\tau$  implies slower losses  
198 through evapotranspiration or runoff, while a smaller  $\tau$  indicates faster depletion.

199  $\tau(t)$  is defined as a function of temperature to represent the seasonal variations in water  
200 residence time.

$$201 \quad \tau(t) = a + b \cdot T_z(t), \quad (2)$$

202 with  $a$  and  $b$  are model parameters to be calibrated.

203 To reduce the influence of extreme temperatures on model stability, the original temperature  
 204  $T(t)$  is first detrended and standardized, then transformed using a sigmoid function:

$$205 \quad T_z = 1 - \tanh\left(\frac{T_0 - \text{mean}(T_0)}{SD(T_0)}\right). \quad (3)$$

206 Since subzero temperatures have negligible effects on residence time (as temperature changes  
 207 exert little influence on storage losses such as evapotranspiration and runoff under freezing or  
 208 insufficient thawing conditions), the original temperature series  $T(t)$  was preprocessed as follows:

$$209 \quad T_0 = \begin{cases} 0, & T < 0 \\ T, & T \geq 0 \end{cases}. \quad (4)$$

210 The initial water storage  $TWS(0)$  is computed as the ratio of long-term mean precipitation to  
 211 mean water loss rate:

$$212 \quad TWS(0) = \frac{\text{mean}(P)}{1 - \text{mean}\left(e^{-\frac{1}{\tau(t)}}\right)}. \quad (5)$$

213 The daily TWS series is aggregated to monthly means  $TWS(t_m)$  and calibrated against  
 214 deseasonalized and detrended GRACE TWSA at the monthly scale. The model parameters  
 215  $a, b$  and  $\beta$  are calibrated using a Markov Chain Monte Carlo (MCMC) approach, with the objective  
 216 of minimizing the sum of squared residuals between the simulated and GRACE-derived TWSA:

$$217 \quad \text{anom}(GRACE(t_m)) = \beta \cdot \text{anom}(TWS(t_m)) + \epsilon. \quad (6)$$

218 The coefficient  $\beta$  serves as a scaling factor obtained from MCMC calibration,  $\epsilon$  denotes the  
 219 residual term, and  $\text{anom}(\cdot)$  represents the anomaly series after removal of the seasonal cycle and  
 220 linear trend.

### 221 **3.1.2 Zhong's Reconstruction Model**

222 The scaling factor  $\beta$  in Eq. (6) lacks clear physical meaning and serves merely as a calibration  
 223 parameter, resulting in a strong dependence of the reconstruction on GRACE observations. Building  
 224 on this framework, Zhong et al. (2025) proposed an improved formulation for reconstructing  
 225 terrestrial water storage :

$$226 \quad TWS(t) = TWS(t-1) \cdot e^{-\frac{1}{\tau(t)}} + \beta \cdot P(t). \quad (7)$$

227 In this formulation, the parameter  $\beta$  has a clear physical interpretation, representing the mean  
 228 fraction of daily precipitation that contributes to water storage. The residence time  $\tau(t)$  is defined  
 229 as in Eq. (2). It is important to note that this  $\beta$  is fundamentally different from the scale factor used  
 230 in Eq. (6).

231 The daily TWS series is aggregated to monthly means  $TWS(t_m)$  and calibrated against  
 232 deseasonalized and detrended GRACE TWSA at the monthly scale. The model parameters  
 233  $a, b$  and  $\beta$  are also calibrated using the MCMC approach:

$$234 \quad \text{anom}(GRACE(t_m)) = \text{anom}(TWS(t_m)) + \epsilon. \quad (8)$$

235 Owing to the high nonlinearity of the estimation model, parameter optimization is prone to  
 236 being trapped in local optima. To reduce the bias caused by incomplete chain convergence, the  
 237 MCMC calibration was repeated 50 times for each basin. All parameter sets falling within the  $\pm 1 \cdot$   
 238  $\sigma$  range were retained, and their averages were taken as the final estimates of parameters  $a, b$  and  $\beta$ .  
 239 The standard deviation of these filtered parameters was then calculated to represent the inter-basin  
 240 uncertainty of  $\beta$ .

## 241 **3.2 Proposed Reconstruction Method**

### 242 **3.2.1 Model formulation**

243 The two preceding approaches describe the residence time (Eq. 2) as a temperature-dependent  
 244 function. However, the parameters  $a$  and  $b$  in Eq. (1) and (7) are obtained by calibration, and they  
 245 are less directly linked to specific water-balance components. In this study, our model is derived  
 246 directly from the basin-scale water balance equation and provided a systematic clarification of both  
 247 the derivation and the physical interpretation of the parameters.

248 We start from the water balance equation:

$$249 \quad TWS(t) - TWS(t - 1) = P(t) - ET(t) - R(t) = P(t) - ETR(t), \quad (9)$$

250 where  $P(t)$ ,  $ET(t)$ , and  $R(t)$  denote precipitation, evapotranspiration, and runoff at time  
 251 step  $t$ , respectively, and  $ETR(t) = ET(t) + R(t)$  represents the total loss term. We assume that  
 252  $ETR$  consists of two components: (1) the contemporaneous precipitation loss  $ETR_1$ , i.e., the direct  
 253 loss occurring as precipitation is converted into water storage; and (2) the storage-release loss  $ETR_2$ ,  
 254 i.e., the loss arising from the depletion of antecedent water storage  $TWS(t - 1)$  in subsequent  
 255 periods, which can be expressed as

$$256 \quad ETR_1(t) = x \cdot P(t), \quad (10)$$

$$257 \quad ETR_2(t) = y \cdot TWS(t - 1), \quad (11)$$

$$258 \quad ETR(t) = ETR_1(t) + ETR_2(t) = x \cdot P(t) + y \cdot TWS(t - 1). \quad (12)$$

259 The assumption is validated in Fig. S2, where the  $ETR$  is plotted against precipitation and  
 260  $TWS$  respectively. Both plots show significant correlations, with  $ETR$  apparently exhibiting a  
 261 stronger correlation with the current precipitation than the antecedent  $TWS$ .

262 We then performed a multiple linear regression analysis to further test this assumption. In this

263 step, temperature modulation was temporarily ignored and the coefficients  $x$  and  $y$  were treated  
 264 as constants. The results show that all basins pass the F-test at the significance level of  $p < 0.01$ .  
 265 Moreover, 74 basins (approximately 64% of all basins) exhibit determination coefficients  $R^2 > 0.6$   
 266 (Fig. S3(a)), indicating that the total loss term  $ETR$  can be statistically explained by the linear  
 267 combination of precipitation  $P(t)$  and antecedent storage  $TWS(t - 1)$ .

268 We compared the full multiple regression model with a reduced model that includes only  
 269 precipitation  $P(t)$  as the predictor to further examine whether antecedent storage  $TWS(t - 1)$   
 270 provides additional explanatory power beyond precipitation alone. Figure S3(b) presents the spatial  
 271 distribution of the  $R^2$  obtained from the precipitation-only regression, while Fig. S3(c) shows the  
 272 spatial difference in  $R^2$  between the full model and the reduced model. Although precipitation alone  
 273 explains a substantial portion of the variability in basin-scale losses, the inclusion of antecedent  
 274 storage  $TWS(t - 1)$  improves model performance. Across all basins, the incorporation of  
 275  $TWS(t - 1)$  increases  $R^2$  by approximately 0.1. Therefore, although  $ETR$  is strongly correlated  
 276 with  $P(t)$ , the results demonstrate that accounting for  $TWS(t - 1)$  is both statistically meaningful  
 277 and physically necessary.

278 On this basis, we consider the modulation effect of temperature on the hydrological cycle by  
 279 assuming that the proportional coefficients  $x$  and  $y$  vary with temperature and apply a Taylor  
 280 expansion, which yields

$$281 \quad x = \epsilon \cdot f(T_z) = \epsilon \cdot (1 + \alpha T_z + o(T_z^2)). \quad (13)$$

282 Retaining only the first-order approximation gives

$$283 \quad x = a' \cdot T_z + b', \quad (14)$$

284 where  $T_z$  denotes standardized temperature.

285 Similarly, we obtain

$$286 \quad y = c' \cdot T_z + d'. \quad (15)$$

287 Substituting Eqs. (12), (14), and (15) into the water balance equation (Eq. 9) leads to

$$288 \quad TWS(t) = (1 - a' \cdot T_z - b') \cdot P(t) + (1 - c' \cdot T_z - d') \cdot TWS(t - 1). \quad (16)$$

289 After merging constant terms simplifying the signs, the recursive formulation adopted in this  
 290 study is obtained as

$$291 \quad TWS(t) = (a \cdot T_z + b) \cdot P(t) + (c \cdot T_z + d) \cdot TWS(t - 1), \quad (17)$$

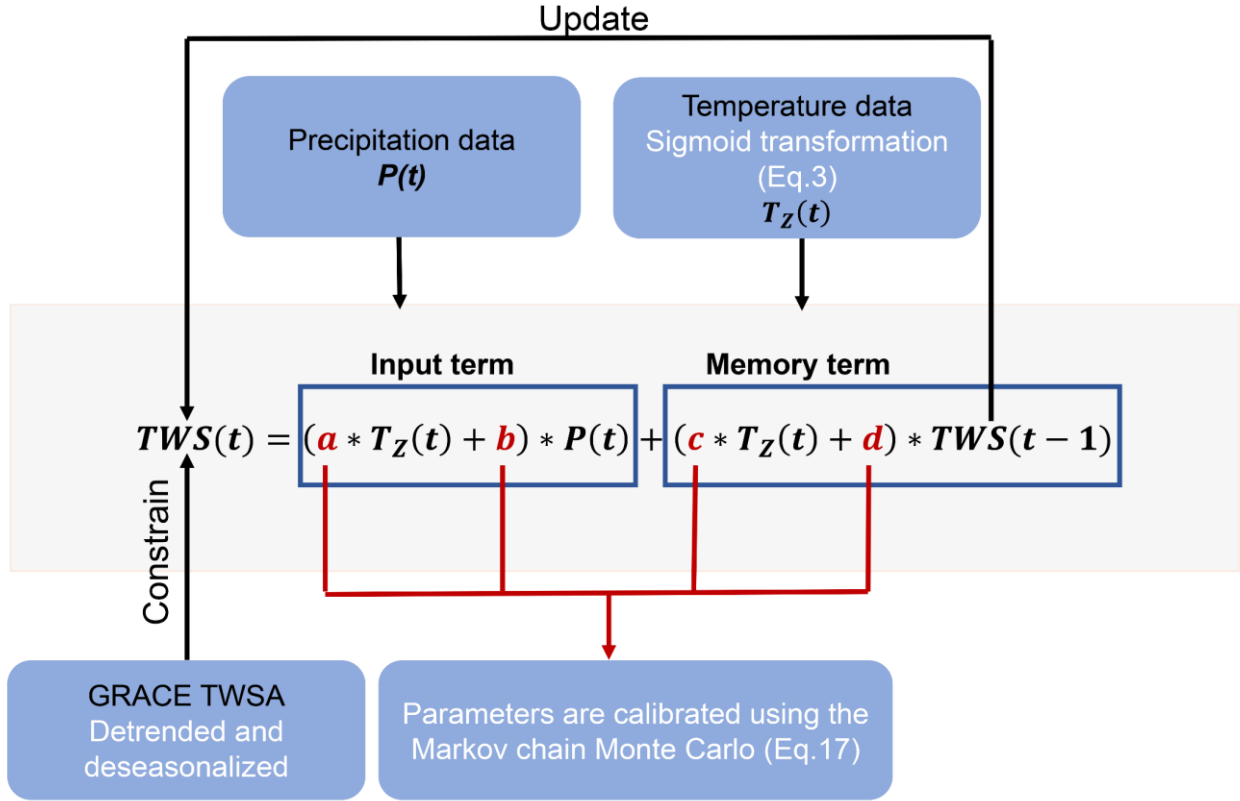
292 which is the reconstruction model adopted in this study. In this form, parameters  $a$  and  $c$  represent  
 293 temperature modulation effects. The term  $1 - b$  can be interpreted as the fraction of precipitation  
 294 that directly contributes to the loss component  $ETR_1(t)$ ; therefore, parameter  $b$  represents the  
 295 effective proportion of precipitation that contributes to  $TWS$  after basin losses. Similarly,  $1 - d$

296 represents the fraction of antecedent storage contributing to the loss term  $ETR_2(t)$ , and thus  
297 parameter  $d$  represents the fraction of previous storage retained in the current storage state. We  
298 will continue to discuss these parameters in subsequent sections.

299       When solving Eq. (1), (7), and (17), it is necessary to specify the initial water storage value,  
300  $TWS(0)$ . Humphrey and Gudmundsson (2019) and Zhong et al. (2025) assumed that the initial water  
301 storage corresponds to an equilibrium state, as defined in Eq. (5). Although using Eq. (5) provides  
302 a unique and reproducible starting condition, it inherently assumes that water storage is close to  
303 equilibrium at the beginning of the reconstruction period. However, if a basin experiences an  
304 extreme hydroclimatic condition at the initial time step, adopting Eq. (5) may introduce systematic  
305 biases and consequently affect the MCMC calibration of other parameters. In this study,  $TWS(0)$   
306 is also treated as a variable and is simultaneously estimated together with other model parameters  
307 through a joint least-squares and MCMC optimization. This approach eliminates the need to assume  
308 equilibrium conditions at the start of the reconstruction period and is therefore more suitable for  
309 basins under nonequilibrium states, such as those influenced by prolonged droughts or intensive  
310 groundwater extraction (see Section 5.1.2 for details).

311       The calibration procedure in this section follows exactly the same steps as described in Section  
312 3.1.2, with Eq. (8) retained throughout the process. The reconstruction logic of TWSA is illustrated  
313 in Fig. 1.

314



315

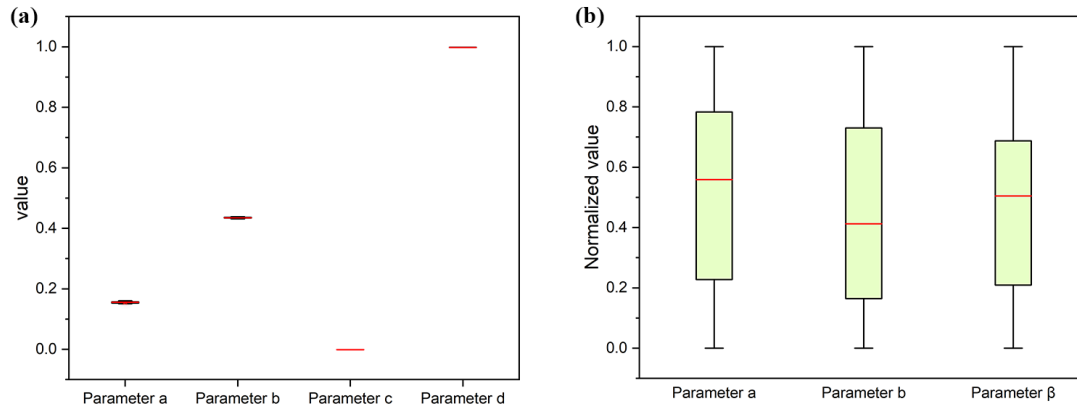
316 Figure 1: Illustration of the physical meaning of parameters.

317 Compared with the methods of Humphrey and Gudmundsson (2019) and Zhong et al. (2025),  
 318 this study linearizes the memory term  $TWS(t - 1) \cdot e^{-\frac{1}{\tau(t)}}$  as  $(c \cdot T_z + d) \cdot TWS(t - 1)$ . A key  
 319 advantage of this linearization is its computational stability and rapid convergence, which  
 320 substantially improves inversion efficiency. Under the same parameter calibration framework as  
 321 Zhong et al. (2025), both our model and the Zhong’s model were run with 50 independent MCMC  
 322 chains for the Yellow River to derive the posterior parameter distributions. The parameter dispersion  
 323 (indicated by the box height) in our model (Fig. 2a) is considerably narrower than that in Zhong’s  
 324 model (Fig. 2b), indicating that our model exhibits smaller parameter uncertainties and faster chain  
 325 convergence. Consequently, a single MCMC calibration is sufficient for each basin in our method,  
 326 greatly reducing computation time without a noticeable loss of fitting accuracy. Specifically, on our  
 327 computing platform (Intel 13th Gen Core i9-13900H processor, 14 cores, base frequency 2.60 GHz,  
 328 turbo frequency up to ~5.4 GHz, and 16 GB RAM), the Zhong’s model required 50 parameter  
 329 calibrations, taking a total of 353.8 s for a single basin, whereas our method required only one  
 330 calibration, completing the computation in 6.8 s.

331 In summary, this study improves upon previous models in three aspects:

332 (1) Linearization of the memory term. The linearization not only accelerates computational

- 333 performance but also enhances the physical interpretability of parameters.
- 334 (2) Incorporation of the coupling between temperature and precipitation. This mechanism
- 335 adjusts the fraction of precipitation converted into water storage, leading to a slight
- 336 improvement in NSE across most basins and substantial enhancement in several cases (see
- 337 Section 5.2.2).
- 338 (3) Consideration of the impact of different initial values on model fitting results (see Section
- 339 5.1.2).



340

341 Figure 2: Comparison of parameter stability in the Yellow River Basin between (a) our reconstruction model and

342 (b) the model of Zhong et al. (2025), based on 50 independent calibrations using the MCMC approach. Parameter

343 medians (red lines) are highlighted.

### 344 3.2.2 Further interpretation of Our Equations

345 To demonstrate the compatibility of the proposed model with existing approaches, we consider

346 two extreme cases:

347 (1) In the model of Zhong et al. (2025), the memory factor is expressed as  $e^{-\frac{1}{\tau(t)}}$ . When

348 applying a first-order Taylor expansion to the exponential term in Eq. (7), it can be approximated

349 as:

$$350 \quad TWS(t) = \beta \cdot P(t) + \left(1 - \frac{1}{\tau(t)}\right) \cdot TWS(t-1). \quad (18)$$

351 Our model (Eq. 17) becomes mathematically similar to Eq. (18) when the temperature and

352 precipitation modulation is neglected (i.e.,  $a = 0$ ). This indicates that our model degenerates to the

353 first-order linear form of the Zhong model when temperature effects are not considered.

354 (2) Assuming that the memory term is unaffected by temperature ( $c = 0$ ) and exhibits perfect

355 conversion efficiency ( $d = 1$ ), Eq. (17) simplifies to:

$$356 \quad TWSC(t) = TWS(t) - TWS(t-1) = (a \cdot T_z + b) \cdot P(t). \quad (19)$$

357 The water balance Eq. is given as:

$$358 \quad TWSC(t) = P(t) - ET(t) - R(t). \quad (20)$$

359 This implicitly assumes a proportional linear relationship between the loss terms,  $ET(t) +$   
360  $R(t)$ , and the precipitation input  $P(t)$ :  $R(t) + ET(t) = (1 - a \cdot T_z - b) \cdot P(t)$ . In other words,  
361 the total output is controlled by the input, so even without explicitly representing the loss terms, the  
362 Eq. is still able to account for their integrated contribution.

### 363 **3.2.3 Uncertainty Estimation**

364 To quantify the uncertainty associated with meteorological forcing data, this study estimates  
365 observational errors based on the differences between multiple data sources. Specifically, the  
366 standard deviation of daily precipitation differences between ERA5-Land and MSWEP is used as  
367 the uncertainty estimate for ERA5-Land precipitation, while the standard deviation of daily  
368 temperature differences between ERA5-Land and GLDAS-GLSM is used to estimate the  
369 uncertainty in ERA5-Land temperature.

370 Temperature uncertainty:

$$371 \quad \sigma_T = std(T_{ERA5-Land} - T_{GLDAS-GLSM}). \quad (21)$$

372 Precipitation uncertainty:

$$373 \quad \sigma_P = std(P_{ERA5-Land} - P_{MSWEP}). \quad (23)$$

374 Building on this, we adopt a Monte Carlo simulation framework in which Gaussian noise with  
375 zero mean and standard deviations of  $\sigma_T$  and  $\sigma_P$  is added to the temperature and precipitation  
376 inputs, respectively. For each perturbed dataset, the model is recalibrated to obtain an optimal set of  
377 parameters under the given realization. After completing 1,000 simulations, all sets of estimated  
378 parameters are collected to derive the uncertainty distributions of both the model parameters and the  
379 reconstructed TWS.

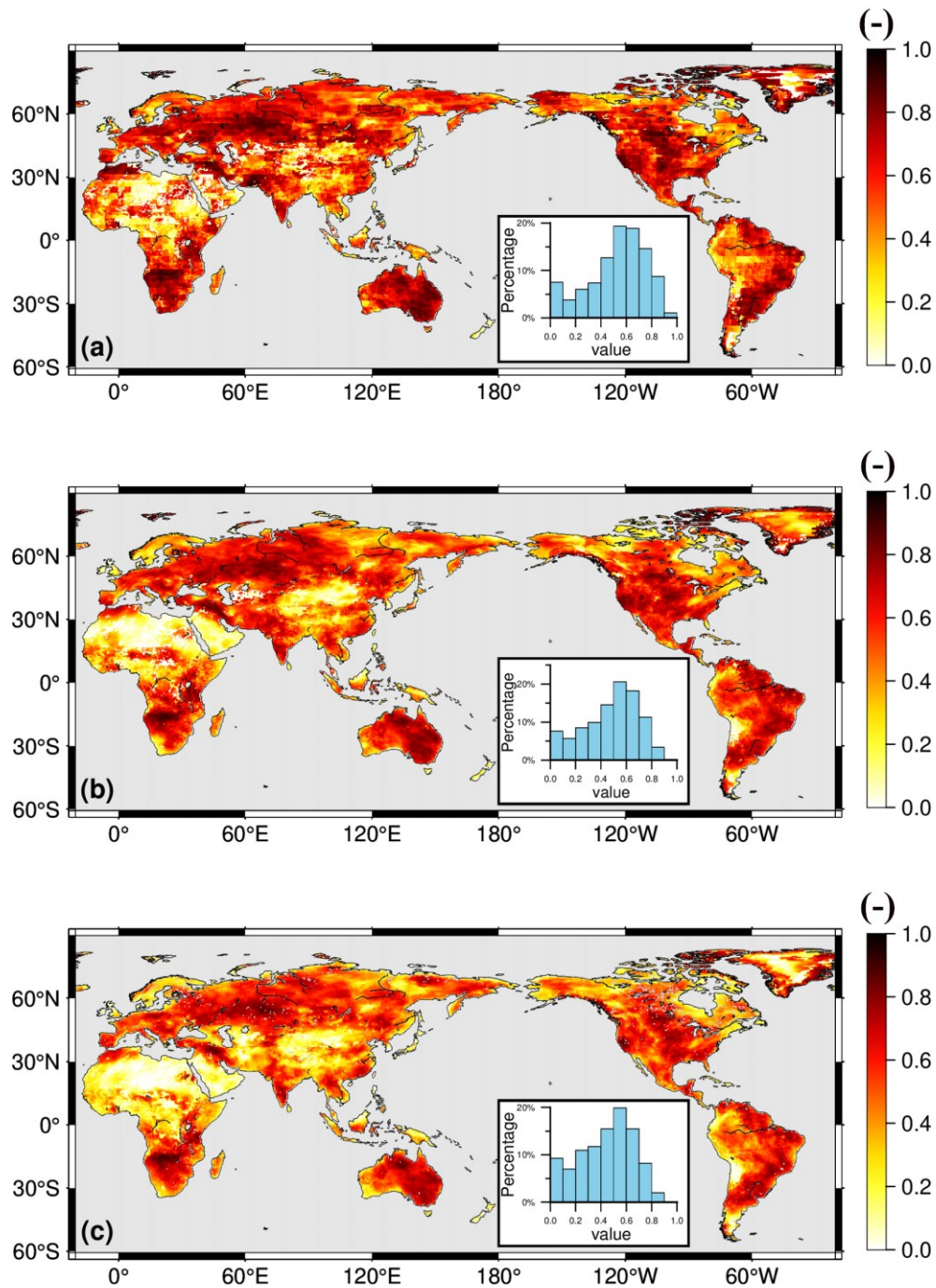
380

## 381 **4 Results**

### 382 **4.1 Performance summary**

383 The performance of the GRACE reconstruction model was evaluated using the *NSE*. Figure 3  
384 presents the spatial distribution of *NSE* values for 116 major river basins worldwide, and the  
385 corresponding numerical results are provided in Table S1. For the reconstruction calibrated against  
386 the JPLM product (hereafter referred to as JPL-REC), 53.3% of the  $0.5^\circ$  grid cells exhibited

387 *NSE* values in the range of 0.5 to 0.8, while 8.6% showed  $NSE > 0.8$ . High *NSE* values were  
388 primarily concentrated in mid- to high-latitude regions such as North America, Europe, and  
389 Australia, where in situ meteorological observations are relatively dense (Fig. 3a). In contrast,  
390 regions such as most of Africa and the arid plateau zones of Central Asia showed relatively low  
391 *NSE* values, likely due to the scarcity of precipitation and temperature observations (Xu et al., 2020;  
392 Chen et al., 2008). The reconstruction calibrated against the GSFC mascon product (hereafter  
393 GSFC-REC) displayed a spatial pattern broadly similar to that of JPL-REC, with 50.2% of grid cells  
394 falling within the range from 0.5 to 0.8 and 2.79% exceeding  $NSE > 0.8$ . We also compared the  
395 reconstruction calibrated against the CSR product (hereafter referred to as CSR-REC) with the  
396 CSR product to further evaluate robustness of the reconstruction method. The spatial distribution  
397 of *NSE* (Fig. 3c) is broadly consistent with those obtained for JPL-REC and GSFC-REC. The *NSE*  
398 distributions derived from the three mascon-based calibrations are largely consistent, indicating that  
399 the reconstruction approach exhibits limited sensitivity to the source of the mascon products, which  
400 is consistent with findings reported in previous studies (Humphrey and Gudmundsson, 2019).  
401



402

403 Figure 3: Spatial distribution of *NSE* (of de-seasonalized, de-trended anomalies) between our reconstruction and

404 (a) JPLM, (b) GSFCM, and (c) CSRM for the period from 2002 to 2021. Histograms show the distribution of

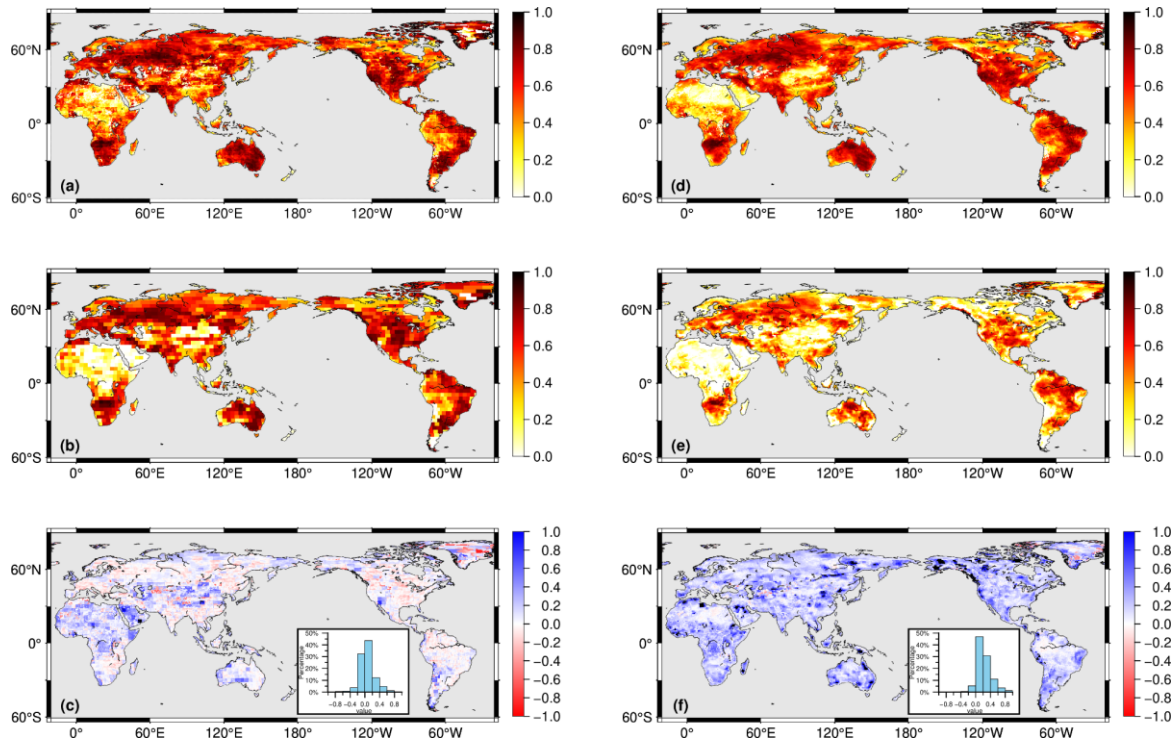
405 *NSE* values across all global grid points.

## 406 4.2 Comparisons

### 407 4.2.1 Comparison with the Humphrey's Model

408 At the grid scale, we first compared the reconstruction results from this study (JPL-REC and

409 GSFC-REC) with those of Humphrey and Gudmundsson (2019), which were trained using the  
410 JPLM and GSFCM products, respectively (hereafter referred to as Humphrey-JPL-REC and  
411 Humphrey-GSFC-REC). Figure 4 presents the spatial distribution of  $NSE$  values between the  
412 reconstructed and GRACE-derived TWSA after deseasonalization and detrending over the period  
413 from 2002 to 2019. Figure 4a and 4d show the  $NSE$  performance of JPL-REC and GSFC-REC  
414 relative to their corresponding mascon products (JPLM and GSFCM), respectively. Both panels  
415 indicate high model performance ( $NSE > 0.7$ ) across large parts of North America, Europe,  
416 Australia, and southern Africa, whereas low  $NSE$  values ( $NSE < 0.4$ ) are evident in parts of North  
417 Africa, Central Asia, and the Arabian Peninsula. Figure 4b and 4e show the  $NSE$  performance of  
418 Humphrey-JPL-REC and Humphrey-GSFC-REC relative to their corresponding mascon products  
419 (JPLM and GSFCM), respectively. While the overall spatial patterns are broadly similar to those in  
420 Fig. 4a and 4d, the extent of low-performance areas in Africa and Central Asia is noticeably larger  
421 in the Humphrey reconstructions. Figure 4c and 4f present the differences in  $NSE$  ( $\Delta NSE$ ),  
422 calculated as Fig. 4a minus Fig. 4b and Fig. 4d minus Fig. 4e, respectively. Blue shading ( $\Delta NSE >$   
423 0) indicates regions where the present method outperforms the approach by Humphrey and  
424 Gudmundsson (2019). Notably, in Fig. 4c, approximately 63% of the grid cells show  $\Delta NSE > 0$ ,  
425 with substantial improvements in regions such as the Arabian Peninsula and sub-Saharan Africa,  
426 along with scattered enhancements across other continents. Figure 4f demonstrates that under  
427 GSFCM constraints, the present method substantially outperforms that of Humphrey and  
428 Gudmundsson (2019). One potential explanation lies in the difference in training datasets: while  
429 (Humphrey and Gudmundsson, 2019) utilized the GSFC mascon v2.4 (an older version that is no  
430 longer publicly available), the present study employs the updated GSFC mascon RL06v2.0. This  
431 difference may account for the discrepancies between our results in Fig. 4e and those reported in the  
432 original publication by Humphrey and Gudmundsson (2019).



433

434 Figure 4: Spatial distribution of *NSE* (of de-seasonalized, de-trended anomalies) for the period from 2002 to  
 435 2019. (a) and (d): *NSE* of JPL-REC and GSFC-REC against their respective GRACE mascon solutions. (b) and  
 436 (e): *NSE* of Humphrey-JPL-REC and Humphrey-GSFC-REC against the corresponding GRACE mascon  
 437 solutions. (c) and (f): Spatial differences in *NSE* between our reconstruction and that of Humphrey and  
 438 Gudmundsson (2019), both relative to the same GRACE mascon solution. Values larger than 0 indicate better  
 439 performance of our method.

#### 440 4.2.2 Basin-scale comparison

441 We compare our reconstruction (JPL-REC) with two representative types of products to  
 442 systematically evaluate model performance at the basin scale: (1) a previous GRACE-based  
 443 reconstruction product (Zhong-REC from (Zhong et al., 2025)), and (2) physically based land  
 444 surface model simulations forced by meteorological data (GLDAS CLSM).

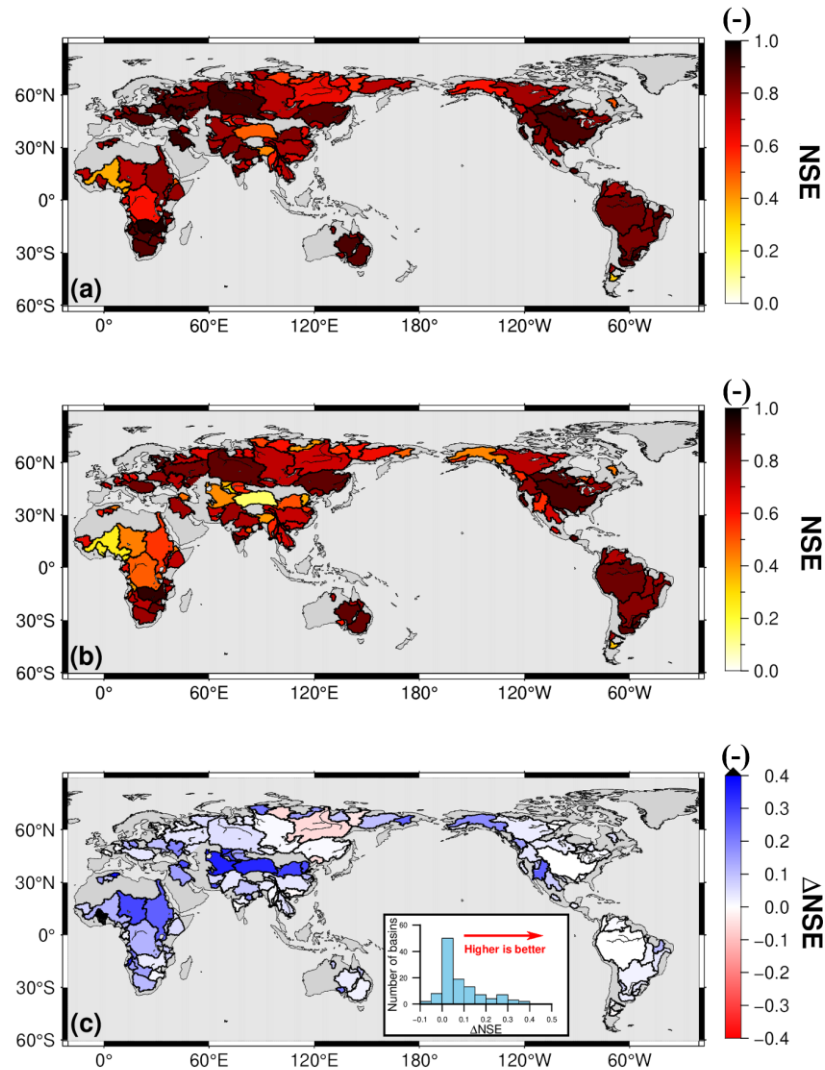
445 First, we adopted the JPLM as a common reference to evaluate the relative performance of  
 446 different statistical reconstruction methods. We compared basin-averaged monthly TWSA from our  
 447 JPL-REC and Zhong-REC against the JPLM and computed the *NSE* for each (Fig. 5). By comparing  
 448 the *NSE* values for the two approaches, we assessed their relative reconstruction accuracy. In the  
 449 comparison across 116 major global river basins, our proposed model achieved a median *NSE* of  
 450 0.76, outperforming Zhong-REC, which had a median *NSE* of 0.70. Among these basins, 84 (73%)  
 451 achieved  $NSE > 0.7$ , compared to only 59 (51%) for Zhong-REC. Model performance was closely  
 452 linked to the quality of the forcing data and the dominant hydrological processes (Yi et al., 2023).

453 Poor performance ( $NSE < 0.5$ ) was observed in basins such as the Congo, Niger, and Sanaga in  
454 Africa, and the Tarim River in Central Asia, where sparse ground-based meteorological observations  
455 result in large uncertainties in climate forcing data (Xu et al., 2020; Chen et al., 2008). Cold-region  
456 basins such as the Yukon, Brahmaputra, Lena, Indigirka, Kolyma, Olenek, Yana, and Khatanga are  
457 strongly influenced by glacial and permafrost-driven seasonal accumulation and ablation processes  
458 (Riegger and Tourian, 2014; Yi et al., 2023). The linear reservoir assumption fails to adequately  
459 capture the mass variations in these basins (Liu et al., 2022), resulting in  $NSE$  values ranging from  
460 0.45 to 0.70. In contrast, river basins where hydrological signals are dominant, such as the Amazon,  
461 Mississippi, Nelson, and Amur, showed good reconstruction performance with  $NSE$  values greater  
462 than 0.80. The histogram of  $\Delta NSE$  (this model minus Zhong-REC) shown in Fig. 5c reveals a  
463 positively skewed distribution, indicating performance improvements in the majority of basins.  
464 Specifically, 37 basins (approximately one-quarter of the sample) exhibited  $\Delta NSE > 0.1$ . The most  
465 pronounced improvements were found in arid and semi-arid regions (Fig. S4), including the Tarim  
466 River, Lake Chad, and Yellow River basins. The spatial distribution of regions with the most notable  
467 improvements in Fig. 5c closely aligns with the  $\Delta NSE$  pattern between the proposed model and  
468 Humphrey-JPL-REC (Fig. S5), further supporting the role of the coupling between temperature and  
469 precipitation and the joint inversion of initial conditions in enhancing reconstruction accuracy in  
470 areas with low  $NSE$ .

471 Second, we compared basin-averaged monthly TWSA from JPL-REC and GLDAS CLSM to  
472 highlight the improvement of our reconstruction method relative to the land surface model (Fig. S6).  
473 Across 116 major river basins worldwide,  $NSE$  values for both products are computed against the  
474 JPLM, and the spatial distribution of the resulting  $\Delta NSE$  is presented (Fig. S6(c)). The results show  
475 that  $NSE$  values from our reconstruction are higher in most basins, with  $\Delta NSE$  predominantly  
476 positive, indicating that the proposed method more effectively captures temporal TWSA variations  
477 at the basin scale.

478 It should be noted that TWSA estimates from GLDAS CLSM are derived from explicit  
479 simulations of individual storage components and are therefore influenced by the completeness and  
480 accuracy of the represented physical processes. In contrast, our approach statistically characterizes  
481 climate-driven storage responses under direct constraints from GRACE observations, implicitly  
482 integrating multiple storage processes. This leads to improved representation of total water storage  
483 variations and better agreement across many river basins.

484



485

486 Figure 5: Spatial distribution of  $NSE$  (de-seasonalized, de-trended anomalies) between JPL mascon and two  
 487 reconstruction models across 116 global river basins for the period 2002 to 2021. The  $NSE$  between JPLM and  
 488 JPL-REC (a); the  $NSE$  between JPLM and the reconstruction by Zhong et al. (2025) (b); and their difference  
 489 ( $\Delta NSE = JPL-REC - Zhong-REC$ ) (c). The inset histogram in (c) illustrates the distribution of  $\Delta NSE$  across all  
 490 basins, where positive values indicate better agreement with GRACE observations by our reconstruction model.

### 491 4.2.3 Comparison of Monthly Time Series

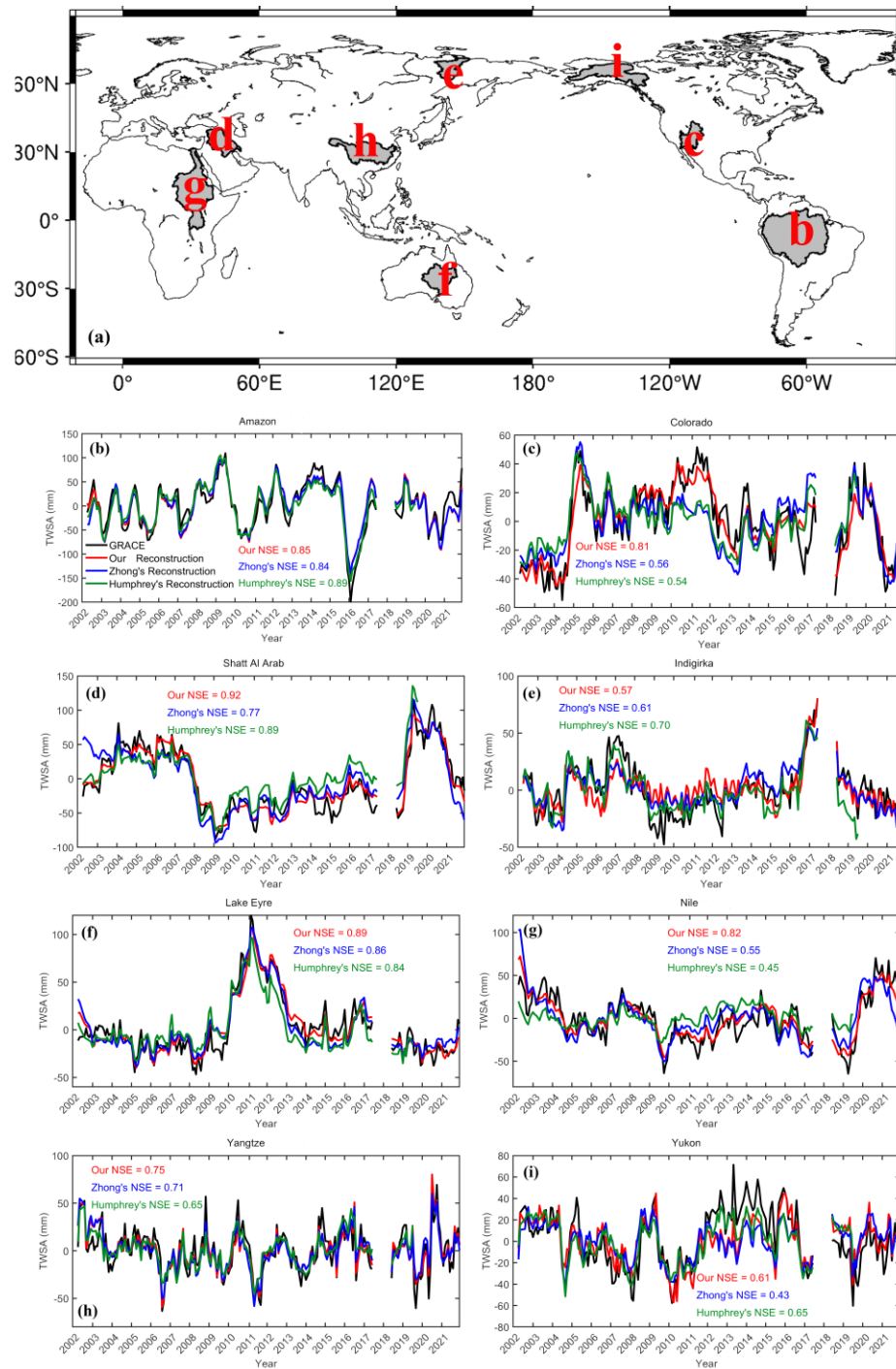
492 This section presents a comparison of deseasonalized and detrended monthly TWSA time  
 493 series across eight representative river basins. Figure 6 illustrates the differences among  
 494 GRACE/GRACE-FO observations (black line), the present reconstruction (JPL-REC, red line),  
 495 Zhong-REC (blue line), and Humphrey-JPL-REC (green line) in terms of basin-averaged TWSA  
 496 time series.

497 Interannual variations in terrestrial water storage are primarily driven by atmospheric  
 498 circulation anomalies induced by ENSO events, which often trigger hydrological extremes such as

499 floods and droughts, thereby significantly altering regional water balance (Chen et al., 2022). In the  
500 Amazon Basin, the La Niña event of 2009 markedly increased regional precipitation (Chen et al.,  
501 2010), while the strong El Niño of 2015–2016 led to severe drought conditions (Tian et al., 2021).  
502 The proposed reconstruction model successfully captured both the water storage peak in 2009 and  
503 the pronounced depletion in 2016. In the Yangtze River Basin, the early 2010 La Niña event was  
504 one of the primary drivers of drought, as it weakened the East Asian summer monsoon and caused  
505 an eastward retreat of the western Pacific subtropical high, thereby reducing moisture transport and  
506 triggering drought in the middle and lower reaches (Li et al., 2023; Zhang et al., 2015). The sharp  
507 decline in TWSA observed by GRACE during this period was accurately reproduced by the present  
508 model. In the Nile Basin, due to the slower response of TWS compared to precipitation, the  
509 reduction in net precipitation caused by the 2008 La Niña event did not immediately lead to a decline  
510 in TWSA. Instead, TWSA exhibited a gradual downward trend from 2008 to 2010 (Forootan et al.,  
511 2019), and this phase lag was also captured by the model. During the strong La Niña of the period  
512 from 2010 to 2011, northern and inland Australia experienced abnormally wet conditions. In the  
513 Lake Eyre Basin, intense rainfall and widespread flooding caused a significant rise in TWSA (Chen  
514 et al., 2022). The reconstructed TWSA time series in this basin is highly consistent with GRACE  
515 observations.

516 In contrast to the basins discussed above, interannual anomalies in river basins such as the  
517 Colorado and Shatt al-Arab are closely coupled with human water extraction. In the Colorado Basin,  
518 groundwater levels experienced a brief recovery from 2009 to 2010 due to increased precipitation.  
519 However, an anomalous drought event in 2012, combined with record-low snowfall in the Rocky  
520 Mountains, led to a rapid decline in water storage (Castle et al., 2014). Similarly, since the onset of  
521 drought in 2007, the Shatt al-Arab Basin has experienced a sharp decline in precipitation,  
522 compounded by a rapid drop in upstream reservoir levels along the Euphrates River. As surface  
523 water supplies became insufficient to meet demand, groundwater abstraction surged and became the  
524 primary source for agricultural and domestic use, leading to a rapid depletion of groundwater storage  
525 (Voss et al., 2013). In the high-latitude Indigirka Basin, winter precipitation is primarily stored in  
526 solid form, with snow water equivalent (SWE) gradually accumulating throughout the frozen season  
527 and then melting intensively in the following summer. GRACE effectively captured this “snow  
528 accumulation–melt” process, showing a spring peak of approximately 50 mm from 2006 to 2007.  
529 By contrast, the four-parameter recursive model developed in this study treats all precipitation  $P(t)$   
530 as an immediate input flux, which is attenuated through a memory term  $(c \cdot T_z + d) \cdot TWS(t - 1)$ .  
531 This assumption leads to premature dissipation of the snow accumulation signal during winter,

532 resulting in significant underestimation of the seasonal peak. Meanwhile, the relatively large  
 533 constant memory coefficient  $d$  caused part of the residual water to be released with a delay,  
 534 resulting in overall higher reconstructed TWSA compared to GRACE observations from 2008 to  
 535 2010. Among all basins examined, this basin exhibited the poorest performance of our method  
 536 relative to previous models.  
 537



538

539 Figure 6: Time series of GRACE/GRACE-FO TWSA and reconstructed TWSA, both de-seasonalized and de-  
540 trended, for the eight selected river basins (b-j) from 2002 to 2021. The global distribution of the eight selected  
541 river basins (a).

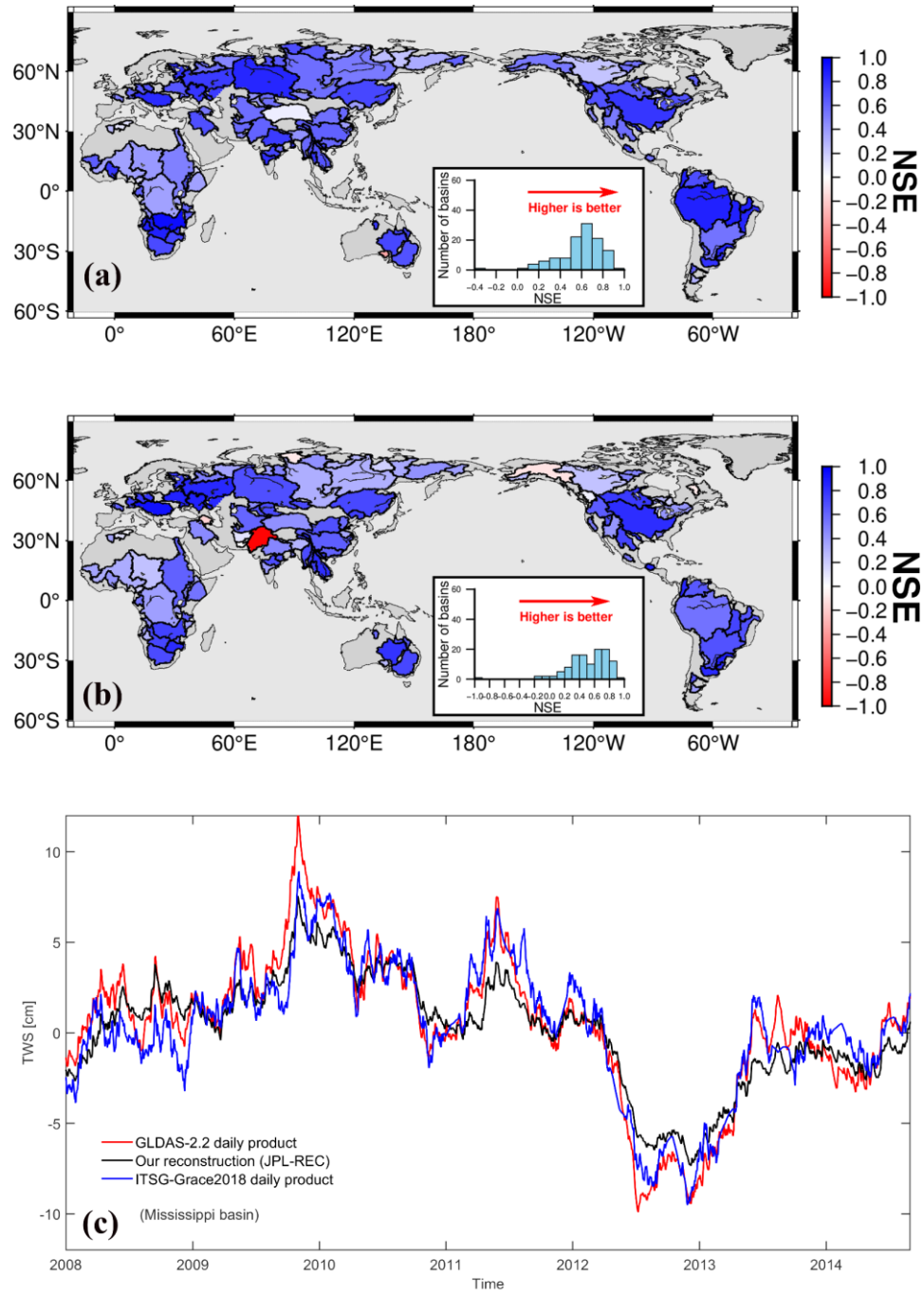
#### 542 **4.2.4 Comparison of Daily Time Series**

543 Although our model is primarily constrained by monthly GRACE observations, the method  
544 can also reconstruct daily characteristics in TWS changes based on the daily climate drivers. Here,  
545 we evaluated our reconstructed daily TWSA by comparing it with two different reference datasets,  
546 namely the daily ITSG-Grace2018 solution derived using a Kalman smoothing approach and the  
547 GLDAS-2.2 Daily product. The comparison with ITSG-Grace2018 evaluates whether the daily  
548 reconstruction is consistent with GRACE-based sub-monthly variability, whereas the comparison  
549 with GLDAS-2.2 assesses whether the daily reconstruction is consistent with physically based land-  
550 surface simulations.

551 We first assessed reconstruction performance at the basin scale by computing basin-averaged  
552 daily TWSA over major river basins worldwide. The difference in the time series was quantified  
553 using *NSE* (Fig. 7a&b). The *NSE* values are positive across most basins, with over half  
554 exceeding 0.5, indicating that our reconstruction results align well with the previous models.  
555 Although the reconstruction performs well overall relative to both reference datasets, its agreement  
556 with ITSG-Grace2018 is stronger in many basins. Specifically, 63 basins show *NSE* values greater  
557 than 0.6 in comparison with ITSG-Grace2018, whereas the count slightly reduces to 51 when  
558 compared with GLDAS-2.2. This is likely because our reconstruction is calibrated under constraints  
559 from GRACE/GRACE-FO TWSA, while ITSG-Grace2018 is also a daily product derived from  
560 GRACE observations.

561 To better illustrate sub-monthly variability, we showed daily TWSA time series in the  
562 Mississippi River basin in Fig. 7c. The results show similar sub-monthly variabilities among the  
563 reconstructed GRACE-REC, ITSG-Grace2018, and GLDAS-2.2 time series. One exception is found  
564 in the Indus basin between our GRACE-REC and GLDAS-2.2. The main reason is that GLDAS-2.2  
565 fails to capture a three-year drought spanning from 2016 and 2018 (see Fig. S7 for details). The  
566 reconstructed daily TWSA is consistent with the corresponding reduction in the 6-month moving  
567 average of precipitation. By contrast, the GLDAS-2.2 daily TWSA remains comparatively flat  
568 during this period, suggesting that it may underestimate the sensitivity of basin-scale storage to  
569 interannual hydroclimatic forcing in this basin. Figure S7(b) further shows that our monthly  
570 reconstruction remains close to the GRACE/GRACE-FO observations, particularly around the  
571 resumption of GRACE-FO observations in June 2018, whereas GLDAS-2.2 tends to overestimate

572 TWSA at that time. This discrepancy highlights the indispensable role of observation-based  
573 reconstruction approaches.



574

575 Figure 7: Spatial distribution of  $NSE$  (of de-seasonalized and de-trended TWSA) between our daily  
576 reconstruction and (a) ITSG-Grace2018 for the period from January 2003 to August 2016 and (b) the GLDAS-2.2  
577 daily product for the period from January 2003 to December 2021 across 116 major river basins worldwide. (c)  
578 Time-series comparison of the three daily datasets for the Mississippi River basin, focusing on the period 2008–  
579 2014 to improve the readability of high-frequency fluctuations.

#### 580 **4.2.5 Independent validation of reconstruction performance**

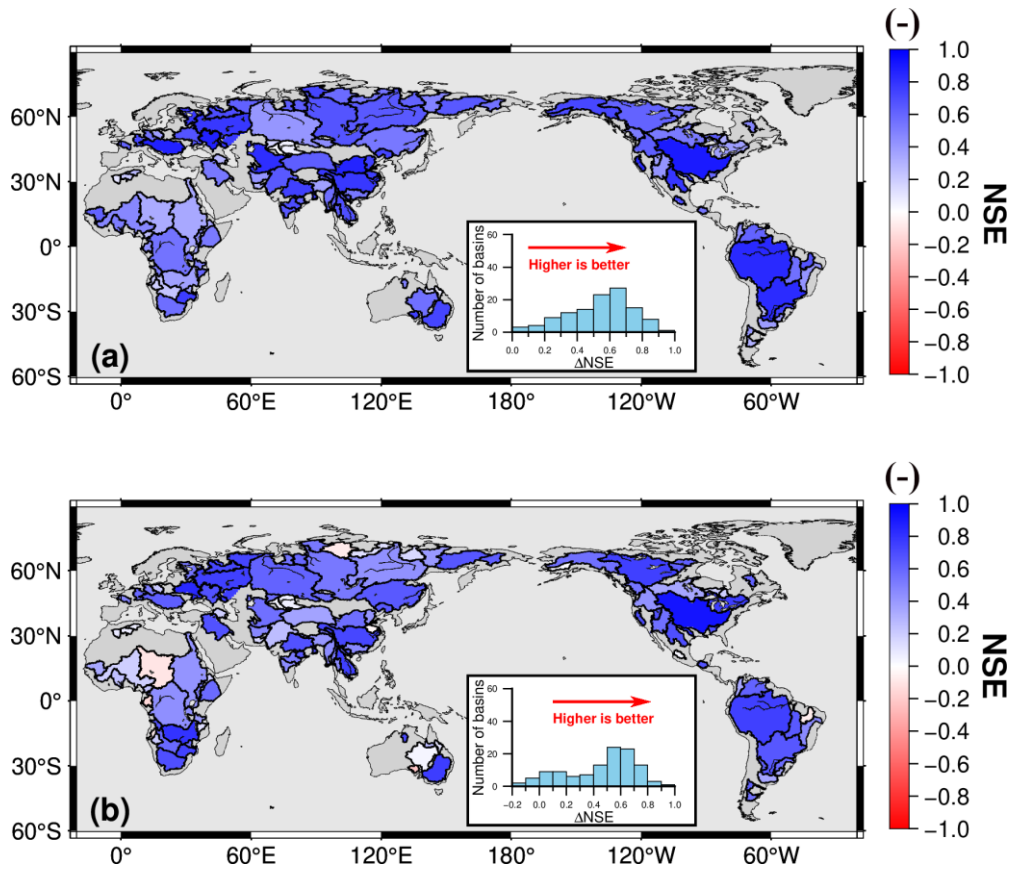
581 We conducted a self-validation experiment to further evaluate the robustness of the proposed  
582 method by dividing the JPLM product into separate training and validation periods. Specifically, the  
583 period from April 2002 to December 2012 was designated as the training period and the model  
584 parameters were estimated, while the period from January 2013 to December 2023 was treated as  
585 an independent validation period.

586 We evaluated the model performance using the  $NSE$  by comparing JPL-REC with the JPLM  
587 in both the training and validation periods (Fig. 8). During the training period (Fig. 8a), most basins  
588 exhibit high reconstruction skill, with  $NSE$  values exceeding 0.6 over large parts of North and  
589 South America as well as Eurasia. A very similar spatial pattern is obtained for the validation period  
590 (Fig. 8b), with only a modest reduction in overall skill. Approximately 60 % of the basins retain  
591  $NSE$  values above 0.5 in the validation period. This indicates that the model retains stable predictive  
592 capability beyond the calibration window and does not exhibit significant overfitting behavior.

593 The spatial pattern of the results during validation period reveals distinct regional  
594 characteristics. High  $NSE$  values are mainly found in large, humid basins with relatively dense  
595 observations, such as the Amazon, Mississippi, Yangtze River basins. In these regions, variability  
596 in TWS is strongly controlled by climate forcing. Consequently, our model can reliably reproduce  
597 the TWSA signals observed by GRACE/GRACE-FO.

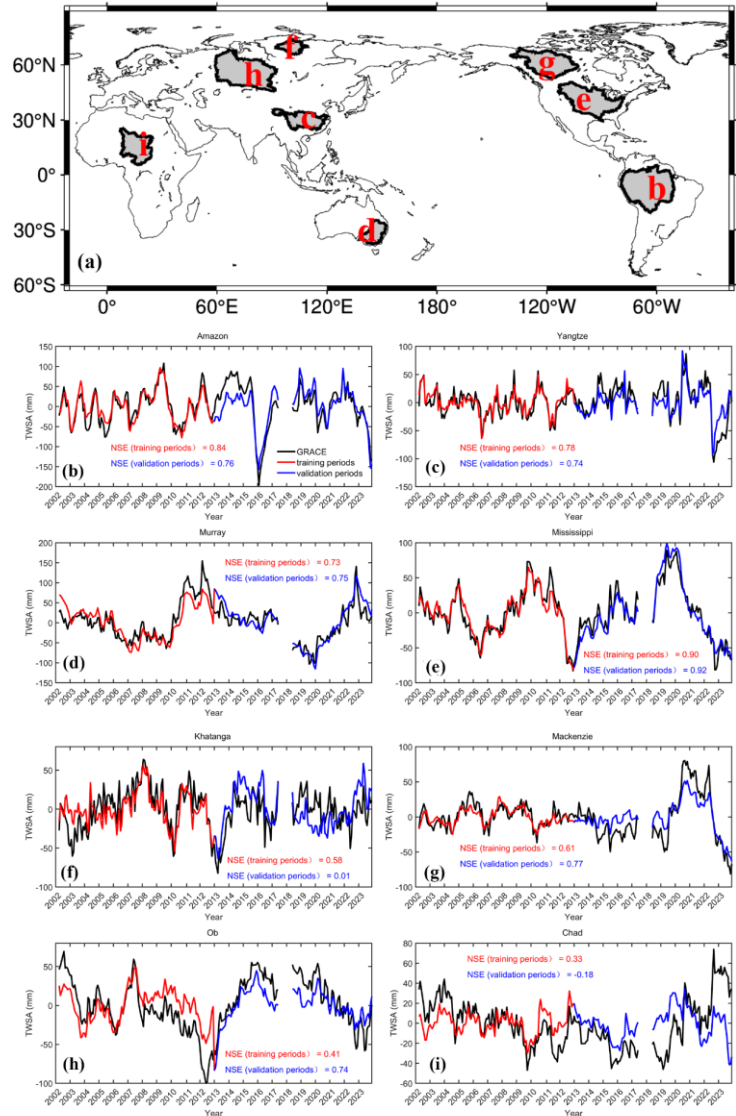
598 We further selected eight representative basins spanning high-latitude cold regions, tropical  
599 humid regions and typical monsoon-affected areas (Fig. 9). For each basin, the validation period  
600 TWSA (blue line) was compared with the reconstructed values during the training period (red line)  
601 and GRACE observations (black line). The results show that in certain basins, such as the Amazon,  
602 Yangtze, Mississippi and Murray, model performance during the extrapolation period remained  
603 comparable to or even exceeded that during the training period. This indicates that under these  
604 hydroclimatic conditions, the fixed-parameter model can reliably capture interannual TWSA  
605 variability driven by precipitation and temperature. In contrast, basins such as Chad and Khatanga  
606 show lower  $NSE$  during the validation period and larger discrepancies in certain years. This  
607 decline may be attributed to the limitations imposed by parameter nonstationarity and errors in the

608 forcing data. These results suggest that when only precipitation and temperature forecasts are  
609 available, the proposed recursive model can be directly applied for estimating TWS changes during  
610 missing periods or into the future. Nevertheless, its applicability varies with hydroclimatic  
611 conditions, and for regions requiring high prediction accuracy, periodic parameter updates or the  
612 inclusion of additional forcing information may still be necessary.



613

614 Figure 8: Comparison of reconstruction performance against JPLM data during the training and validation periods.  
 615 (a) Spatial distribution of the *NSE* between JPL-REC and JPLM data during the training period (April 2002–  
 616 December 2012); (b) Spatial distribution of *NSE* during the validation period (January 2013–December 2023)



617  
 618 Figure 9: Time series comparison of JPL-REC and JPLM, both de-seasonalized and de-trended, for the eight  
 619 selected river basins (b-i) during training and validation periods. Black lines represent GRACE TWSA, red lines  
 620 indicate model reconstructions during the training period, and blue lines during the validation period. The global  
 621 distribution of the eight selected river basins (a).

622

## 623 5 Discussions

### 624 5.1 Sensitivity Analysis of the Model

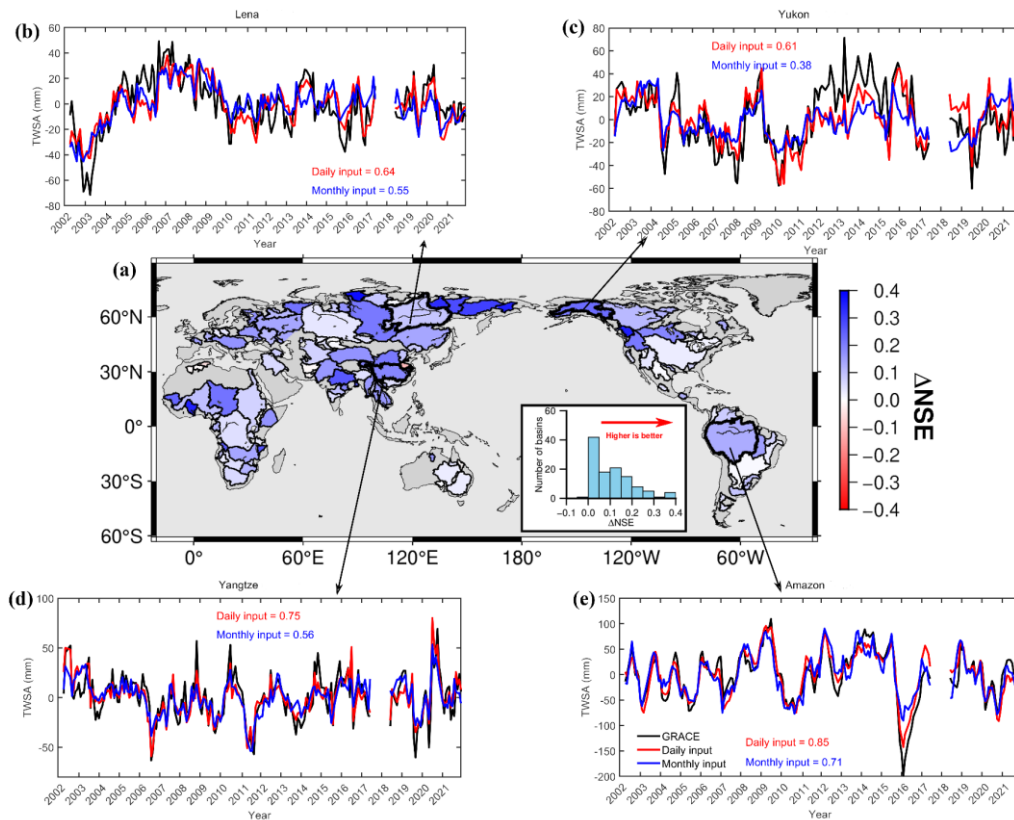
#### 625 5.1.1 Sensitivity to Temporal Resolution

626 We first evaluated the impact of the temporal resolution of meteorological forcing data on  
627 reconstruction performance, specifically whether daily forcing provides additional benefit compared  
628 to monthly forcing. To this end, all 116 basins were recalibrated using both daily and monthly ERA5-  
629 Land precipitation and temperature. The monthly forcing datasets were aggregated from daily values:  
630 precipitation was summed to monthly totals, and temperature was averaged over each month. The  
631 deseasonalized and detrended TWSA reconstructions were then compared on a monthly basis, and  
632 NSE was computed for each case. The difference between the two forcing schemes was quantified  
633 as  $\Delta NSE = NSE_{Daily} - NSE_{Monthly}$ , with its spatial distribution shown in Fig. 8a. Positive values  
634 indicate better performance under daily forcing.

635 The results show that 108 out of 116 basins (approximately 95%) exhibit positive  $\Delta NSE$   
636 values, suggesting that daily forcing generally outperforms monthly forcing. The most significant  
637 improvements were observed in high-latitude permafrost–glacier transition zones such as Siberia  
638 and Alaska. In contrast, tropical humid basins (e.g., the Paraná Basin) exhibited minimal differences  
639 ( $\Delta NSE \leq \pm 0.05$ ). Only seven basins (approximately 5%), such as the Huai River Basin, showed  
640 slightly negative differences ( $-0.1 < \Delta NSE < 0$ ), mostly located in regions with strong human  
641 regulation or sparse observational data.

642 From a time series perspective, the reconstructed TWSA driven by monthly climate forcing  
643 exhibits significantly lower amplitude compared to that driven by daily forcing. This is mainly due  
644 to two factors: daily forcing preserves extreme precipitation events and high-frequency temperature  
645 fluctuations on a day-to-day basis, allowing short-term water flux pulses to enter the recursive  
646 system as sharp peaks. In contrast, monthly forcing involves pre-integrating or averaging over 30-  
647 day precipitation and temperature records before inputting them into the model, which smooths out  
648 extreme events and leads to underestimated amplitude in the simulations (Humphrey et al., 2016).  
649 In addition, the memory coefficient calibrated at the daily time step typically ranges from  $d_{day} \approx$   
650 0.97 to 0.995, indicating very limited attenuation on a daily basis. When directly scaled to a  
651 monthly time step, this corresponds to  $d_{day}^{30} \approx 0.74$ ; alternatively, recalibration at the monthly scale  
652 yields an equivalent memory coefficient of approximately  $d_{day}^{30} \approx 0.7$  to 0.8. Both cases imply a  
653 stronger low-pass filtering effect, which further dampens peak values and elevates troughs.

654 Therefore, we argue that daily forcing is essential for improving model fidelity across nearly all  
 655 basins and should be considered necessary for accurate reconstruction.  
 656

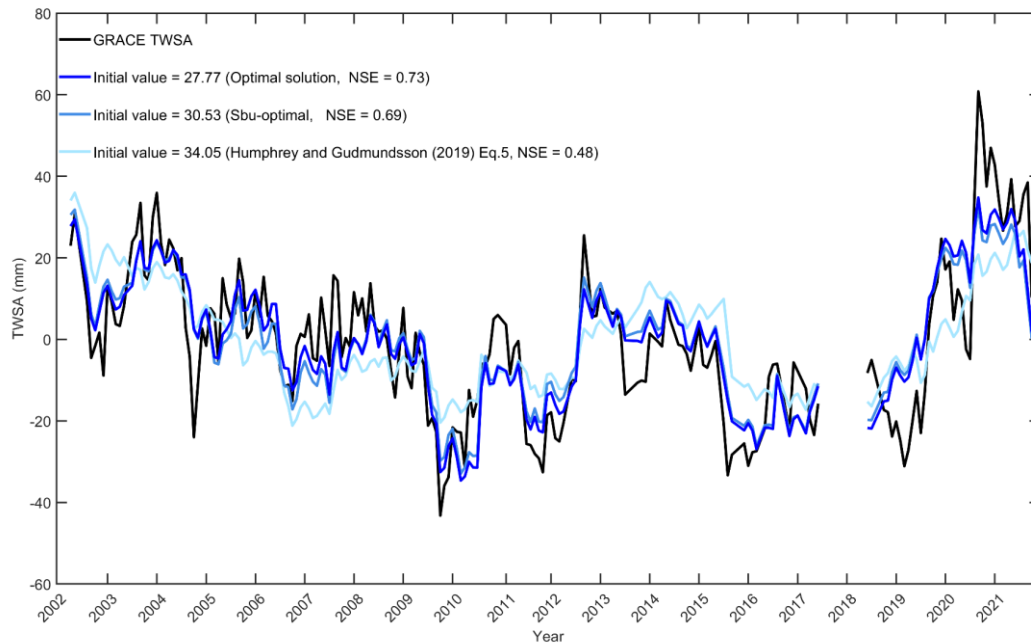


657  
 658 Figure 8: Spatial distribution of the difference in *NSE* between TWSA reconstructions forced by daily and  
 659 monthly meteorological data, with respect to GRACE JPLM from 2002 to 2021 (a). Time series of  
 660 GRACE/GRACE-FO TWSA and reconstructed TWSA (de-seasonalized and de-trended) using daily and monthly  
 661 meteorological data in the (b) Lena, (c) Yukon, (d) Yangtze, and (e) Amazon basins.

### 662 5.1.2 Sensitivity to Initial Conditions

663 This section investigates the influence of the initial terrestrial water storage,  $TWS(0)$ , on  
 664 reconstruction performance. Humphrey and Gudmundsson (2019) demonstrated through  
 665 experiments that if  $TWS(0)$  is simply set to zero, the model requires a spin-up period before  
 666 reaching equilibrium. To avoid discarding multiple years of data, they derived an analytical  
 667 equilibrium value to serve as a uniform initialization for  $TWS(0)$ . However, when a basin is not in  
 668 climatic equilibrium, this equilibrium-based initialization can introduce systematic bias. Since the  
 669 model's memory decay coefficient ( $c \cdot T_z + d$ ) is approximately equal to 1, such bias can persist  
 670 for several years in basins with long water residence times, making the initial condition a critical  
 671 factor influencing reconstruction accuracy in long-memory basins. To mitigate this bias,  $TWS(0)$

672 in this study is jointly estimated along with the model parameters ( $a, b, c, d$ ). Three initialization  
 673 strategies were compared (Fig. 9): (1) the optimal solution where  $TWS(0)$  is treated as an unknown  
 674 and solved through least squares estimation; (2) a suboptimal solution, obtained by increasing the  
 675 optimal value by 10%; and (3) the equilibrium solution, derived from Eq. (5). Using the Chad Basin  
 676 as an example, the NSE of the equilibrium initialization was 0.48, which increased to 0.57 under the  
 677 suboptimal case and further improved to 0.73 under the globally optimal initialization. Therefore,  
 678 in this study,  $TWS(0)$  is retrieved through parameter inversion for all basins, substantially  
 679 reducing systematic bias and enhancing reconstruction accuracy.



680  
 681 Figure 9: Time series of TWSA in the Chad Basin from GRACE observations and model reconstructions under  
 682 three different initial condition strategies. GRACE data (black) are based on the JPLM; model reconstructions are  
 683 obtained using (1) an optimized initialization (dark blue), (2) a sub-optimal initialization (medium blue), and (3)  
 684 an equilibrium initialization following Eq. (5) of Humphrey and Gudmundsson (2019) (light blue).

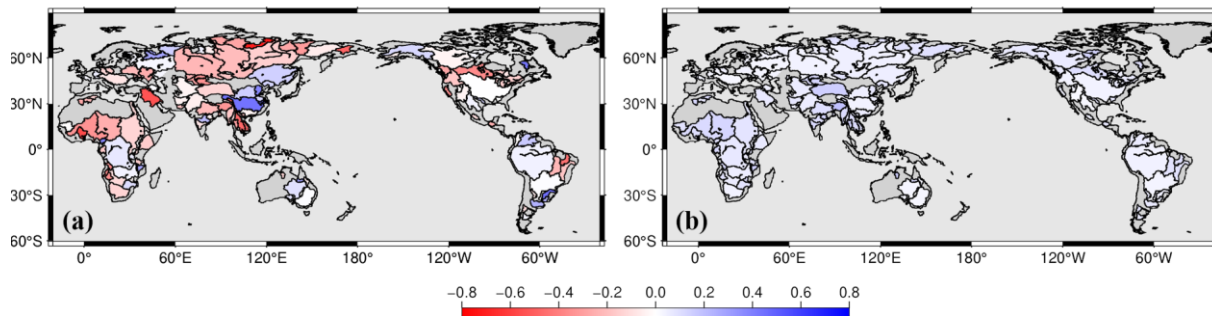
## 685 5.2 Discussion of Parameter Sensitivity

### 686 5.2.1 Parameter $a$

687 Figure 10 illustrates the global distribution of parameter  $a$  and its associated uncertainty. That  
 688 parameter  $a$  arises from the first-order Taylor expansion of the temperature modulation term in the  
 689 model formulation (Eq. 13). Therefore, it represents a mathematical sensitivity coefficient  
 690 describing how temperature perturbs the relationship between precipitation input and TWS response.  
 691 As shown in Fig. 10a, the sign of parameter  $a$  varies among basins, with positive values in some  
 692 basins and negative values in others. In many high-latitude regions, the calibrated values tend to be

693 negative, whereas some humid basins exhibit positive values. These spatial differences likely reflect  
694 variations in regional hydroclimatic conditions and the ways in which temperature influences  
695 hydrological processes such as snowmelt, evapotranspiration, and soil moisture dynamics.

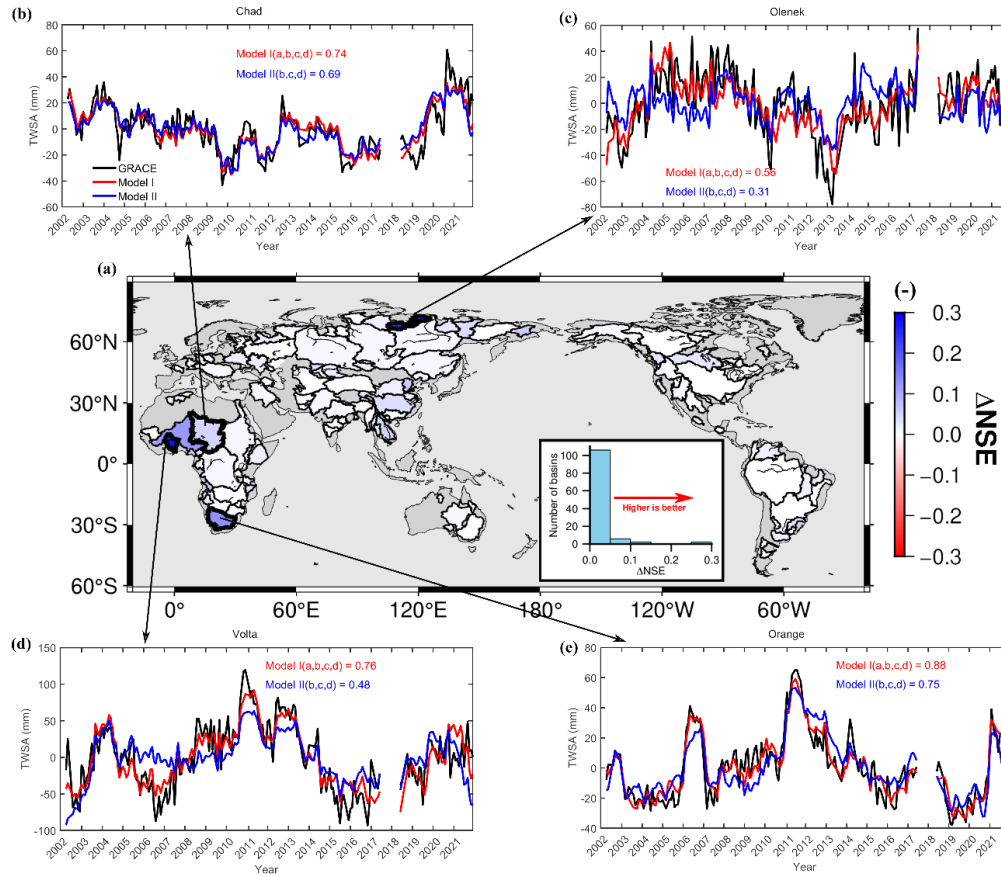
696 Figure 11a compares the full four-parameter model with a reduced three-parameter version in  
697 which the temperature and precipitation coupling term (i.e., parameter  $a$ ) is excluded. The global  
698 distribution of  $\Delta NSE$  indicates that, on average, the contribution of parameter  $a$  is limited.  
699 However, under specific hydroclimatic conditions, incorporating parameter  $a$  can lead to  
700 improvements in model performance with  $\Delta NSE > 0.1$ . Notable improvements are concentrated in  
701 high-latitude snow-dominated basins, such as the Olenek River, and in certain arid regions with high  
702 evapotranspiration, such as the Volta Basin. For the remaining 112 basins,  $\Delta NSE$  values are below  
703 0.1, suggesting no statistically significant difference. From a time series perspective, excluding  
704 parameter  $a$  in the Volta Basin leads to an overestimation of summer peaks during 2005–2006 and  
705 an underestimation during 2009–2012. In regions with high temperatures, removing parameter  $a$   
706 impairs the model’s ability to capture interannual peak and trough variability.



707

708 Figure 10: Spatial distribution of parameter  $a$  (a), and parameter  $a$  errors (b) across global 116 river basins.  
709 Parameters calibrated by monthly TWSA from JPLM based on ERA5-Land precipitation and temperature from  
710 2002 to 2021. Uncertainty estimation of parameters based on Monte Carlo simulations using precipitation and  
711 temperature errors across each basin. Precipitation errors were derived from the standard deviation between ERA5-  
712 Land and MSWEP datasets. Temperature errors were derived from the standard deviation between ERA5-Land  
713 and GLDAS-2.2 datasets.

714



715

716 Figure 11: Spatial distribution of the *NSE* difference between TWSA reconstructions from a four-parameter and  
 717 a three-parameter (excluding parameter *a*) daily recursive model with respect to JPLM across 116 global river  
 718 basins during 2002–2021 (a). Time series of GRACE/GRACE-FO TWSA and reconstructed TWSA (de-  
 719 seasonalized and de-trended) in the (b) Chad, (c) Olenek, (d) Volta, and (e) Orange basins.

## 720 5.2.2 Parameter *b*

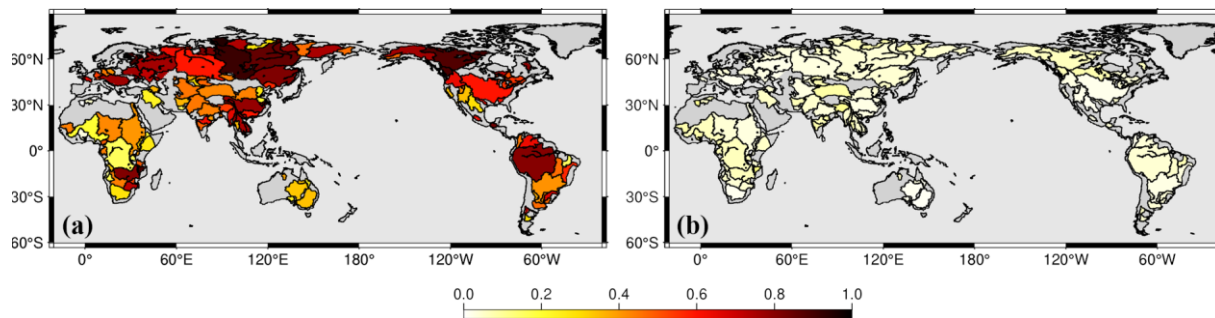
721 From Eq. (17) it follows that parameter *b* represents the fraction of precipitation that is  
 722 ultimately converted into TWS: a larger *b* value indicates that a higher fraction of precipitation  
 723 contributes to storage, whereas a smaller *b* suggests that a greater proportion of precipitation is  
 724 dissipated through evapotranspiration or rapid runoff before becoming part of the basin water  
 725 storage. In this sense, parameter *b* characterizes whether precipitation tends to be directly lost from  
 726 the system or effectively retained as storage.

727 Figure 12 shows the global distribution of parameter *b* and its associated uncertainty. Overall,  
 728 parameter *b* exhibits spatial heterogeneity across different regions. As shown in a Fig. 12,  
 729 relatively low parameter *b* values are found in the Congo, Niger, and Orange River basins in Africa;  
 730 the Lake Eyre and Murray basins in Australia; and the Rio Grande and Colorado basins in North  
 731 America. These regions are characterized by relatively high long-term mean ET/P (Fig. S8); a large

732 fraction of precipitation is exhausted by evaporation and runoff before contributing to *TWS*, which  
 733 yields a low conversion efficiency. Consequently, the efficiency of precipitation-to-storage  
 734 conversion is low, resulting in smaller calibrated *b* values.

735 In contrast, higher *b* values are observed in high-latitude basins such as the Lena and Yenisey.  
 736 In these regions, winter precipitation is predominantly stored as snow or ice, delaying immediate  
 737 losses to the atmosphere or river discharge. During the snowmelt season, the accumulated solid  
 738 water is released and contributes to *TWS*. This storage–delay–release mechanism reduces the  
 739 fraction of direct precipitation loss and enhances the effective conversion of precipitation into *TWS*,  
 740 thereby leading to higher *b* values.

741 To further substantiate this physical interpretation, we examined the statistical relationship  
 742 between parameter *b* and the multi-year mean loss ratio  $(ET + R)/P$  at the basin scale (Fig. S9).  
 743 The results reveal a negative correlation between parameter *b* and  $(ET + R)/P$ . This statistical  
 744 relationship provides independent observational support for interpreting parameter *b*. Furthermore,  
 745 the global spatial pattern of *b* is consistent with the findings of (Zhong et al., 2025).



746  
 747 Figure 12: Spatial distribution of parameter *b* (a), and parameter *b* errors (b) across global 116 river basins.  
 748 Parameters calibrated by monthly TWSA from JPLM based on ERA5-Land precipitation and temperature from  
 749 2002 to 2021. Uncertainty estimation of parameters based on Monte Carlo simulations using precipitation and  
 750 temperature errors across each basin. Precipitation errors were derived from the standard deviation between ERA5-  
 751 Land and MSWEP datasets. Temperature errors were derived from the standard deviation between ERA5-Land  
 752 and GLDAS-2.2 datasets.

### 753 5.2.3 Parameter *c*

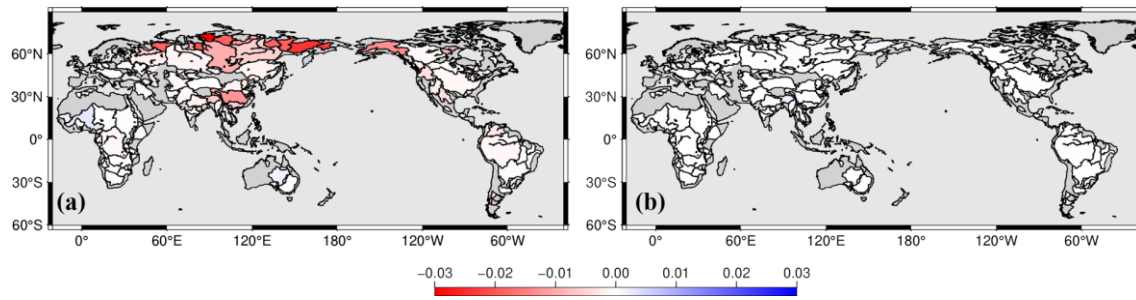
754 Figure 13 shows the global distribution of parameter *c* and its associated uncertainty. The  
 755 parameter *c* arises from the first-order Taylor expansion of the temperature modulation term in the  
 756 model formulation (Eq. 15). Therefore, it represents a mathematical sensitivity coefficient  
 757 describing how temperature perturbs the storage memory term. The parameter *c* exhibit clear  
 758 spatial variability. In many high-latitude basins, the magnitude of *c* is relatively larger (mostly  
 759 negative), whereas in most mid- and low-latitude basins the values are close to zero. This pattern

760 indicates that the influence of temperature on the model's water storage memory term is more  
761 pronounced in colder regions, while in many other basins the recursive water storage component is  
762 largely temperature-independent within our reconstruction model. The spatial distribution of  
763 parameter uncertainty (Fig. 13b) further shows that estimation errors are relatively small in most  
764 regions, suggesting that the adopted calibration approach is stable and robust across the majority of  
765 basins.

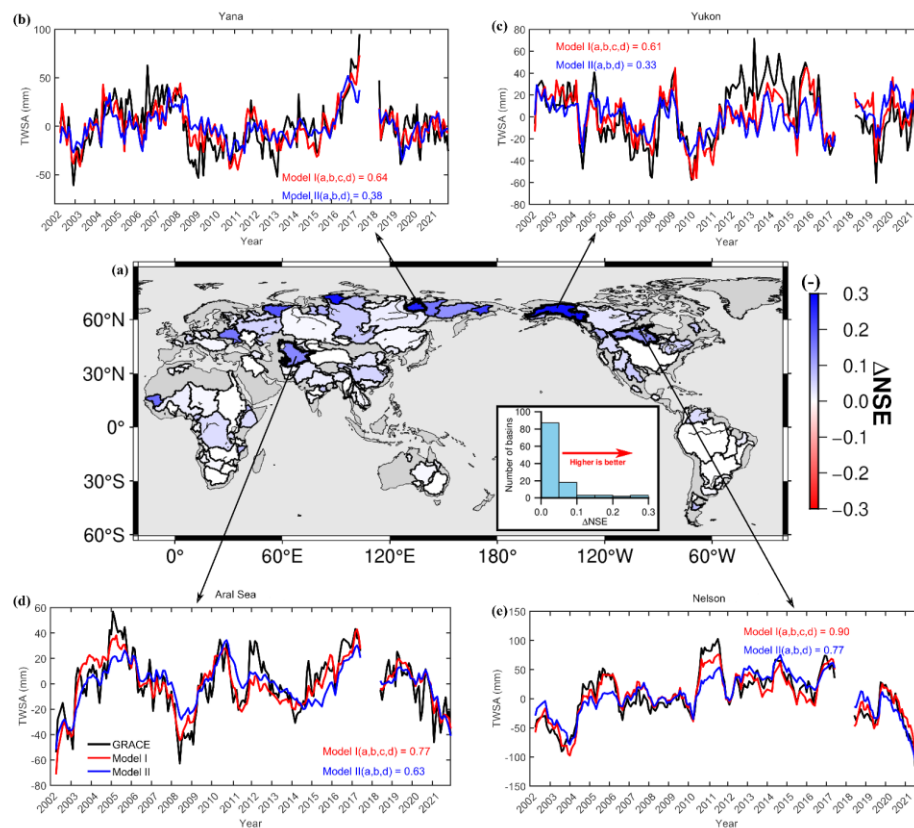
766 Figure 14 compares the full four-parameter model with a reduced three-parameter version  
767 excluding parameter  $c$ . Among the 116 basins analyzed, most exhibit limited sensitivity to  
768 parameter  $c$ ; however, incorporating  $c$  still yields measurable performance improvements in some  
769 cases, with 11 basins showing  $\Delta NSE > 0.1$ . Notable improvements are concentrated in high-  
770 latitude cold regions around the Arctic (e.g., Yana, Yukon, and Lena), while limited improvements  
771 are found in parts of North America (e.g., Columbia, Colorado, and Rio Grande) and in certain  
772 Chinese basins (e.g., Yangtze and Yellow River).

773 From a time series perspective, in the Yukon Basin, the average temperatures during the  
774 snowmelt seasons of the period from 2006 to 2007 and in 2009 were markedly below normal (Fig.  
775 S10). Under these conditions, the decay factor  $(c \cdot T_z + d)$  remained close to the baseline value  $d$ ,  
776 resulting in a slower dissipation rate. Meanwhile, SWE during these years was also below the multi-  
777 year average, limiting the contribution from snowmelt. GRACE observations captured a slight  
778 increase in TWS during spring, followed by a continuous decline. The four-parameter model  
779 successfully reproduced this slow depletion process (Fig. 14c). In contrast, the three-parameter  
780 model, due to the absence of a temperature modulation term, maintains a higher retention rate under  
781 the same precipitation input and consequently produces an overall overestimation in the  
782 reconstructed TWS curve. Conversely, 2019 was an exceptionally warm year, with precipitation  
783 anomalies remaining positive from 2017 to 2021. Rising temperatures substantially reduce  $(c \cdot T_z +$   
784  $d)$ , thereby accelerating the depletion of antecedent water storage. The abundant snowpack was  
785 rapidly flushed during the melt season, resulting in a deeper TWS trough recorded by GRACE  
786 compared to that of 2006 to 2007. The four-parameter model captured this deep trough due to the  
787 dynamic weakening effect of the parameter  $c$ , whereas the three-parameter model produced a  
788 shallower minimum. A similar pattern is observed in the Yana Basin: exceptionally high SWE in  
789 2007 led to elevated TWS levels, while reduced snowmelt and elevated temperatures in the  
790 following year (Fig. S11) triggered rapid snowmelt and a sharp decline in TWS during 2008. The  
791 four-parameter model exhibits better agreement with GRACE observations, whereas the three-  
792 parameter model shows excessive smoothing (Fig. 14b). These comparisons highlight the critical

793 role of the temperature modulation term  $c$  in accurately capturing TWS extremes during warm years  
 794 or periods of rapid snowmelt.



795  
 796 Figure 13: Spatial distribution of parameter  $c$  (a), and parameter  $c$  uncertainties (b) across global 116 river  
 797 basins.



799  
 800 Figure 14: Spatial distribution of the  $NSE$  difference between TWSA reconstructions from a four-parameter and  
 801 a three-parameter (excluding parameter  $c$ ) daily recursive model with respect to JPLM across 116 global river  
 802 basins during 2002-2021 (a). Time series of GRACE/GRACE-FO TWSA and reconstructed TWSA (de-  
 803 seasonalized and de-trended) in the (b) Yana, (c) Yukon, (d) Aral Sea, and (e) Nelson basins.

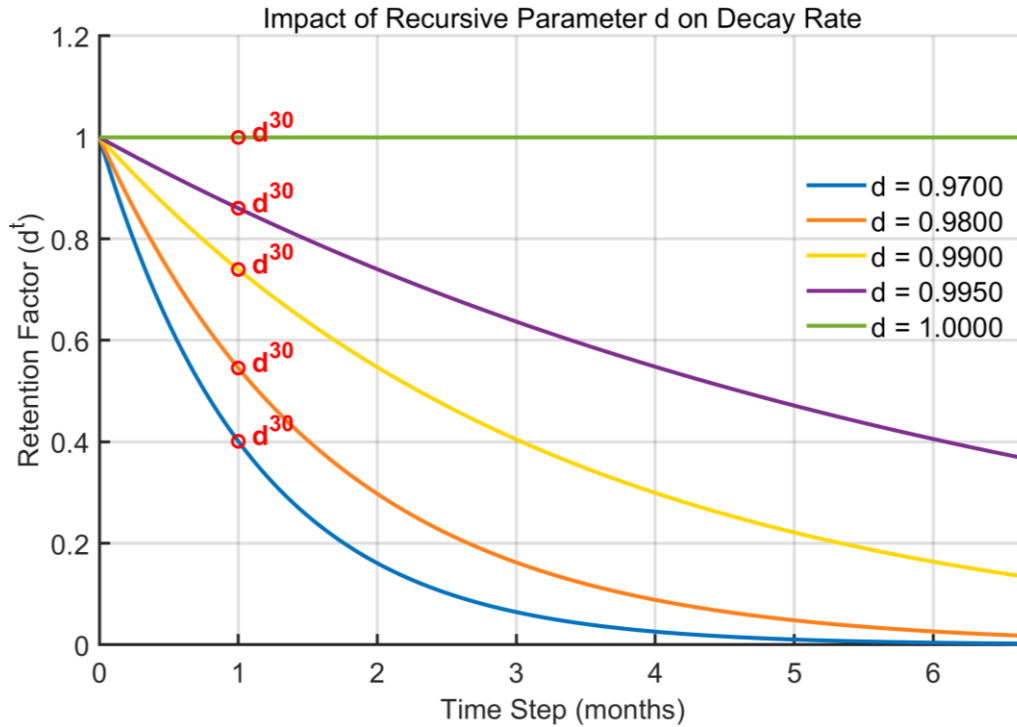
#### 804 5.2.4 Parameter $d$

805 In the linear storage model, when no additional precipitation input is introduced, water

806 depletion follows an exponential decay process. As shown in Fig. 15, when parameter  $d = 0.97$ , and  
807 in the absence of both additional precipitation input and temperature effects, only 40% of the stored  
808 water remains after one month. In contrast, when  $d = 0.99$ , more than 70% of the stored water can  
809 still be retained after one month under the same condition. This indicates that the interannual  
810 variability of TWS is highly sensitive to the value of parameter  $d$ .

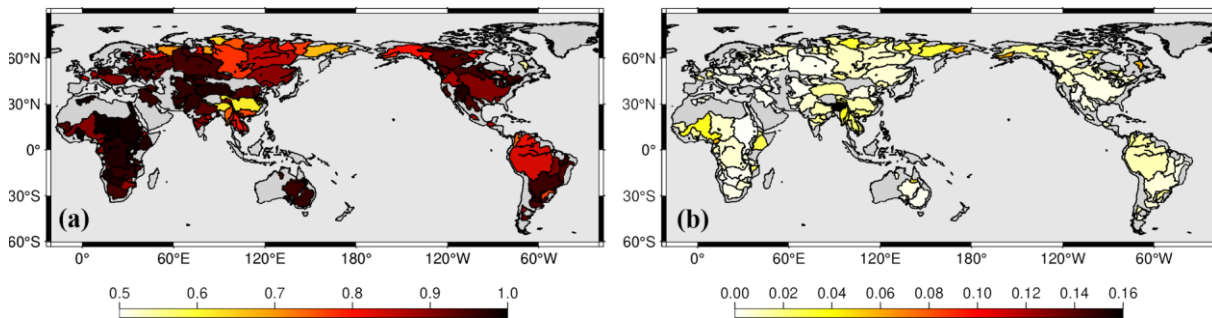
811 To examine spatial variability, Fig. 16 presents the global distribution of the 30-day retention  
812 factor ( $d^{30}$ ) and its associated uncertainty (Fig. 16). According to the model formulation, parameter  
813  $d$  in Eq. (17) represents the fraction of antecedent storage that remains in the system after a period  
814 of decay. From Eq. (10) it follows that, when  $d$  approaches 1, the storage at the previous time step  
815 is almost entirely preserved, implying that the current total loss ( $ET + R$ ) is primarily attributable  
816 to contemporaneous precipitation losses. Conversely, when  $d$  deviates from 1, current losses are  
817 also driven by the release of antecedent storage. This mechanism is consistent with the spatial pattern  
818 shown in Fig. 16a. In high-latitude regions of the Northern Hemisphere, the widespread presence of  
819 permafrost limits the infiltration of meltwater during the spring thaw, causing most of the water to  
820 be rapidly discharged as surface runoff. As a result, the calibrated parameter  $d$  is relatively low.

821 Notably, the Ob and Mackenzie River basins exhibit considerably higher water retention after  
822 30 days of decay compared to other high-latitude basins (Fig. 16a). In the Ob basin, approximately  
823 15% of the upstream and midstream region between the Irtysh River and the upper Ob is occupied  
824 by endorheic depressions, where inflow is retained and does not contribute to the main channel (Yi  
825 et al., 2023). Additionally, springtime ice jams, channel backwater, and overbank flooding further  
826 prolong the average residence time of water, resulting in a higher calibrated 30-day retention  
827 coefficient ( $d^{30}$ ). In the Mackenzie basin, the presence of multiple large lakes leads to a highly  
828 fragmented river network (Yi et al., 2023), impeding the continuous drainage of liquid water and  
829 thereby extending the residence time of water within the basin.



830

831 Figure 15: Comparison of the temporal evolution of the decay factor under different values of parameter  $d$  ( $t$  in  
832 days).



833

834 Figure 16: Spatial distribution of the 30-day retention factor  $d^{30}$  (a) and the uncertainty of parameter  $d^{30}$  (b)  
835 across 116 global river basins.

836

### 837 5.3 Limitations of the study

838 This approach relies on external precipitation and temperature forecasts for predicting future  
839 TWSA. In regions where precipitation and temperature forecasts are relatively accurate, the model  
840 is capable of producing reliable climate-driven TWSA predictions. Conversely, uncertainties in the  
841 forcing data can significantly affect the reconstruction results. Therefore, both the applicability and  
842 the upper limit of prediction accuracy are constrained by the quality of the forcing inputs. In addition,

843 the current model is designed to simulate deseasonalized and detrended TWSA only, and is not  
844 capable of independently reproducing the full seasonal component or long-term trend.

845 In permafrost- and glacier-dominated basins, model performance is further limited. Taking the  
846 Yukon Basin as an example (Fig. 6a), during the period from 2011 to 2014, the average temperatures  
847 during the snowmelt seasons were significantly lower than normal for three consecutive years, while  
848 annual precipitation slightly exceeded the long-term mean (Fig. S10). GRACE observations show  
849 that TWSA remained at a consistently high level during this period, whereas our reconstruction  
850 results are notably lower than those from GRACE. The primary source of error is likely associated  
851 with the sub-daily freezing and thawing cycles within the snowpack, which is not explicitly  
852 represented in the model. Meltwater generated during daytime warming is treated as immediate  
853 runoff and removed from the system, while at night, when temperatures drop below 0 °C, a portion  
854 of this water refreezes and is retained in the snowpack. Since the model does not incorporate this  
855 refreezing into the subsequent day's initial storage, it systematically underestimates the net  
856 accumulation within the snow layer. As a result, the peak in winter and spring is dampened, and the  
857 overall reconstructed TWSA is biased low.

858

## 859 **6 Conclusions**

860 Building upon the classical linear water storage framework, this study introduces coupled  
861 temperature and precipitation factors to develop a four-parameter daily recursive model for  
862 reconstructing climate-driven TWSA at a  $0.5^\circ \times 0.5^\circ$  spatial resolution over global land areas  
863 (excluding Antarctica) for the period 2002–2021. The model is calibrated against the JPL RL06 and  
864 GSFC RL06v2 mascon solutions, using hourly ERA5-Land precipitation and temperature data  
865 aggregated to daily time steps as forcing inputs.

866 By treating  $TWS(0)$  as an additional parameter to be estimated jointly with the other model  
867 parameters, the influence of initial condition errors is substantially reduced, leading to improved  
868 reconstruction accuracy. Independent calibrations conducted for 116 major river basins demonstrate  
869 that the model parameters are more stable compared to existing approaches, while the reduced  
870 number of calibration iterations significantly enhances computational efficiency. At the gridded  
871 scale, the reconstruction closely aligns with GRACE/GRACE-FO mascon observations, with 62%  
872 of grid cells achieving  $NSE > 0.5$  for JPL-REC and 53% for GSFC-REC. A sensitivity analysis  
873 further clarifies the physical interpretation of the four model parameters and their respective roles  
874 in governing hydrological processes.

875 Comparison with existing statistical products demonstrates that the proposed model performs  
876 well across multiple spatial scales. At the basin scale, reconstructed TWSA results were compared  
877 with those from Zhong-REC. Among the 116 global river basins examined, 84 achieved a NSE  
878 greater than 0.7 using the proposed model, whereas only 59 basins (51%) exceeded this threshold  
879 with Zhong-REC. In most basins, the model yields improved performance, with the most substantial  
880 gains observed in arid and semi-arid regions. At the gridded scale, the reconstructed results were  
881 also compared with those from the Humphrey's model. After subtracting the Humphrey-JPL-REC  
882 results from the JPL-REC reconstruction, 63% of the grid cells have  $\Delta NSE$  values greater than  
883 zero, indicating an overall improvement in model performance. The most pronounced improvements  
884 are observed across the Arabian Peninsula and large parts of Africa.

885 At the daily scale, the reconstructed TWSA agrees well overall with the ITSG-Grace2018 daily  
886 solution and GLDAS-2.2, indicating that the model can capture sub-monthly storage variability. The  
887 reconstruction demonstrates comparable performance during independent training and validation  
888 periods. The validation results show that about 60% of the basins retain  $NSE$  values greater than  
889 0.5, and in representative basins such as the Amazon, Yangtze, Mississippi, and Murray, the  
890 validation period reconstructions remain consistent with GRACE/GRACE-FO observations in both  
891 amplitude and phase. These results indicate that the proposed framework does not exhibit overfitting  
892 and has good temporal extrapolation capability.

893 In summary, this study develops a climate-driven model of global TWS variability at the  
894 interannual scale, based on a simple four-parameter linear response framework that relies solely on  
895 precipitation and temperature as inputs. This study introduces optimization strategies to improve  
896 both computational efficiency and reconstruction accuracy. The method can be further extended to  
897 incorporate additional factors such as wind speed, radiation, and topography, or to adopt nonlinear  
898 relationships using machine learning approaches. However, the simplicity and high computational  
899 efficiency demonstrated by the proposed framework greatly enhance its applicability and  
900 interpretability. Future research will focus on simulating the seasonal and long-term trends of water  
901 storage and ultimately aim to separate and quantify the relative impacts of climatic factors and  
902 human activities on water resources across multiple temporal scales.

903

## 904 **Competing interests**

905 The authors declare that they have no known competing financial interests or personal  
906 relationships that could have appeared to influence the work reported in this paper.

907

## 908 **Author contribution**

909 P.X. conducted the data curation, formal analysis, and investigation; developed the model and  
910 software; and prepared the original manuscript draft. S.Y. conceptualized the study, acquired the  
911 funding, and supervised the research. Both authors contributed to the methodology and participated  
912 in the review and editing of the manuscript.

## 913 **Acknowledgements**

914 This research is funded by the National Natural Science Foundation of China (42374103,  
915 E214040201), the University of Chinese Academy of Sciences Research Start-up Grant  
916 (E3ER0402A2, 110400M003).  
917

## 918 **Data availability**

919 The authors are grateful to the NASA Goddard Space Flight Center  
920 (<https://earth.gsfc.nasa.gov/geo/data/grace-mascons>), Jet Propulsion Laboratory  
921 ([https://podaac.jpl.nasa.gov/dataset/TELLUS\\_GRAC-  
922 GRFO\\_MASCON\\_CRI\\_GRID\\_RL06.1\\_V3](https://podaac.jpl.nasa.gov/dataset/TELLUS_GRAC-GRFO_MASCON_CRI_GRID_RL06.1_V3)), and the Center for Space Research, University of  
923 Texas at Austin (<http://www2.csr.utexas.edu/grace>) for providing the GRACE mascon solutions  
924 (last access: 31 July 2025). The authors thank Zhong et al. (2025) for providing publicly available  
925 datasets and source code that supported the reconstruction of terrestrial water storage anomalies,  
926 which are accessible from <https://doi.org/10.1038/s43247-024-01967-7>. The authors also are  
927 grateful for the water storage reconstruction dataset provided by Humphrey and Gudmundsson  
928 (2019) (<https://doi.org/10.6084/m9.figshare.7670849>). ERA5-Land precipitation, temperature and  
929 evapotranspiration datasets are provided by [https://doi.org/10.24381/cds.e2161bac  
930 63](https://doi.org/10.24381/cds.e2161bac63)(last access: 31 July 2025). MSWEP precipitation can be publicly obtained from [http://  
931 www.gloh2o.org/mswep/65](http://www.gloh2o.org/mswep/65)(last access: 31 July 2025). GLDAS-GLSM data can be publicly  
932 obtained from <https://disc.gsfc.nasa.gov/datasets> (last access: 31 July 2025). The boundary of global  
933 river basins is available at [https://www.bafg.de/SharedDocs/ExterneLinks/GRDC/mrb\\_shp\\_zip.  
934 html#62](https://www.bafg.de/SharedDocs/ExterneLinks/GRDC/mrb_shp_zip.html#62) (last access: 31 July 2025). Aridity index can be publicly obtained from  
935 <https://doi.org/10.6084/m9.figshare.7504448.v451> (last access: 31 July 2025).

937 **References**

- 938 Abbott, B. W., Bishop, K., Zarnetske, J. P., Minaudo, C., Chapin, F. S., Krause, S., Hannah, D.  
 939 M., Conner, L., Ellison, D., Godsey, S. E., Plont, S., Marçais, J., Kolbe, T., Huebner, A., Frei, R. J.,  
 940 Hampton, T., Gu, S., Buhman, M., Sara Sayedi, S., Ursache, O., Chapin, M., Henderson, K. D., and  
 941 Pinay, G.: Human domination of the global water cycle absent from depictions and perceptions, *Nat.*  
 942 *Geosci.*, 12, 533-540, 10.1038/s41561-019-0374-y, 2019.
- 943 An, L., Wang, J., Huang, J., Pokhrel, Y., Hugonnet, R., Wada, Y., Cáceres, D., Müller Schmied,  
 944 H., Song, C., Berthier, E., Yu, H., and Zhang, G.: Divergent Causes of Terrestrial Water Storage  
 945 Decline Between Drylands and Humid Regions Globally, *Geophys. Res. Lett.*, 48, e2021GL095035,  
 946 <https://doi.org/10.1029/2021GL095035>, 2021.
- 947 Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer,  
 948 C., Davis, K., and Evans, R.: FLUXNET: A new tool to study the temporal and spatial variability of  
 949 ecosystem-scale carbon dioxide, water vapor, and energy flux densities, *Bull. Am. Meteorol. Soc.*,  
 950 82, 2415-2434, 2001.
- 951 Beck, H. E., van Dijk, A. I. J. M., Levizzani, V., Schellekens, J., Miralles, D. G., Martens, B.,  
 952 and de Roo, A.: MSWEP: 3-hourly 0.25° global gridded precipitation (1979–2015) by merging  
 953 gauge, satellite, and reanalysis data, *Hydrol. Earth Syst. Sci.*, 21, 589-615, 10.5194/hess-21-589-  
 954 2017, 2017.
- 955 Beck, H. E., Wood, E. F., Pan, M., Fisher, C. K., Miralles, D. G., van Dijk, A. I. J. M., McVicar,  
 956 T. R., and Adler, R. F.: MSWEP V2 Global 3-Hourly 0.1° Precipitation: Methodology and  
 957 Quantitative Assessment, *Bull. Am. Meteorol. Soc.*, 100, 473-500, [https://doi.org/10.1175/BAMS-](https://doi.org/10.1175/BAMS-D-17-0138.1)  
 958 [D-17-0138.1](https://doi.org/10.1175/BAMS-D-17-0138.1), 2019.
- 959 Burek, P. and Smilovic, M.: The use of GRDC gauging stations for calibrating large-scale  
 960 hydrological models, *Earth Syst. Sci. Data*, 15, 5617-5629, 2023.
- 961 Castle, S. L., Thomas, B. F., Reager, J. T., Rodell, M., Swenson, S. C., and Famiglietti, J. S.:  
 962 Groundwater depletion during drought threatens future water security of the Colorado River Basin,  
 963 *Geophys. Res. Lett.*, 41, 5904-5911, 2014.
- 964 Chen, J., Cazenave, A., Dahle, C., Llovel, W., Panet, I., Pfeffer, J., and Moreira, L.:  
 965 Applications and challenges of GRACE and GRACE follow-on satellite gravimetry, *Surv. Geophys.*,  
 966 43, 305-345, 2022.
- 967 Chen, J. L., Wilson, C. R., and Tapley, B. D.: The 2009 exceptional Amazon flood and  
 968 interannual terrestrial water storage change observed by GRACE, *Water Resour. Res.*, 46, 2010.
- 969 Chen, M., Shi, W., Xie, P., Silva, V. B., Kousky, V. E., Wayne Higgins, R., and Janowiak, J. E.:  
 970 Assessing objective techniques for gauge - based analyses of global daily precipitation, *J. Geophys.*  
 971 *Res.-Atmos.*, 113, 2008.
- 972 Crochemore, L., Isberg, K., Pimentel, R., Pineda, L., Hasan, A., and Arheimer, B.: Lessons  
 973 learnt from checking the quality of openly accessible river flow data worldwide, *Hydrol. Sci. J.*, 65,  
 974 699-711, 2020.
- 975 Duan, A., Zhong, Y., Xu, G., Yang, K., Tian, B., Wu, Y., Bai, H., and Hu, E.: Quantifying the  
 976 2022 extreme drought in the Yangtze River Basin using GRACE-FO, *J. Hydrol.*, 630, 130680, 2024.
- 977 Fan, C., Song, C., Liu, K., Ke, L., Xue, B., Chen, T., Fu, C., and Cheng, J.: Century - scale  
 978 reconstruction of water storage changes of the largest lake in the inner Mongolia plateau using a  
 979 machine learning approach, *Water Resour. Res.*, 57, e2020WR028831, 2021.
- 980 Fekete, B. M., Looser, U., Pietroniro, A., and Robarts, R. D.: Rationale for monitoring  
 981 discharge on the ground, *J. Hydrometeorol.*, 13, 1977-1986, 2012.

982 Feng, J., Li, B., Song, J., Tang, B., Nyein, M. M., and Tani, B. P.: Spatiotemporal Variations of  
983 Terrestrial Water Storage and Driving Factors in the Water Towers of Northwest China Based on  
984 GRACE and Multi-Source Data Sets, *Water Resour. Res.*, 61, e2024WR039490,  
985 <https://doi.org/10.1029/2024WR039490>, 2025.

986 Fernández-Alvarez, J. C., Nieto, R., Vicente-Serrano, S. M., Carvalho, D., and Gimeno, L.:  
987 Moisture and temperature sources were key drivers of the anomalies for the record-breaking of 2023  
988 Amazon drought, *Commun. Earth Environ.*, 6, 801, 10.1038/s43247-025-02771-7, 2025.

989 Forootan, E., Khaki, M., Schumacher, M., Wulfmeyer, V., Mehrnegar, N., van Dijk, A. I.,  
990 Brocca, L., Farzaneh, S., Akinluyi, F., and Ramillien, G.: Understanding the global hydrological  
991 droughts of 2003–2016 and their relationships with teleconnections, *Sci. Total Environ.*, 650, 2587-  
992 2604, 2019.

993 Gao, Y., Luo, Z., Liu, H., Wang, L., Chen, X., and Li, H.: Reconstruction of global long-term  
994 daily streamflow dataset using machine learning models for revealing streamflow changes, *Journal  
995 of Hydrology: Regional Studies*, 64, 103148, 2026.

996 Humphrey, V. and Gudmundsson, L.: GRACE-REC: a reconstruction of climate-driven water  
997 storage changes over the last century, *Earth Syst. Sci. Data*, 11, 1153-1170, 10.5194/essd-11-1153-  
998 2019, 2019.

999 Humphrey, V., Gudmundsson, L., and Seneviratne, S. I.: Assessing Global Water Storage  
1000 Variability from GRACE: Trends, Seasonal Cycle, Subseasonal Anomalies and Extremes, *Surv.  
1001 Geophys.*, 37, 357-395, 10.1007/s10712-016-9367-1, 2016.

1002 Humphrey, V., Rodell, M., and Eicker, A.: Using Satellite-Based Terrestrial Water Storage Data:  
1003 A Review, *Surv. Geophys.*, 44, 1489-1517, 10.1007/s10712-022-09754-9, 2023.

1004 Laudon, H., Spence, C., Buttle, J., Carey, S. K., McDonnell, J. J., McNamara, J. P., Soulsby,  
1005 C., and Tetzlaff, D.: Save northern high-latitude catchments, *Nat. Geosci.*, 10, 324-325, 2017.

1006 Li, B., Rodell, M., Kumar, S., Beaudoin, H. K., Getirana, A., Zaitchik, B. F., de Goncalves, L.  
1007 G., Cossetin, C., Bhanja, S., Mukherjee, A., Tian, S., Tangdamrongsub, N., Long, D., Nanteza, J.,  
1008 Lee, J., Policelli, F., Goni, I. B., Daira, D., Bila, M., de Lannoy, G., Mocko, D., Steele-Dunne, S. C.,  
1009 Save, H., and Bettadpur, S.: Global GRACE Data Assimilation for Groundwater and Drought  
1010 Monitoring: Advances and Challenges, *Water Resour. Res.*, 55, 7564-7586,  
1011 <https://doi.org/10.1029/2018WR024618>, 2019.

1012 Li, F., Kusche, J., Chao, N., Wang, Z., and Löcher, A.: Long - term (1979 - present) total water  
1013 storage anomalies over the global land derived by reconstructing GRACE data, *Geophys. Res. Lett.*,  
1014 48, e2021GL093492, 2021.

1015 Li, F., Kusche, J., Rietbroek, R., Wang, Z., Forootan, E., Schulze, K., and Lück, C.: Comparison  
1016 of data - driven techniques to reconstruct (1992 - 2002) and predict (2017 - 2018) GRACE - like  
1017 gridded total water storage changes using climate inputs, *Water Resour. Res.*, 56, e2019WR026551,  
1018 2020.

1019 Li, X., Jin, T., Liu, B., Chao, N., Li, F., and Cai, Z.: The Influence of ENSO on the Long -  
1020 Term Water Storage Anomalies in the Middle - Lower Reaches of the Yangtze River Basin:  
1021 Evaluation and Analysis, *Earth Space Sci.*, 10, e2023EA003007, 2023.

1022 Liu, B., Zou, X., Yi, S., Sneeuw, N., Cai, J., and Li, J.: Identifying and separating climate-and  
1023 human-driven water storage anomalies using GRACE satellite data, *Remote Sens. Environ.*, 263,  
1024 112559, 2021.

1025 Liu, B., Zou, X., Yi, S., Sneeuw, N., Li, J., and Cai, J.: Reconstructing GRACE-like time series  
1026 of high mountain glacier mass anomalies, *Remote Sens. Environ.*, 280, 113177, 2022.

1027 Liu, X., Feng, X., Ciais, P., and Fu, B.: Widespread decline in terrestrial water storage and its  
1028 link to teleconnections across Asia and eastern Europe, *Hydrol. Earth Syst. Sci.*, 24, 3663-3676,  
1029 10.5194/hess-24-3663-2020, 2020.

1030 Long, D., Longuevergne, L., and Scanlon, B. R.: Global analysis of approaches for deriving

1031 total water storage changes from GRACE satellites, *Water Resour. Res.*, 51, 2574-2594, 2015.

1032 Long, D., Shen, Y., Sun, A., Hong, Y., Longuevergne, L., Yang, Y., Li, B., and Chen, L.:

1033 Drought and flood monitoring for a large karst plateau in Southwest China using extended GRACE

1034 data, *Remote Sens. Environ.*, 155, 145-160, 2014.

1035 Loomis, B., Luthcke, S., and Sabaka, T.: Regularization and error characterization of GRACE

1036 mascons, *Journal of geodesy*, 93, 1381-1398, 2019.

1037 Ma, N., Zhang, Y., and Szilagyi, J.: Water-balance-based evapotranspiration for 56 large river

1038 basins: A benchmarking dataset for global terrestrial evapotranspiration modeling, *J. Hydrol.*, 630,

1039 130607, 2024.

1040 Muñoz-Sabater, J.: ERA5-Land hourly data from 1981 to present, Copernicus climate change

1041 service (C3S) climate data store (CDS), 10, 2019.

1042 Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G.,

1043 Boussetta, S., Choulga, M., Harrigan, S., and Hersbach, H.: ERA5-Land: A state-of-the-art global

1044 reanalysis dataset for land applications, *Earth Syst. Sci. Data*, 13, 4349-4383, 2021.

1045 Ni, S., Chen, J., Wilson, C. R., Li, J., Hu, X., and Fu, R.: Global Terrestrial Water Storage

1046 Changes and Connections to ENSO Events, *Surv. Geophys.*, 39, 1-22, 10.1007/s10712-017-9421-7,

1047 2018.

1048 Palazzoli, I., Ceola, S., and Gentile, P.: GRACE: reconstructing terrestrial water storage

1049 anomalies with recurrent neural networks, *Sci. Data*, 12, 146, 2025.

1050 Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y.-W., Poindexter, C.,

1051 Chen, J., Elbashandy, A., and Humphrey, M.: The FLUXNET2015 dataset and the ONEFlux

1052 processing pipeline for eddy covariance data, *Sci. Data*, 7, 225, 2020.

1053 Riegger, J. and Tourian, M. J.: Characterization of runoff - storage relationships by satellite

1054 gravimetry and remote sensing, *Water Resour. Res.*, 50, 3444-3466, 2014.

1055 Rodell, M., Velicogna, I., and Famiglietti, J. S.: Satellite-based estimates of groundwater

1056 depletion in India, *Nature*, 460, 999-1002, 2009.

1057 Rodell, M., Famiglietti, J. S., Wiese, D. N., Reager, J., Beaudoing, H. K., Landerer, F. W., and

1058 Lo, M.-H.: Emerging trends in global freshwater availability, *Nature*, 557, 651-659, 2018.

1059 Save, H., Bettadpur, S., and Tapley, B. D.: High - resolution CSR GRACE RL05 mascons,

1060 *Journal of Geophysical Research: Solid Earth*, 121, 7547-7569, 2016.

1061 Sun, A. Y., Scanlon, B. R., Save, H., and Rateb, A.: Reconstruction of GRACE total water

1062 storage through automated machine learning, *Water Resour. Res.*, 57, e2020WR028666, 2021.

1063 Sun, Z., Long, D., Yang, W., Li, X., and Pan, Y.: Reconstruction of GRACE data on changes in

1064 total water storage over the global land surface and 60 basins, *Water Resour. Res.*, 56,

1065 e2019WR026250, 2020.

1066 Tapley, B. D., Watkins, M. M., Flechtner, F., Reigber, C., Bettadpur, S., Rodell, M., Sasgen, I.,

1067 Famiglietti, J. S., Landerer, F. W., and Chambers, D. P.: Contributions of GRACE to understanding

1068 climate change, *Nature climate change*, 9, 358-369, 2019.

1069 Tian, K., Wang, Z., Li, F., Gao, Y., Xiao, Y., and Liu, C.: Drought Events over the Amazon

1070 River Basin (1993–2019) as Detected by the Climate-Driven Total Water Storage Change, *Remote*

1071 *Sens.*, 13, 1124, 2021.

1072 Trabucco, A. and Zomer, R. J.: Global Aridity Index and Potential Evapotranspiration (ET0)

1073 Climate Database v3, International Center for Tropical Agriculture (CIAT) and World Agroforestry

1074 Centre (ICRAF) [dataset], 10.6084/m9.figshare.7504448.v3, 2019.

1075 Voss, K. A., Famiglietti, J. S., Lo, M., De Linage, C., Rodell, M., and Swenson, S. C.:

1076 Groundwater depletion in the Middle East from GRACE with implications for transboundary water

1077 management in the Tigris - Euphrates - Western Iran region, *Water Resour. Res.*, 49, 904-914, 2013.

1078 Wiese, D., Yuan, D., Boening, C., Landerer, F. W., and Watkins, M.: JPL GRACE and GRACE-

1079 FO Mascon Ocean, Ice, and Hydrology Equivalent Water Height Coastal Resolution Improvement

1080 (CRI) Filtered Release 06 Version 02, DAAC: Pasadena, CA, USA, 2019.

1081 Xiao, C., Zhong, Y., Wu, Y., Zhang, Z., Bai, H., and Li, Z.: Flood Evolution in the Past 60 Years

1082 Revealed by Reconstructed Daily Terrestrial Water Storage Anomalies in China, *Water Resour. Res.*,

1083 61, e2024WR038712, <https://doi.org/10.1029/2024WR038712>, 2025.

1084 Xu, L., Chen, N., Moradkhani, H., Zhang, X., and Hu, C.: Improving global monthly and daily

1085 precipitation estimation by fusing gauge observations, remote sensing, and reanalysis data sets,

1086 *Water Resour. Res.*, 56, e2019WR026444, 2020.

1087 Yi, S., Saemian, P., Sneeuw, N., and Tourian, M. J.: Estimating runoff from pan-Arctic drainage

1088 basins for 2002–2019 using an improved runoff-storage relationship, *Remote Sens. Environ.*, 298,

1089 113816, 2023.

1090 Yi, S., Sun, W., Feng, W., and Chen, J.: Anthropogenic and climate-driven water depletion in

1091 Asia, *Geophys. Res. Lett.*, 43, 9061-9069, <https://doi.org/10.1002/2016GL069985>, 2016.

1092 Yin, J., Slater, L. J., Khouakhi, A., Yu, L., Liu, P., Li, F., Pokhrel, Y., and Gentine, P.: GTWS-

1093 MLrec: global terrestrial water storage reconstruction by machine learning from 1940 to present,

1094 *Earth Syst. Sci. Data*, 15, 5597-5615, 2023.

1095 Zhang, D., Zhang, Q., Werner, A. D., and Liu, X.: GRACE-based hydrological drought

1096 evaluation of the Yangtze River Basin, China, *J. Hydrometeorol.*, 17, 811-828, 2016.

1097 Zhang, J., Liesch, T., and Goldscheider, N.: Impacts of climate change and human activities on

1098 global groundwater storage from 2003 to 2022, *J. Hydrol.*, 664, 134298,

1099 <https://doi.org/10.1016/j.jhydrol.2025.134298>, 2025.

1100 Zhang, L. and Sun, W.: Progress and prospect of GRACE Mascon product and its application,

1101 *Reviews of Geophysics and Planetary Physics*, 53, 35-52, 2022.

1102 Zhang, Z., Chao, B., Chen, J., and Wilson, C.: Terrestrial water storage anomalies of Yangtze

1103 River Basin droughts observed by GRACE and connections with ENSO, *Glob. Planet. Change*, 126,

1104 35-45, 2015.

1105 Zheng, S., Zhang, Z., Yan, H., Zhao, Y., and Li, Z.: Characterizing drought events occurred in

1106 the Yangtze River Basin from 1979 to 2017 by reconstructing water storage anomalies based on

1107 GRACE and meteorological data, *Sci. Total Environ.*, 868, 161755,

1108 <https://doi.org/10.1016/j.scitotenv.2023.161755>, 2023.

1109 Zhong, Y., Feng, W., Humphrey, V., and Zhong, M.: Human-Induced and Climate-Driven

1110 Contributions to Water Storage Variations in the Haihe River Basin, China, *Remote Sens.*, 11, 3050,

1111 2019.

1112 Zhong, Y., Tian, B., Kim, H., Yuan, X., Liu, X., Zhu, E., Wu, Y., Wang, L., and Wang, L.: Over

1113 60% precipitation transformed into terrestrial water storage in global river basins from 2002 to 2021,

1114 *Commun. Earth Environ.*, 6, 53, 2025.

1115 Zomer, R. J., Xu, J., and Trabucco, A.: Version 3 of the Global Aridity Index and Potential

1116 Evapotranspiration Database, *Sci. Data*, 9, 409, [10.1038/s41597-022-01493-1](https://doi.org/10.1038/s41597-022-01493-1), 2022.

1117