nature water

Article https://doi.org/10.1038/s44221-024-00194-w

# Global high-resolution total water storage anomalies from self-supervised data assimilation using deep learning algorithms

In the format provided by the authors and unedited

## Contents

# 1 Introduction

This document includes the results of extended experiments to investigate the quality of our downscaled product. The downscaled product are compared with two other high-resolution TWSA products, WaterGAP Hydrology Model (WGHM)[1] and global land water storage data set release 2 (GLWS)[2] to highlight the feasibility of our approach. Both of them have the same spatial resolution as our downscaled product (0.5 degrees). In Section 2, we examine the performance of the three products in conserving global mass. Section 3 reports the results of basin-wise evaluation, focusing on the reliability of temporal information, such as long-term trends and annual phases. In Section 4, we introduce satellite altimetry to validate the three products externally. In the end, the generalizability of our deep learning model is discussed in Section 5.

# 2 Mass conservation of global terrestrial water storage

The Gravity Recovery and Climate Experiment (GRACE) and its follow-on (GRACE-FO) missions (hereafter GRACE) measure the global total water storage (TWS) with high accuracy. Therefore, the downscaled TWS anomaly (TWSA) product should preserve mass conservation globally. Fig. 1a) shows the global terrestrial water storage variations estimated from four different products. Our downscaled product agrees well with the GRACE measurements, both in terms of seasonality and long-term trends. On the contrary, WGHM and GLWS products differ from GRACE measurements due to their underestimation of the decreasing trend after 2012. The differences between the high-resolution TWSA products and GRACE measurements are shown in Fig. 1b), where we observe clear seasonality in the errors of WGHM and GLWS products. In contrast, the seasonal errors are not present in the downscaled TWSAs, with only a minor error of around $-400$ Gt remaining.

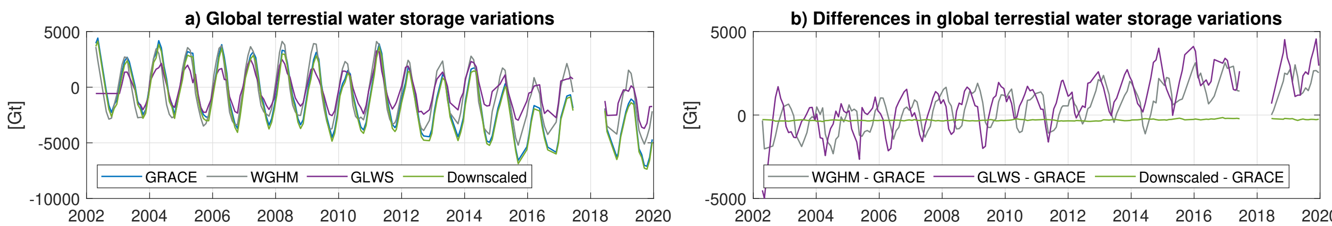


**Figure 1.** Global total water storage variations. a) Global total water storage variations estimated from GRACE, WGHM, GLWS, and the downscaled product. b) Differences between the estimations using WGHM, GLWS or the downscaled product with GRACE measurements.

To further compare the spatial variability of the three products with GRACE measurements, we downsampled the three high-resolution products to mascon-like products by computing the mascon-wise average values. Then, the mean errors (ME) and root-mean-square errors (RMSE) between the two mascon-like products and JPL mascons are computed, as shown in Fig. 2. We note that the downsamples mascon-like products are not entirely equivalent to the JPL mascons but only provide intuition about the spatial variability without limitation of basin boundaries. The ME patterns of the downscaled product are similar to the ones of the WGHM and GLWS products but with generally smaller magnitudes, indicating the benefits of the proposed method for providing more accurate values. Differences between the patterns are found in Alaska and Siberia, where the downscaled product has negative MEs. These errors may explain the systematic negative bias shown in Fig. 1b).

In terms of RMSE, the downscaled product agrees better with the GRACE measurements than WGHM or GLWS TWSAs (Fig. 2b), especially in the Middle East, Central Africa, and South Asia. Two hot spots remain in the downscaled product: the Amazon basin and Alaska. The considerable disagreements in the Amazon basin are related to the relatively large amplitudes, which result in significant uncertainties. The errors in the Alaska region are caused by the insufficient modelling of glaciers in WGHM simulations, as discussed in the main text. The deep learning model tries to constrain mass conservation over a region of $16 \times 16$ degrees, which may cause leakage errors in the neighbouring regions. Therefore, the RMSEs in some non-glaciated areas are more significant than the WGHM simulations, but they are smaller in the glaciated regions of Alaska.

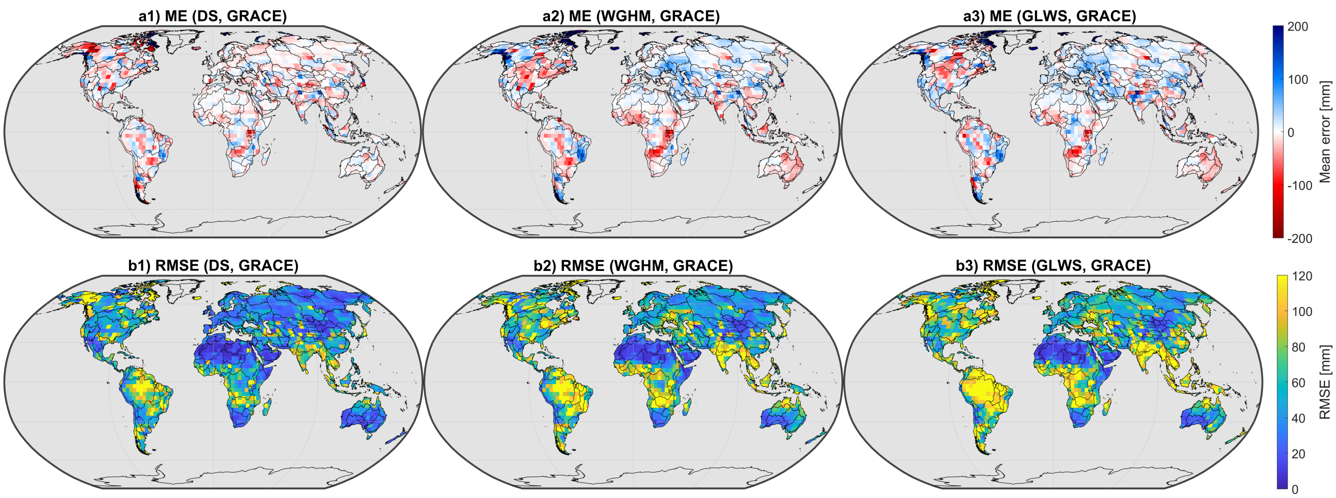


**Figure 2.** Comparison of GRACE TWSAs with the downscaled (DS), WGHM and GLWS TWSAs resampled to the mascon cells. a) The mascon-wise mean errors. b) The mascon-wise root-mean-square errors.

## 3 Basin-wise evaluation

In this section, we report results on examining the mass conservation compared to GRACE measurements on the basin scale. Fig. 3 and 4 indicate the long-trend trends and signal phases of the GRACE TWSAs and the three high-resolution products in 160 major basins larger than 200 000 $km^2$. The downscaled trends agree better with GRACE measurements than WGHM simulations, especially in the Arctic basin, the centre of South America, the centre of Africa, the Middle East and Siberia. These results demonstrate the ability of our downscaling algorithm to find the optimized relationship based on global TWSA products. The most dominant trend errors are found in the glaciated regions of Alaska (Fig. 3b). Our downscaled product and GLWS underestimate the decreasing trends (Fig. 3c, whereas the WGHM simulations do not reflect the trends at all (Fig. 3a and b2). The results indicate the benefits of assimilating GRACE measurements to the high-resolution products. However, the trend errors of GLWS are still larger than our downscaled product, proving the performance of our method.

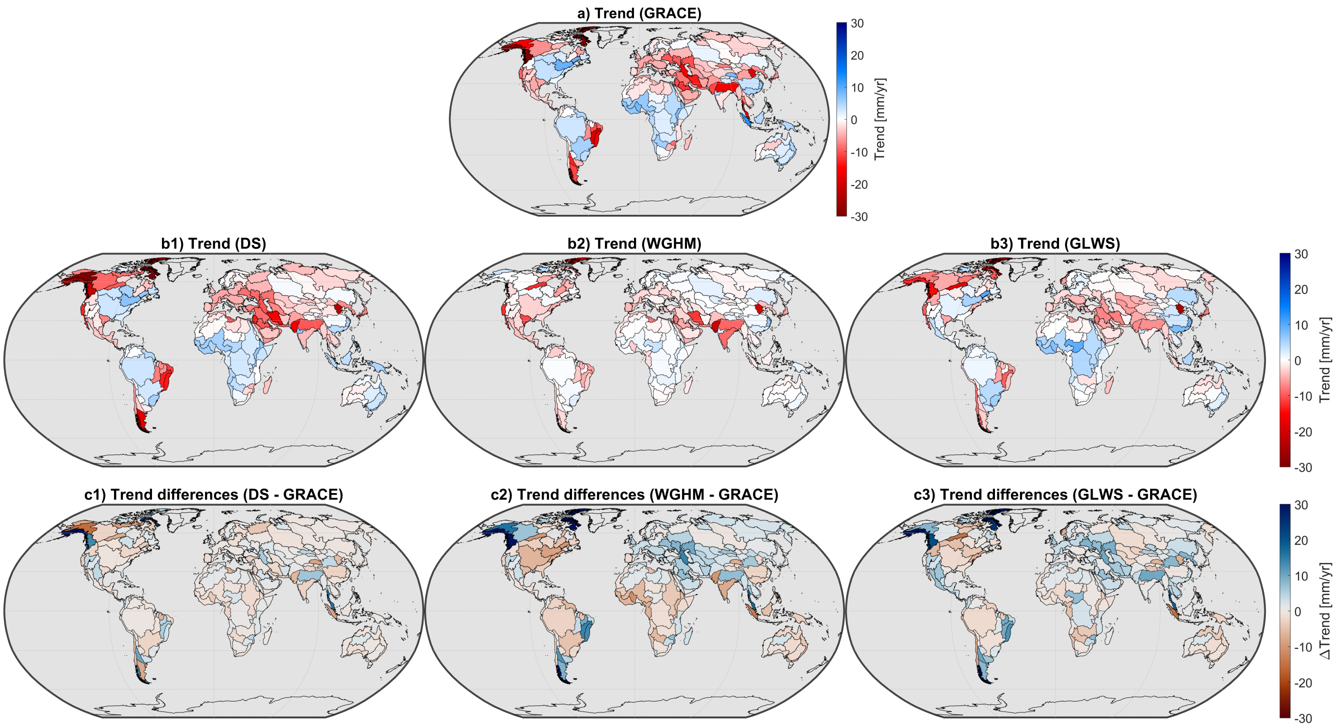


**Figure 3.** Comparison of basin-wise long-term trends. a) The long-term trends estimated from GRACE measurements. b) The long-term trends estimated from the three high-resolution TWSA products. c) The differences in long-term trends between the high-resolution TWSA products and GRACE measurements.

We do not observe apparent phase shifts between our downscaled TWSAs and GRACE measurements in the seasonal domain (Fig. 4a and b1). The only significant differences are found in the arid region of the north of Africa (Fig. 4c1), where the peaks may not be clearly defined due to the weakness of hydrological signals. Both WGHM and GLWS show some systematic errors of two to three months in many basins, which indicate potential shifts in their modelling of seasonal terms. These phase shifts may be one of the reasons behind the periodicity in Fig. 1b).

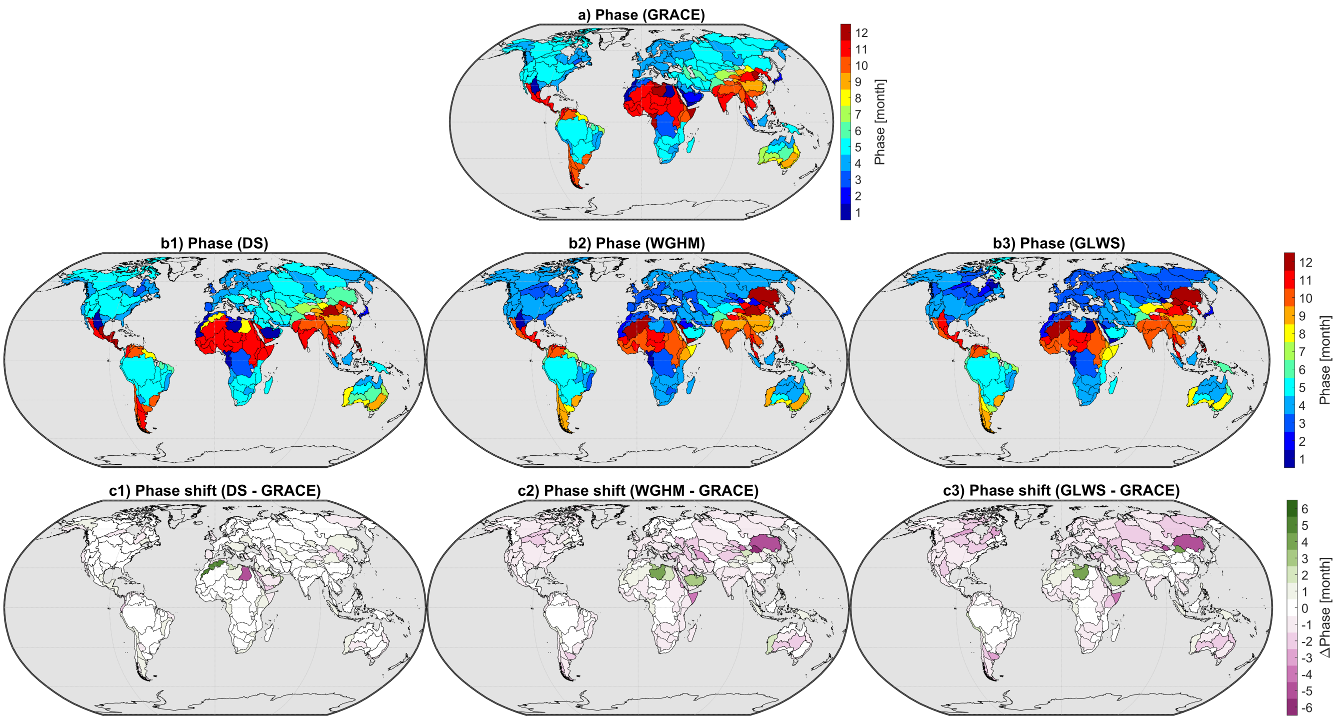


**Figure 4.** Comparison of basin-wise peak month of annual signals. a) The peak months from GRACE measurements. b) The peak months from the three high-resolution TWSA products. c) The phase shifts between the high-resolution TWSA products and GRACE measurements.

# 4 Validation against satellite altimetry

In this section, we compare the high-resolution TWSA products with the inland satellite altimetry measurements over ten major river systems to prove the quality of the high-resolution information. Since the river dynamics should dominate the variations in TWSAs at the river cells[3–5], relatively high correlations are expected in these cells. The water level changes of global rivers measured by multiple satellite altimetry missions are available at Hydroweb[6,7]. As we cannot accurately estimate the surface water storage changes due to the lack of high-quality river profiles, we rely on the water level changes. To quantify non-linear monotonic relationships, we computed Spearman's rank correlations ($\rho$) between TWSAs and water level changes in ten major river basins. The spatial patterns of individual measuring points are shown in Fig. 5 with statistical results given in Table 1. The basin-wise mean and median correlations are computed based on correlations at individual virtual stations within the same basin. In most cases, the downscaled TWSAs give the most statistically significant correlations ($\alpha < 0.05$, two-sided) and the highest basin-wise mean and median correlations, which reveals the benefit of the accurate high-resolution information. Most remarkable improvements compared to GRACE measurements can be found in the Lena, Yangtze, and Parana rivers, the dynamics of which are not effectively resolved by the original GRACE resolution. The high-resolution products can separate the river signals from their surrounding environments to a certain degree.

**Table 1.** Statistics of Spearman's rank correlations between the water level measured by satellite altimetry missions and the three TWSA products in ten major river basins from 2002 to 2019. The confidence level for the statistical significance test is 95 %. The best performance is highlighted.

| Basin | Total cells | t-test significant ($\alpha < 0.05$) | | | | basin-wise median of $\rho_i$ | | | | basin-wise mean of $\rho_i$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | DS | GRACE | WGHM | GLWS | DS | GRACE | WGHM | GLWS | DS | GRACE | WGHM | GLWS |
| Amazonas | 493 | **486** | 483 | 483 | 437 | **0.84** | 0.82 | **0.84** | 0.69 | **0.80** | 0.78 | 0.79 | 0.64 |
| Mississippi | 9 | 6 | **8** | 6 | 2 | 0.40 | **0.46** | 0.37 | 0.32 | 0.40 | **0.44** | 0.33 | 0.28 |
| Congo | 163 | **154** | 147 | 152 | 119 | **0.75** | 0.67 | 0.72 | 0.50 | **0.72** | 0.63 | 0.67 | 0.46 |
| Lena | 16 | **12** | 0 | **12** | 11 | **0.55** | 0.12 | 0.51 | 0.46 | 0.48 | 0.14 | **0.50** | 0.40 |
| Yangtze | 13 | **12** | 8 | **12** | **12** | **0.59** | 0.33 | 0.52 | 0.42 | **0.50** | 0.32 | 0.48 | 0.42 |
| Nile | 12 | **9** | 5 | 7 | 5 | **0.36** | 0.25 | 0.24 | 0.22 | **0.45** | 0.38 | 0.32 | 0.28 |
| GBM | 41 | **39** | 38 | 38 | **39** | **0.83** | 0.78 | 0.79 | 0.70 | **0.77** | 0.68 | 0.76 | 0.66 |
| Niger | 15 | **14** | 13 | 12 | 12 | **0.85** | 0.77 | 0.81 | 0.59 | **0.75** | 0.67 | 0.68 | 0.51 |
| Danube | 20 | **18** | **18** | 16 | 14 | **0.63** | 0.61 | 0.58 | 0.43 | **0.61** | 0.55 | 0.52 | 0.40 |
| Parana | 26 | **26** | 15 | 25 | 15 | 0.65 | 0.42 | **0.70** | 0.37 | 0.64 | 0.44 | **0.65** | 0.34 |

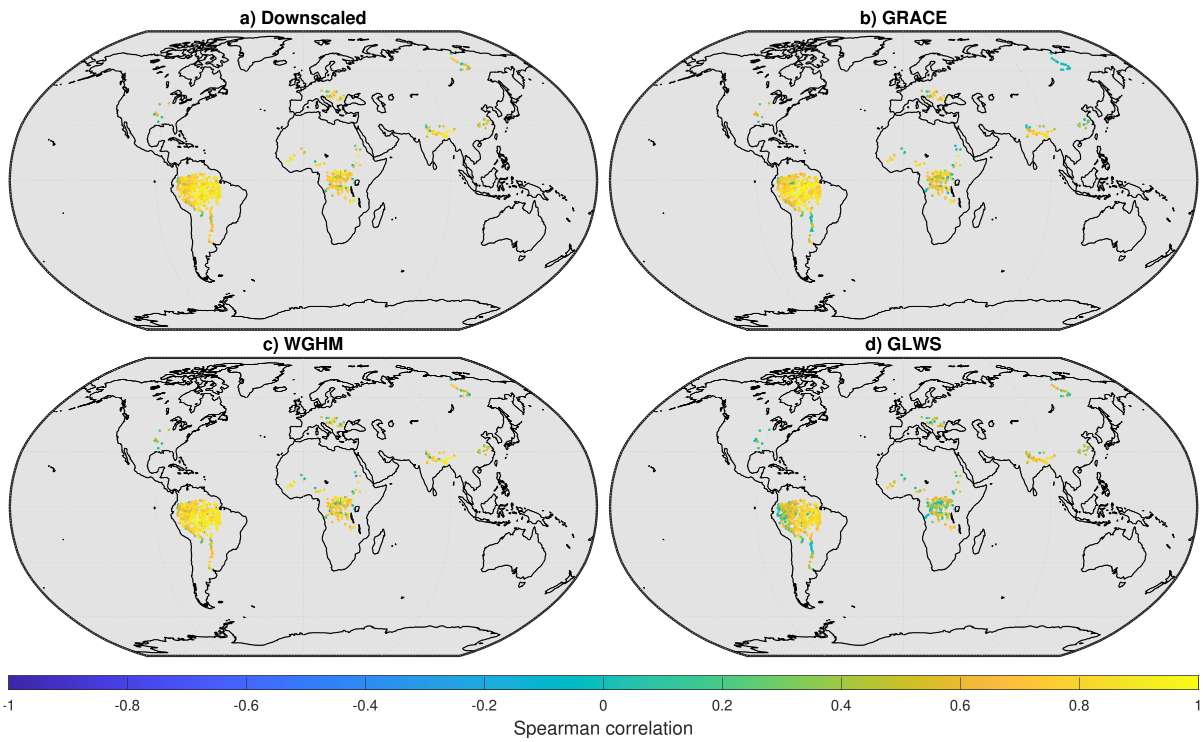


**Figure 5.** Spearman's rank correlations between the four TWSA products and water level anomalies measured by multiple satellite altimetry missions in ten major river basins.

## 5 Generalizability of the global model

A unique contribution of the proposed algorithm is its high generalizability, which allows us to obtain high-quality global monthly TWSAs over 17 years using one neural network. In this study, we also trained basin-wise and continental models whose training data only included one specific basin or one continent. The generalizability of the global model is proven by comparing its results with the other spatially more constrained models. We note that the basin-wise and continental models have simpler architectures than the global model to reduce the number of trainable parameters due to the limited training samples. Fig. 6 depicts the results in the Amazon basin over different seasons. First, we observe that the degradation of the continental model compared to the regional model is negligible. Since the Amazon basin is the dominant basin in South America, the signals from the Amazon basin also play an important role in the continental model. For the global model, the drastically increased amount of training samples is sufficient to train a larger neural network with better capacity. Therefore, the global model achieves impressive generalizability for the whole globe while the refined local structures are retained. We observe that the Amazon River and its smaller branches are partially smoothed out compared to the basin-wise model. We consider this issue as a reasonable compromise considering the advantages of global coverage. Fig. 6 also provides visual evidence that our proposed method remarkably compresses the noise of the WGHM simulations. The anomalies along the western boundary of the Amazon basin are clearly reduced.

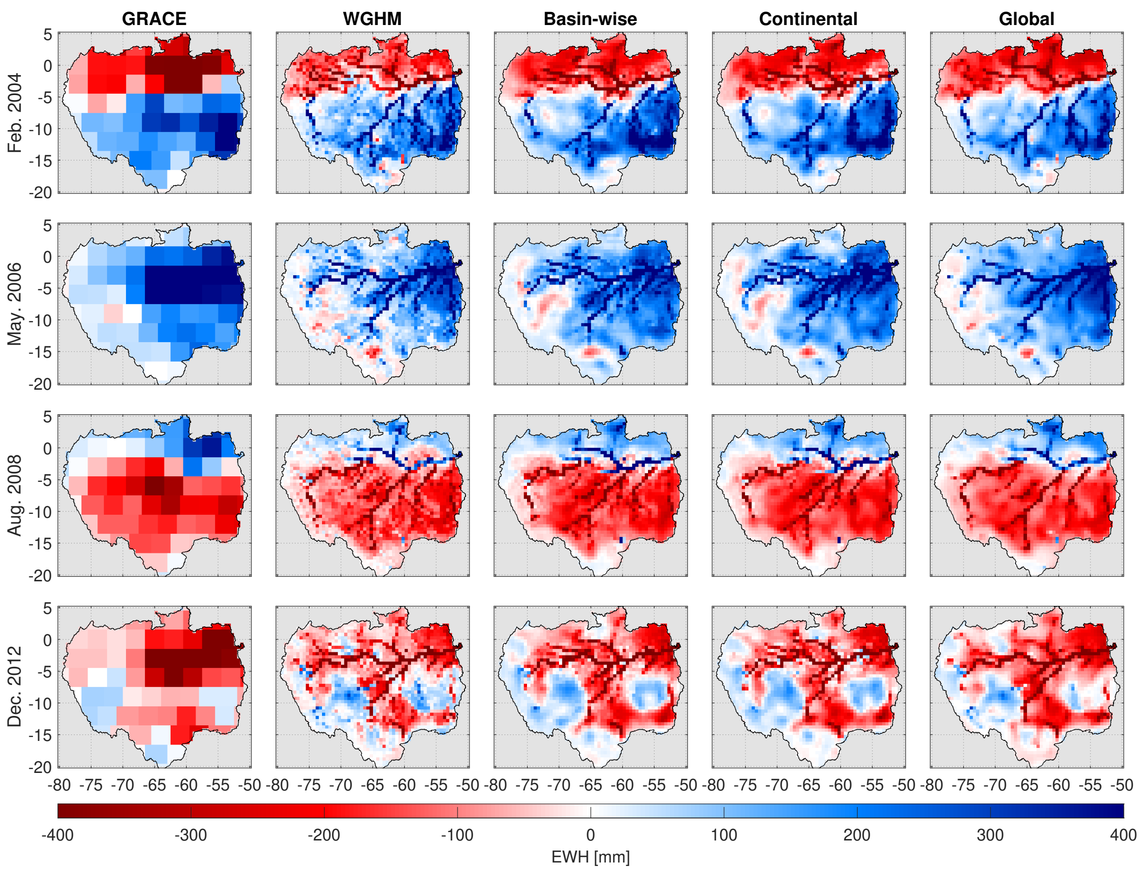


**Figure 6.** The TWSAs in the Amazon basin from five different sources. The columns from left to right represent the TWSAs from GRACE measurements, WGHM simulations, basin-wise, continental, and global deep learning models, respectively. The four rows represent four months in different seasons over the GRACE mission interval.

## Data availability

The raw data used in this study are publicly available. The water level changes products from multiple satellite altimetry missions are available at Hydroweb (https://hydroweb.theia-land.fr/). The downscaled TWSA product generated in this study is publicly available at https://doi.org/10.3929/ethz-b-000648738[8]. he WGHM TWSA simulations are available at https://doi.org/10.1594/PANGAEA.948461[9]. The global land water storage data set release 2 (GLWS) is available at https://doi.org/10.1594/PANGAEA.954742[10].

## References


1. Müller Schmied, H. *et al.* The global water resources and use model WaterGAP v2.2d: Model description and evaluation. *Geosci. Model. Dev.* **14**, 1037–1079, DOI: https://doi.org/10.5194/gmd-14-1037-2021 (2021).
2. Gerdener, H., Kusche, J., Schulze, K., Döll, P. & Klos, A. The global land water storage data set release 2 (GLWS2. 0) derived via assimilating GRACE and GRACE-FO data into a global hydrological model. *J. Geod.* **97**, 73, DOI: https://doi.org/10.1007/s00190-023-01763-9 (2023).
3. Tourian, M. J. *et al.* A copula-supported Bayesian framework for spatial downscaling of GRACE-derived terrestrial water storage flux. *Remote. Sens. Environ.* **295**, 113685, DOI: https://doi.org/10.1016/j.rse.2023.113685 (2023).
4. Getirana, A., Kumar, S., Girotto, M. & Rodell, M. Rivers and floodplains as key components of global terrestrial water storage variability. *Geophys. Res. Lett.* **44**, 10–359, DOI: https://doi.org/10.1002/2017GL074684 (2017).
5. Kim, H., Yeh, P. J.-F., Oki, T. & Kanae, S. Role of rivers in the seasonal variations of terrestrial water storage over global basins. *Geophys. Res. Lett.* **36**, DOI: https://doi.org/10.1029/2009GL039006 (2009).
6. Da Silva, J. S. *et al.* Water levels in the Amazon basin derived from the ERS 2 and ENVISAT radar altimetry missions. *Remote. sensing environment* **114**, 2160–2181, DOI: https://doi.org/10.1016/j.rse.2010.04.020 (2010).
7. Normandin, C. *et al.* Evolution of the performances of radar altimetry missions from ERS-2 to Sentinel-3A over the Inner Niger Delta. *Remote. Sens.* **10**, 833, DOI: https://doi.org/10.3390/rs10060833 (2018).
8. Gou, J. & Soja, B. GRACE-SeDA: A global total water storage anomaly product with a spatial resolution of 0.5 degrees from self-supervised data assimilation. ETH Research Collection, DOI: https://doi.org/10.3929/ethz-b-000648738 (2023).
9. Müller Schmied, H. *et al.* The global water resources and use model WaterGAP v2.2d - Alternative model output driven with gswp3-w5e5. PANGAEA, DOI: https://doi.org/10.1594/PANGAEA.948461 (2022).
10. Gerdener, H., Schulze, K. & Kusche, J. GLWS 2.0: A global product that provides total water storage anomalies, groundwater, soil moisture and surface water with a spatial resolution of 0.5° from 2003 to 2019. PANGAEA, DOI: https://doi.org/10.1594/PANGAEA.954742 (2023).