Author Comments to Referee Comment 2

Manuscript number: EGUSPHERE-2025-5379

Title: High-resolution terrestrial water storage dynamics in Central Asia: Evaluating hydrological forcing datasets for GRACE downscaling

Dear Reviewer,

We are very grateful for your constructive comments and suggestions to our manuscript entitled "High-resolution terrestrial water storage dynamics in Central Asia: Evaluating hydrological forcing datasets for GRACE downscaling" [#EGUSPHERE-2025-5379]. These comments are very valuable and helpful for improving our manuscript. In the following, the comments are in black and numbered. Our responses are in blue, and quoted texts from the manuscript are in orange and italic.

General comment

I have read the article Liu et al, with an interest. Though, I find the idea and goal of this study interesting and relevant to HESS, I am not convinced by the choice of case study, its approach and the validation meths. There are a number of technical issues exist that need to be addresses.

Response: We sincerely thank you for the insightful comments and constructive suggestions, and we will try our best to address each of your concerns.

Comments

1. There are a number of studies on downscaling of TWSC through statistical machine learning, regression, data assimilation and Bayesian fusion. None of these approaches are mentioned and the benefit of the chosen approach is not demonstrated.

Response: Thank you for the comment. In the following, we expand the relevant discussion to include hydrological data assimilation, Bayesian fusion, statistical/regression-based downscaling, and machine learning-based downscaling methods.

The coarse spatial (~300 km) and temporal (~monthly) resolutions of GRACE limit its direct application in regional hydrological studies. To address this limitation, several methods have been developed. For example, hydrological data assimilation has been widely used to integrate GRACE observations into hydrological or land surface models (e.g. Tangdamrongsub et al., 2015; Li et al., 2019). Its primary purpose is to constrain or update model states using GRACE observations, often at the basin scale. Through model-based state updating, it can also help disaggregate total water storage into individual vertical hydrological components, such as soil moisture, groundwater, and surface water. However, recent work has pointed out that assimilating GRACE/GRACE-FO data into hydrological models may risk overfitting observations to model simulations, thereby distorting the original satellite observations and biasing results towards model outputs rather than measurements (Gerdener et al., 2023). In addition, both hydrological data assimilation and Bayesian fusion require explicit assumptions about error distributions, which are difficult to specify and independently validate.

An alternative and widely used method is downscaling, which has been developed to enhance the spatial and temporal resolution of GRACE data. These methods are generally classified as dynamic, statistical, or machine learning-based.

Dynamic downscaling employs physically based numerical models, e.g. regional climate or hydrological models, to generate high-resolution (HR) outputs while preserving underlying physical processes (Maraun et al., 2010). Statistical downscaling, often implemented through regression, derives fine-scale estimates from coarse-resolution data by establishing relationships between large-scale predictors and local-scale predictands (Maraun et al., 2010). The latter is data-driven, computationally efficient, and thus widely used in GRACE applications. We found most existing studies rely on a fixed set of forcing variables, such as P, ET, or R, derived from reanalysis or land surface models (e.g. Arshad et al., 2022; Pellet et al., 2024; Kalu et al., 2024). However, the influence of input hydrological forcing data selection on downscaling performance has rarely been assessed systematically. To the best of the authors' knowledge, Pellet et al. (2024) is the only recent study that developed a hybrid statistical dynamical approach to produce daily TWSC estimates at 1 km resolution by integrating GRACE data with auxiliary variables such as P, ET, R, and river network topography. This approach performs downscaling in both temporal and spatial domains, and therefore provides a suitable basis for evaluating how different input hydrological forcing data affect GRACE downscaling performance.

Machine learning methods have also shown considerable potential for GRACE downscaling. However, one of the main objectives of this study is to evaluate the influence of different hydrological input datasets on the downscaling results. If a machine learning framework were used, each combination of GRACE TWS and hydrological inputs would generally require a separate model training procedure. In our case, this would mean training four different models for the four input combinations considered. As a result, differences in hyperparameter tuning, model initialization, optimization could also affect the final outputs. Therefore, the differences among the downscaled products would not be attributable solely to the hydrological input datasets, but could also reflect variability introduced by the model training process itself. In addition, many machine learning models, especially more complex non-linear models, are less interpretable than more parsimonious approaches, making it more difficult to clearly identify how different hydrological inputs influence the downscaling results. For these reasons, exploring such methods is beyond the scope of this study.

We have revised the second paragraph in INTRODUCTION section to read:

"*To address this limitation, several methods have been developed. One example is hydrological data assimilation, which has been widely used to integrate GRACE observations into hydrological or land surface models (e.g. Tangdamrongsub et al., 2015; Li et al., 2019). However, recent studies have suggested that assimilating GRACE/GRACE-FO data into hydrological models may lead to an excessive adjustment of the observations towards model states, potentially weakening the original satellite signal and shifting the results closer to model simulations than to the measurements themselves (Gerdener et al., 2023). An alternative and widely used method is downscaling, which aims to enhance the spatial and temporal resolution of GRACE data. These methods are generally classified as dynamic, statistical, or machine learning-based. Dynamic downscaling employs physically based numerical models, e.g. regional climate or hydrological models, to generate high-resolution (HR) outputs while preserving underlying physical processes (Maraun et al., 2010). Statistical downscaling derives fine-scale estimates from coarse-resolution data by establishing relationships between large-scale predictors and local-scale predictands (Maraun et al., 2010). The latter is data-driven, computationally efficient, and thus widely used in GRACE applications. In recent years, machine learning techniques have also been increasingly applied in GRACE downscaling to capture complex, non-*

linear relationships between large-scale gravity signals and high-resolution hydroclimatic variables. Yet this approach faces challenges, including high computational demands and limited interpretability."

2. The article uses fluxes and related to the TWSC, it is shown in the past, that the estimation of derivatives introduces phase shift and smoothness, see e.g., https://doi.org/10.1002/2015JD023808. The situation is even worse when the monthly data is used. Ho is this effect controlled especially for computing daily time-steps?

Response: Thank you for your comment. We agree that the centred difference scheme introduces a smoothing effect. To address this, we have revised our method following Eicker et al. (2015). The same processing was also performed in Petch et al. (2023).

The terrestrial water balance expresses the change in total water storage as:

$$\frac{dS}{dt} = P - ET - R. \quad \text{(R1)}$$

For monthly data, the storage change $\frac{dS}{dt}$ for month i is calculated from the two preceding and following months using central weights $-\frac{1}{8}, -\frac{1}{4}, \frac{1}{4}, \frac{1}{8}$. Because this differentiation filter has a smoothing effect, we also apply a central phase-preserving filter to the right-hand side of Eq. (R1), with weights $\frac{1}{22}, \frac{1}{4}, \frac{9}{22}, \frac{1}{4}, \frac{1}{22}$. These coefficients ensure that the amplitude of a purely sinusoidal annual signal is damped in exactly the same way as it being applied to the left-hand side of Eq. (R1).

This revision will be documented appropriately in Sect. 4.1. Our revised results show that this modification does not change the main conclusions of the manuscript. The corresponding results presented in the RESULTS section will be updated accordingly.

3. Comparison with daily ITSG is rather strange, especially when the goal is high resolution downscaling is. In theory, daily ITSG represents the dominant gravity signal contents at daily time-step but up to degree and order 40. The presented approach by authors likely does not contain the gravity information at daily time-scale because all information comes from daily fluxes. I cannot follow the justification of mismatches between spatial resolutions and differences in the signal content.

Response: Thanks for your comment. We consider the comparison between the temporally downscaled TWSC in our study and ITSG-Grace2018 daily solution to be appropriate because both products represent the same underlying terrestrial water storage signal. ITSG-Grace2018 provides daily time-variable gravity fields, which can be converted into gridded total water storage anomalies. Our temporally downscaled TWSC combines daily water balance fluxes with the monthly JPL mascon product, which it itself derived from GRACE obeservations and provided in the form of gridded TWS anomalies. Therefore, our downscaled TWSC is not determined by the daily fluxes alone, but also contains information from GRACE-observed gravity field. In this sense, both datasets represent terrestrial water storage variations and are therefore comparable.

The comparison of our downscaled TWSC and ITSG-Grace2018 is conducted at the basin scale. First, TWSC is calculated from ITSG-Grace2018. Then, basin-mean time series are derived for both datasets and evaluated using the correlation coefficients and root mean square error, as shown in Fig. 6 of the main manuscript.

4. Since the method is general and data are global, I suggest to use the approach for a data rich area where the signal can be compared with independent data sets such as dense groundwater and GNSS network in parts of US or Europe. I believe there are already processes data in previous studies, at the range of a few km resolution, which could be used for a comparison.

Response: Thank you for this valuable suggestion. We agree that groundwater and GNSS observations can provide useful reference information for evaluating downscaled terrestrial water storage changes. However, it should be noted that both groundwater and GNSS mainly allow indirect evaluation, because TWS represents the integrated signal of multiple storage components, including surface water, groundwater, soil moisture, snow, and glacier storages (Rodell and Famiglietti, 2001). In this sense, soil moisture data can also provide a valuable form of indirect evaluation, as they reflect an important component of the total storage signal. Therefore, we have added one additional comparison using soil moisture data, together with a comparison against two existing global downscaled TWS products. Please see Comment 5 for details of these two comparison results.

This study combines methodology with regional application. One of its objectives is to develop validation strategies for data-scarce regions, where dense groundwater and GNSS stations are generally unavailable. Another motivation is that FLDAS Central Asia is specifically designed for Central Asia but has not yet been extensively investigated in hydrological applications. The selected study area therefore provides a valuable opportunity to assess the performance of this dataset. In addition, this basin is hydrologically important because of its complex topography and strong upstream – downstream linkages.

At the same time, the region is of clear societal relevance. It has a pronounced transboundary character, spanning parts of Kyrgyzstan and Uzbekistan, and faces substantial water-allocation and coordination challenges. Improving water information here is therefore important for irrigation management, agricultural productivity, and broader water security. Ongoing international programmes (https://weact-project.eu/, https://documents.worldbank.org/en/publication/documents-reports/documentdetail/466081496780097071) in the Naryn – Kara Darya catchments and the Fergana Valley further indicate the need for better water allocation, improved irrigation and drainage service delivery, and stronger adaptation to growing climate pressure in this transboundary setting.

5. The treatment of biases between monthly GRACE data and daily fluxes is basically nothing rather than distributing these biases in days based on the water balance increments. This does not guarantee any physical meaningfulness, unless the authors argue differently. Therefore, the suggestion of independent validation in 4 seems necessary to me.

Response: Thank you for this important comment. We agree that the temporal correction step does not, by itself, guarantee full physical realism at the daily scale. Our intention was not to claim that the daily TWSC fields are independently resolved by GRACE, but rather that they are reconstructed by combining (i) sub-monthly variability from daily water balance fluxes and (ii) monthly large-scale storage constraints from GRACE. In this sense, the method is interpreted as a GRACE-constrained and hydrologically informed reconstruction.

Furthermore, we have strengthened the validation section by explicitly recognizing the necessity of independent evaluation. In addition to the comparison with the ITSG-Grace2018 daily solution, the upscaling-back consistency test, and the event-based analysis, we have also added an evaluation based on soil moisture and a comparison with two existing global downscaled TWS products. The comparison results are shown below. We sincerely thank you again for your valuable and insightful feedback that helps to improve the quality of our manuscript.

Evaluation using soil moisture data and existing global GRACE downscaled products (the following text is the same with our response to Comment 10 of Referee Comment 1)

We have conducted an indirect evaluation using soil moisture data. Specifically, we use the ESACCI-Zheng product (Zheng et al., 2023), which provides global gap-free surface soil moisture at daily 1 km resolution for the period 2000 – 2020. It enables evaluation at comparable spatio-temporal scales. The dataset can be downloaded at https://data.tpdc.ac.cn/zh-hans/data/30131436-88d1-4be3-8e3d-14905a29d6d6. Daily soil moisture data is stored in MOIDIS Sinusoidal Tile Grid. For our study region, four tiles h23v04, h23v05, h24v04, h24v05 for each year are downloaded.

Due to the temporal coverage of soil moisture (SM) data, the analysis is restricted to April 2002 – December 2020. TWSC is converted to TWS (here simply referred to as TWS although it represents storage relative to TWS at epoch 0) because TWS is comparable with SM. Prior to the comparison, the linear trend and long-term mean are removed from the time series of SM and TWS. The resulting detrended series are hereafter referred to as SM and TWS anomalies. It should be noted that a direct comparison of absolute values between SM and TWS is not feasible due to differences in integration depths and physical units. Therefore, the Pearson's correlation coefficient is used to analyze their relationship. And time lags between TWS and SM anomalies are computed using a cross-correlation analysis.

Figure R1 presents the basin-averaged time series of soil moisture and TWS anomalies from the downscaled products in the four hydrological forcing scenarios (GLDAS, FLDAS-CA, ERA5-Land, and Mix). TWS exhibits comparatively smooth behavior at short timescales and a dominant seasonal signal, whereas the SM time series shows substantially higher variability at short timescales.

A general correspondence is observed between TWS and SM anomalies in terms of their seasonal dynamics, with a strong rise during spring followed by a quick decline, indicating strong climatic control on terrestrial water variability. In addition to the primary seasonal peak, a minor secondary maximum is observed in the SM time series during October and November. Correspondingly, a more rapid increase in TWS can partially be observed during this period. This can be attributed to reduced evapotranspiration after the summer period, which allows water to accumulate. The correlation coefficients between TWS and SM are 0.50, 0.45, 0.46 and 0.48 for the four downscaled products (GLDAS, FLDAS-CA, ERA5-land and Mix), respectively. Their relative performance is consistent with that already analyzed in the main manuscript. The GLDAS performs well in the basin-averaged values. The Mix downscaled product's overall performance is stable according to evaluation shown in the main manuscript and this additional evaluation with SM data.

The time shift is identified by computing the cross-correlation between two time series, resulting in a lag of approximately one month (31 – 34 days for the four downscaled results). The seasonal maxima and minima of the TWS time series occur about one month later than those of SM. This is reasonable and can be attributed to the slower and delayed response of deeper water storage components observed by GRACE, whereby transitions between dry and wet conditions evolve more slowly than in near surface soil layers.

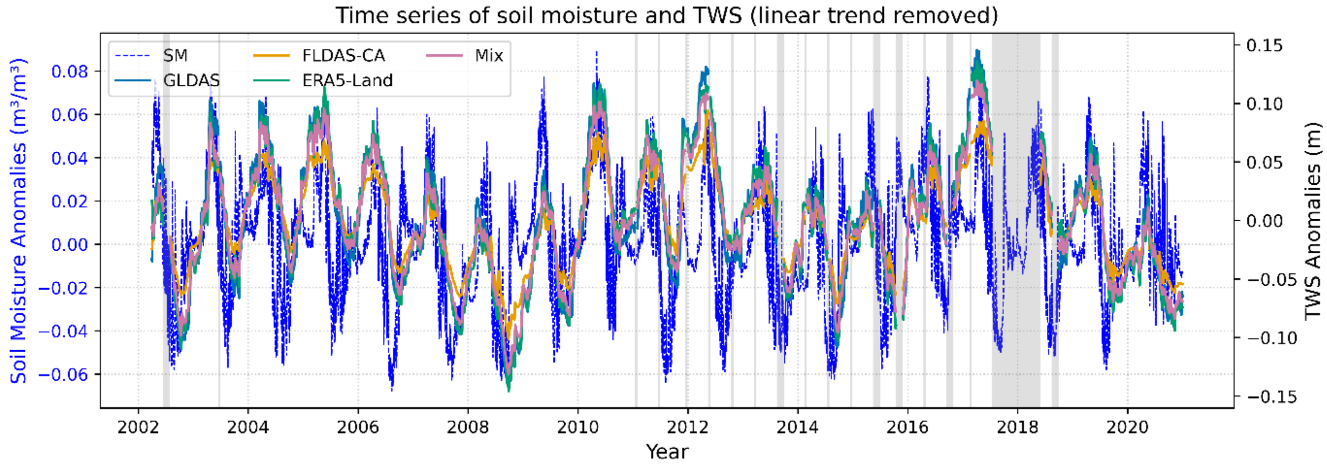

Figure R1: Time series of soil moisture (left axis) and TWS anomalies derived from the downscaled products in the four hydrological forcing scenarios (GLDAS, FLDAS-CA, ERA5-Land, and Mix) (right axis), averaged over the study region for the period 2002 – 2020. Grey-shaded areas indicate periods when GRACE data are unavailable.

Figure R2 displays the correlation coefficients between daily time series of TWS from downscaled products and SM data for each grid cell along with the corresponding time shifts derived from cross-correlation analysis. All downscaled products exhibit predominantly positive correlations with SM across the study region. Negative correlations are mainly found in areas containing reservoirs (Toktogul Reservoir), lakes (Song Kul Lake), and southern glacier regions, which are known to be challenging for reliable SM observations.

The spatial patterns of time shifts indicate that TWS variations generally lag behind SM, as reflected by the dominance of positive time shifts (green shading). This lag is particularly pronounced in the central and southern regions, where delays of approximately 30 – 90 days are common. Such behavior reflects the slower and integrated response of terrestrial water storage relative to surface soil moisture, likely mediated by subsurface processes and soil water redistribution. It should be noted that time shifts reaching ±120 days do not necessarily indicate an exact lag of 120 days, but rather reflect cases where the maximum allowable lag is exceeded. This suggests weak correlations in these regions, which generally correspond to areas with low or negative correlation coefficients in the correlation map (left column).

It can be observed that the correlation between two variables in highly irrigated area is relatively lower than that in the surrounding area, and the lag of TWS relative to SM is also longer. This suggests that additional water redistribution processes may be present in highly irrigated area, e.g., potential return flows from surface water irrigation.

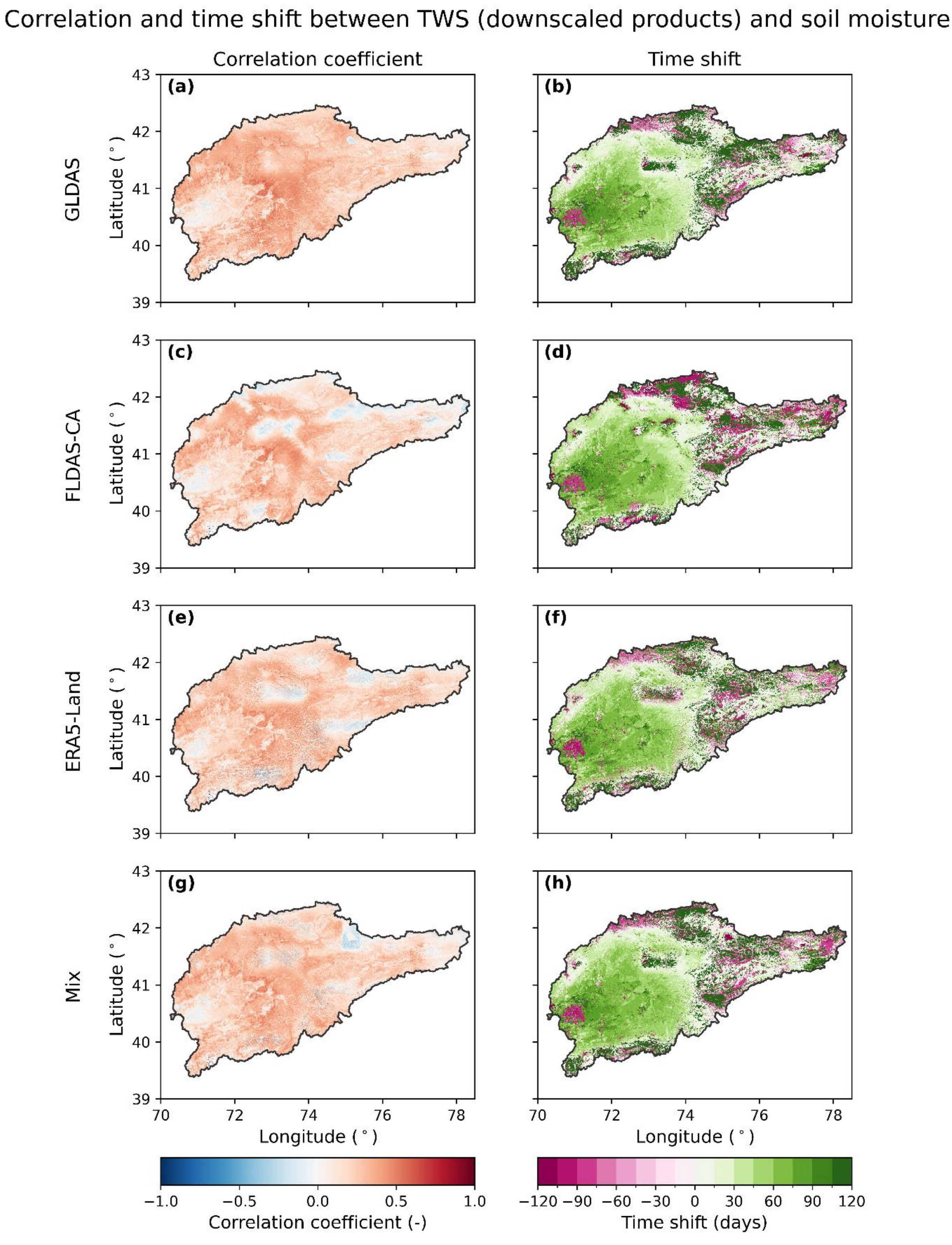

Figure R2: Spatial distribution of correlation coefficients and time shifts between TWS anomalies derived from the downscaled products in the four hydrological forcing scenarios (GLDAS, FLDAS-CA, ERA5-Land, and Mix) and SM anomalies at each grid point. Positive time shifts mean that TWS is delayed in comparison to SM.

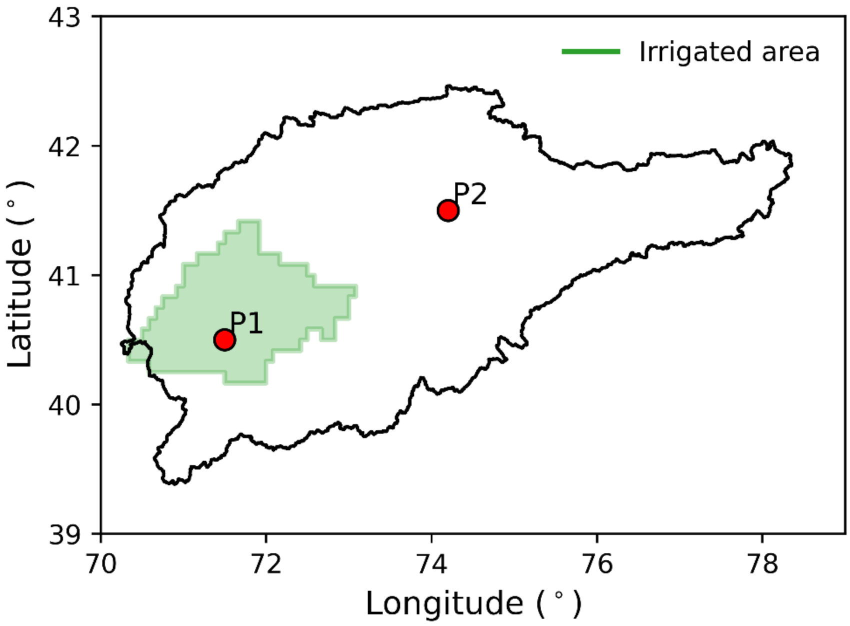

Figure R3: The study region and highly irrigated area. P1 and P2 are chosen to show SM and TWS time series.

We choose two grid cells, P1 and P2 as shown in Fig. R3, located in highly irrigated and non-irrigated area, respectively. P2 is located at an elevation of approximately 2000 m, on moderately sloping terrain that is typical in non-irrigated area. It is situated far from glacier region and is therefore not directly influenced by glacier processes. Figures R4a and R4c show the daily TWS anomalies in these grid cells in comparison with SM. Seasonal patterns are clearly evident at both locations. The seasonal cycle at P2 is comparatively more moderate than that observed at P1. For both sites, SM exhibits substantially higher variability than TWS products, with pronounced short-term fluctuations superimposed on a clear seasonal cycle. In contrast, all TWS datasets (GLDAS, ERA5-Land, FLDAS-CA, and Mix) display smoother temporal variations, reflecting their integrated nature. Despite these differences in variability, the temporal evolution of TWS anomalies is broadly consistent with that of SM. However, a temporal lag between TWS and SM is observed at both sites. TWS variations generally lag behind SM. This lag is more pronounced at P1 (70 – 80 days for the four downscaled products) than at P2 (20 – 30 days for the four downscaled products), indicating potential spatial differences in hydrological processes.

Figures R4b and R4d illustrate the corresponding daily climatology. For both sites, TWS reaches its maximum during late spring to early summer (approximately April – June), followed by a gradual decline towards autumn and a recovery during winter. This seasonal variation is consistent across all downscaled products, although ERA5-Land generally exhibit slightly higher amplitudes compared to others. In contrast, SM shows a different seasonal phase relative to TWS. At P1, SM peaks earlier in the year (late winter to early spring) and declines during summer, reaching minimum values around August – September. At P2, SM reaches its maximum around April, followed by a decline through summer, with minimum values also occurring in August – September. Subsequently, SM exhibits a secondary peak around November, after which it decreases slightly and remains relatively stable from December to March. We recognize during winter, microwave-based soil moisture retrievals are known to be unreliable due to frozen soil and snow cover, often resulting in missing or underestimated values. As temperatures rise in early spring, the rapid thawing of the soil and snowpack leads to a sudden increase in liquid water availability, which is captured by the satellite signal as a sharp rise in soil moisture. However, SM at P1 remains consistently high throughout winter

(December to March). This high level indicates a prolonged wet state during the cold season, suggesting stronger water accumulation prior to winter, potentially associated with moisture accumulation following the summer irrigation.

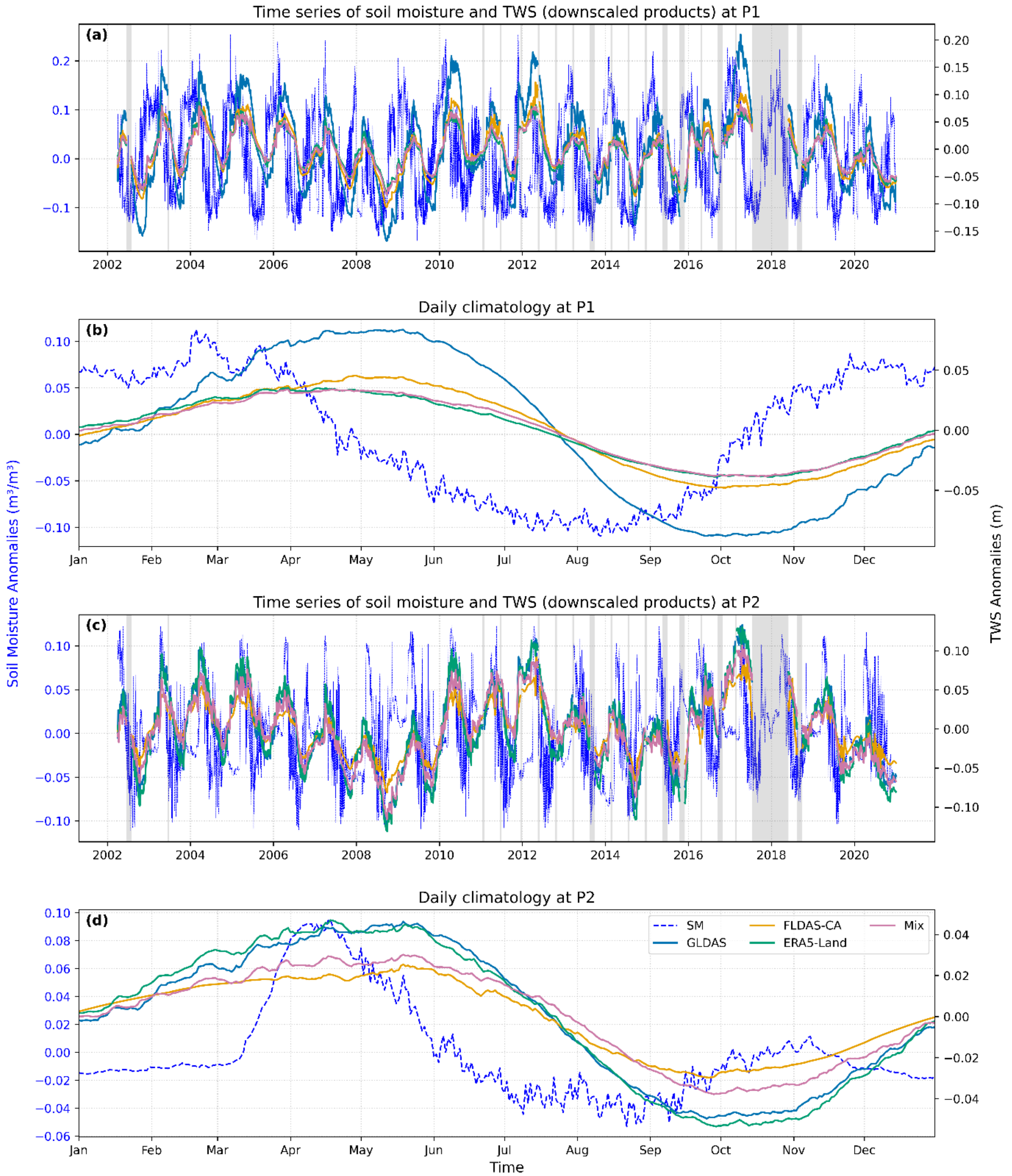

Figure R4: Time series of SM and TWS anomalies from four downscaled products (GLDAS, FLDAS-CA, ERA5-Land, and Mix) at P1 and P2. The linear trend and long-term mean have been removed from all time series. (a) SM (left axis) and TWS anomalies (right axis) from downscaled products at P1. (b) Daily climatology of SM (left axis) and TWS anomalies (right axis) at P1. (c) SM (left axis) and TWS anomalies (right axis) from downscaled products at P2. (d) Daily climatology of SM (left axis) and TWS anomalies (right axis) at P2.

To isolate sub-monthly variations, following Blank et al. (2023), a third-order Butterworth high-pass filter with a cutoff frequency of 30 d is applied in the forward and backward directions. This filter preserves the phase while removing signals with periods longer than 30 days, which dominate the original time series. Figure R5 illustrates the high-pass filtered (sub-monthly) variations of SM and TWS from the four downscaled products (GLDAS, FLDAS-CA, ERA5-Land, and Mix) at P1 and P2 between May 2010 and October 2011, highlighting short-term hydrological fluctuations after removing low-frequency components.

Compared with the total signals shown in Figs. R4a and R4c, the high-frequency component of SM accounts for a substantially larger proportion of the total signal than that of TWS at both P1 and P2. SM exhibits larger and more frequent sub-monthly fluctuations than TWS. This reflects its rapid response to precipitation and evapotranspiration at the land surface, whereas TWS integrates water storage changes across soil, groundwater, and surface water components, resulting in attenuation of high-frequency signals. Moreover, the intrinsic characteristics of GRACE observations, including their coarse spatial resolution and inherent temporal smoothing, further reduce the sensitivity of TWS to sub-monthly variability. The correlations between the high-pass filtered TWS and SM time series are considerably lower than those of the unfiltered time series. This is expected, as the seasonal storage variations that largely drive high correlations for the unfiltered time series are not present anymore in the filtered ones. Nevertheless, the correlation between SM and TWS at sub-monthly timescales remains moderate (approximately $r \approx 0.3$), with the Mix forcing achieving the highest values around 0.35 at both sites. This suggests that downscaled TWS is still able to capture part of the day-to-day variability in SM.

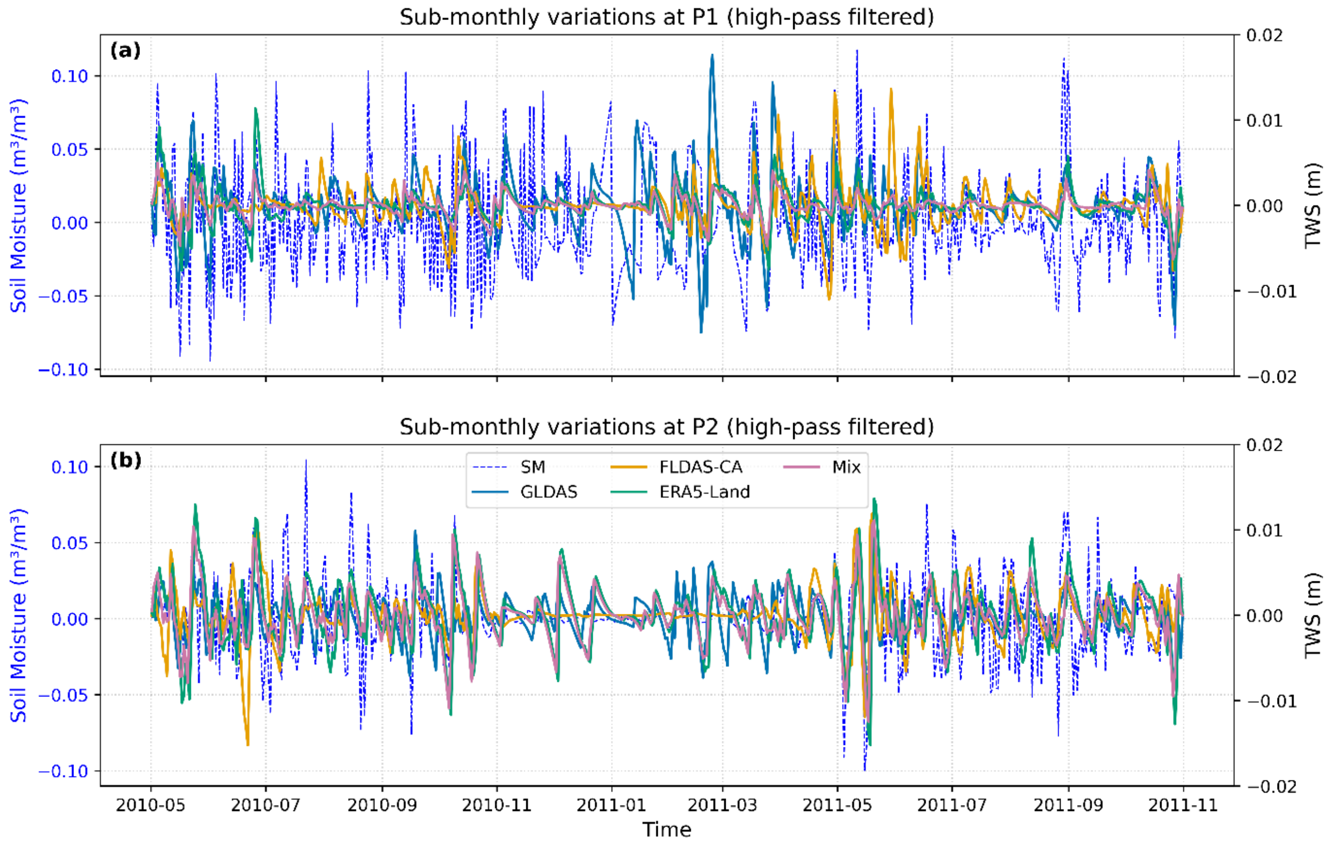

Figure R5: High-pass filtered time series of SM and TWS anomalies from four downscaled products (GLDAS, FLDAS-CA, ERA5-Land, and Mix) at P1 and P2.

Next, we computed the correlation coefficients among daily soil moisture changes (ΔSM), daily TWSC from the Mix downscaled product, and daily P in highly irrigated and non-irrigated regions. This is also a supplementary analysis to Fig. 10 in the main manuscript. ΔSM is used instead of SM to ensure comparability with P, which is defined as a daily flux. As shown in Fig. R6, in both regions, TWSC exhibits a strong positive correlation with P (*r* = 0.80/0.81), indicating that precipitation is one of the dominant drivers of terrestrial water storage variability in the downscaling. This strong relationship is consistent across irrigated and non-irrigated areas, suggesting that large-scale water storage dynamics are primarily controlled by hydrological forcing inputs.

In contrast, the correlations between ΔSM and both TWSC and P are notably weaker. In highly irrigated areas (Fig. R6a), ΔSM shows low correlations with TWSC (*r* = 0.16) and P (*r* = 0.15). This weak relationship likely reflects the influence of irrigation, which introduces additional water inputs independent of precipitation. In non-irrigated areas (Fig. R6b), the correlations between ΔSM and both TWSC and P are slightly higher (*r* ≈ 0.22). The absence of irrigation allows soil moisture to respond more directly to precipitation and to co-vary more consistently with terrestrial water storage.

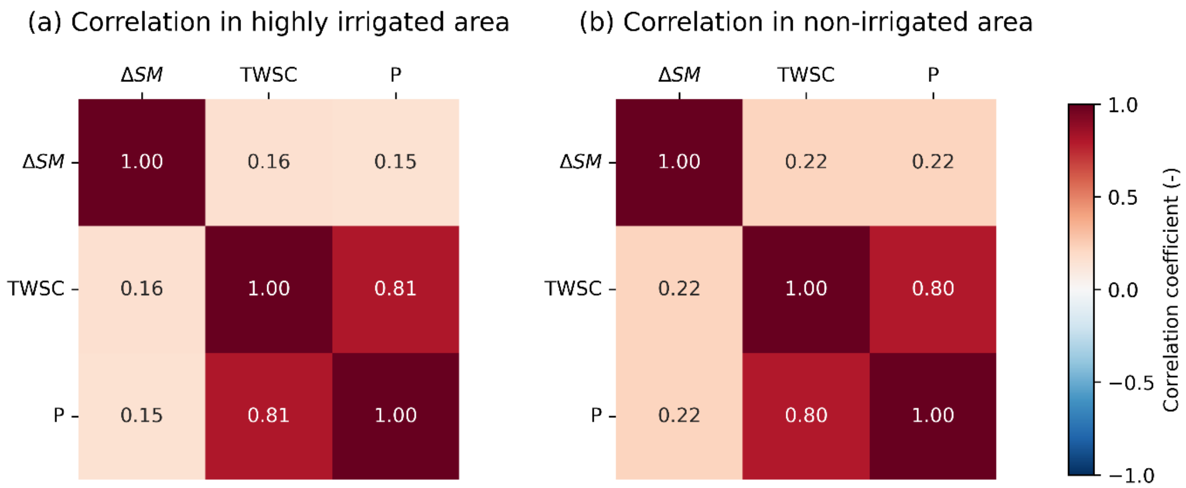

Figure R6: Correlation coefficients among region-averaged ΔSM, TWSC, and P from Mix downscaled product in highly irrigated and non-irrigated areas.

In addition to the evaluation using soil moisture data, we include two existing GRACE downscaling datasets (Gou and Soja, 2024; Li and Kusche, 2026) for comparison. To the best of our knowledge, these are currently the only two open-access global downscaled total water storage datasets available. The correlation coefficients and root mean square error (RMSE) between the downscaled products and the JPL mascon are computed at each mascon grid point, as shown in Fig. R7. This is also a supplementary analysis to Fig. 8 in the main manuscript.

All downscaled results in our study (Figs. R7a, c, e, g) exhibit consistently high correlations with the JPL mascon, with values approaching or exceeding 0.8 across most of the study region. The mean correlation coefficients are 0.88, 0.79, 0.82, and 0.85 for the GLDAS, FLDAS-CA, ERA5-Land, and Mix downscaled results, respectively. In comparison, the result of Gou and Soja (2024) in Fig. R7i displays lower and more spatially heterogeneous correlations, particularly in the central and northern regions, with a mean value of 0.76. The results of Li and Kusche (2026) in Fig. R7k shows uniforms high correlations across the domain, with a mean value of 0.98, exceeding those of other downscaled products.

The RMSE patterns (Figs. R7b, d, f, h) further differentiate the performance of these datasets. The downscaled results in our study generally exhibit low to moderate RMSE values, with relatively higher errors concentrated in the southern region. The basin-averaged RMSE values are 1.90, 2.44, 2.04, and 1.94 cm for the GLDAS, FLDAS-CA, ERA5-Land, and Mix downscaled results, respectively. In contrast, the result of Gou and Soja (2024) in Fig. R7j shows substantially higher RMSE values across most of the region, with a mean value of 8.42 cm. The result of Li and Kusche (2026) in Fig. R7l exhibits pronounced spatial variability in RMSE (mean value of 4.52 cm), with particularly high errors in the central areas, despite its high correlation values. This suggests that while this product captures temporal variability well, it may suffer from amplitude mismatches.

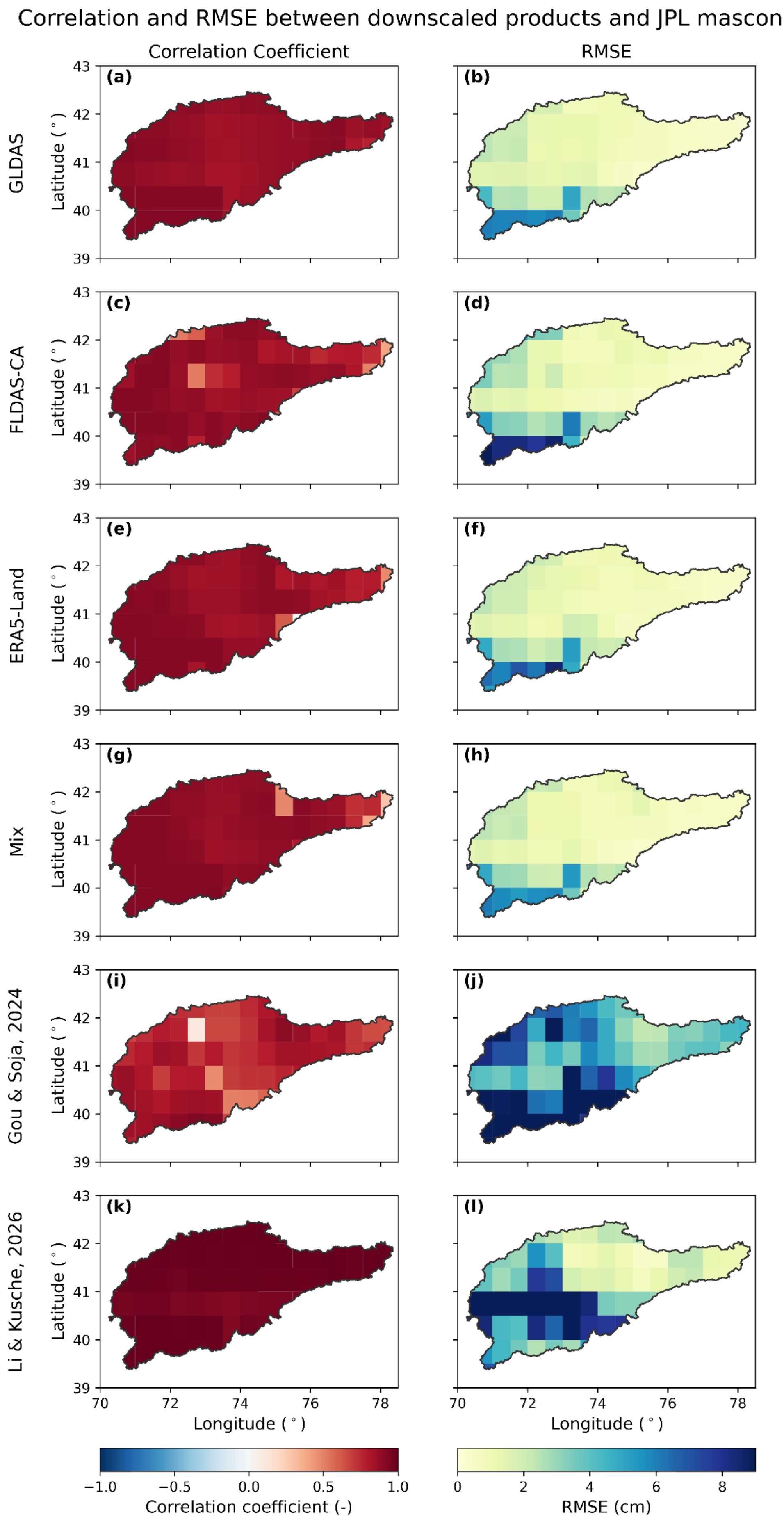

Figure R7: Correlation coefficients and RMSEs between monthly TWSCs from the downscaled products obtained using different hydrological forcing datasets (GLDAS, FLDAS-CA, ERA5-Land, and Mix) and the JPL mascon product. Results from Gou and Soja (2024, https://doi.org/10.1038/s44221-024-00194-w) and Li and Kusche (2026, https://doi.org/10.1029/2025GL119881), two global downscaled total water storage anomaly products are also included.

References

Arshad, A., Mirchi, A., Samimi, M., and Ahmad, B.: Combining downscaled-GRACE data with SWAT to improve the estimation of groundwater storage and depletion variations in the Irrigated Indus Basin (IIB), Sci. Total Environ., 838, 156044, https://doi.org/10.1016/j.scitotenv.2022.156044, 2022.

Blank, D., Eicker, A., Jensen, L., and Güntner, A.: A global analysis of water storage variations from remotely sensed soil moisture and daily satellite gravimetry, Hydrol. Earth Syst. Sci., 27, 2413–2435, https://doi.org/10.5194/hess-27-2413-2023, 2023.

Eicker, A., Forootan, E., Springer, A., Longuevergne, L., and Kusche, J.: Does GRACE see the terrestrial water cycle "intensifying"?, J. Geophys. Res. Atmos., 121(2), 733-745, https://doi.org/10.1002/2015JD023808, 2015.

Gerdener, H., Kusche, J., Schulze, K., Döll, P., and 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, Journal of Geodesy, 97(7), 73, https://doi.org/10.1007/s00190-023-01763-9, 2023.

Gou, J., and Soja, B.: Global high-resolution total water storage anomalies from self-supervised data assimilation using deep learning algorithms, Nature Water, 2(2), 139-150, https://doi.org/10.1038/s44221-024-00194-w, 2024.

Kalu, I., Ndehedehe, C. E., Ferreira, V. G., Janardhanan, S., Currell, M., and Kennard, M. J.: Statistical downscaling of GRACE terrestrial water storage changes based on the Australian Water Outlook model, Sci. Rep., 14(1), 10113, https://doi.org/10.1038/s41598-024-60366-2, 2024.

Li, B., Rodell, M., Kumar, S., Beaudoing, H. K., Getirana, A., Zaitchik, B. F., de Goncalves, L. G., Cossetin, C., Bhanja, S., Mukherjee, A., Tian, S., Tangdamrongsub, N., Long, D., Nanteza, J., Lee, J., Policelli, F., Goni, I. B., Daira, D., Bila, M., de Lannoy, G., Mocko, D., Steele-Dunne, S. C., Save, H., and Bettadpur, S.: Global GRACE data assimilation for groundwater and drought monitoring: Advances and challenges, Water Resour. Res., 55(9), 7564-7586, https://doi.org/10.1029/2018WR024618, 2019.

Li, F., and Kusche, J.: Reproducing GRACE total water storage change at finer spatial scales, Geophys. Res. Lett., 53(1), e2025GL119881, https://doi.org/10.1029/2025GL119881, 2026.

Maraun, D., Wetterhall, F., Ireson, A. M., Chandler, R. E., Kendon, E. J., Widmann, M., Brienen, S., Rust, H. W., Sauter, T., Themeßl, M., Venema, V. K. C., Chun, K. P., Goodess, C. M., Jones, R. G., Onof, C., Vrac, M., and Thiele-Eich, I.: Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user, Rev. Geophys., 48(3), https://doi.org/10.1029/2009RG000314, 2010.

Pellet, V., Aires, F., Alfieri, L., and Bruno, G.: A physical/statistical data-fusion for the dynamical downscaling of GRACE data at daily and 1 km resolution, J. Hydrol., 628, 130565, https://doi.org/10.1016/j.jhydrol.2023.130565, 2024.

Petch, S., Dong, B., Quaife, T., King, R. P., and Haines, K.: Water and energy budgets over hydrological basins on short and long timescales, Hydrol. Earth Syst. Sci., 27, 1723–1744, https://doi.org/10.5194/hess-27-1723-2023, 2023.

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

Tangdamrongsub, N., Steele-Dunne, S. C., Gunter, B. C., Ditmar, P. G., and Weerts, A. H.: Data assimilation of GRACE terrestrial water storage estimates into a regional hydrological model of the Rhine River basin, Hydrol. Earth Syst. Sci., 19, 2079–2100, https://doi.org/10.5194/hess-19-2079-2015, 2015.

Zheng, C., Jia, L., and Zhao, T.: A 21-year dataset (2000–2020) of gap-free global daily surface soil moisture at 1-km grid resolution, Scientific Data, 10(1), 139, https://doi.org/10.1038/s41597-023-01991-w, 2023.