Reviewer 1

The authors did a fantastic job responding to my comments. I really enjoyed reading this paper (acknowledging I am biased because of my research interests). Please consider a few minor comments below before publication. Well done.

We thank the reviewer for his positive feedback on our manuscript and thank him for the constructive comments to improve our manuscript.

1) **I think it can be stated earlier that water limitation can reduce evaluation of source of soil water for roots. Energy-limitation will result in low correlations between any soil column water availability and NIRv. However, there might be differences in use of TWS and soil moisture for the plant water use though it can't be detected because of the lack of water limitation. For example, this can be mentioned earlier in line 170-172 about why some regions have negative correlations. This could partially be due to energy limitation.**

In line 181 of the revised manuscript, we added:

Note that predominant energy limitation of the vegetation prevents the evaluation of the relevance of soil moisture vs. terrestrial water storage as partial correlations will become insignificant when temperature or radiation are mainly controlling vegetation functioning.

2) **I think it is promising that use of SIF at 0.5 degrees shows similar results as shown in the SI. However, I do think it is worth at least mentioning some redundance and double counting of relationships that will be show in the boxes in Figs. 2 and 3. Specifically, GRACE is gridded at 0.5 degrees and really at 2 degrees. Matching 0.05 resolution NIRv from many pixels to the same GRACE anomalies can present large spatial autocorrelation to the results. I suggest mentioning this issue when mentioning the SIF results in the robustness section.**

In last paragraph of the section 3.4, we added the following paragraphs:

When analyzing partial correlations between Total Water Storage (TWS) and vegetation metrics (NIRv or SIF) at finer resolutions (0.05 degrees for NIRv or 0.5 degrees for SIF), it is crucial to acknowledge the potential emergence of significant spatial autocorrelation. This is attributed to the fact that the actual spatial resolution of the satellite signal underlying the TWS data is 2-3 degrees.

3) **I am satisfied with the bootstrapping on Fig. 2c. This does show a level of significant differences through aggregation. Additionally, consider moving Fig. S9 to the main text somewhere (I find this to be great). Or consider showing stippling on Fig. 1C which would require bootstrapping tests within individual pixels to see whether the correlation distribution for TWS-NIRv falls outside the bounds of the correlation distribution for soil moisture-NIRv. I leave this up to the authors.**

Following these suggestions, we utilized the bootstrapping technique to infer the significance of the results in the map in Figure 1c. Accordingly, a supplementary figure S4 has been included, illustrating that, for the majority of grid cells, the correlation distribution of NIRv ~ TWS falls outside the bounds of the correlation distribution for NIRv ~ SSM, indicating that the correlation difference displayed in Figure 1c is significant in many regions. Subsequently, Figure 1 has been revised to show the bootstrapping mean partial correlation which is supposed to be more robust than the actual partial correlations displayed before.

To convey this information, we incorporated the following details into line 167 in the Methodology section (2.2.2).

To derive partial correlation estimates between NIRv and the water storages, we employed a bootstrapping approach (resampling with replacement from the original data) within each grid cell, with 1000 repetitions to compute bootstrap means and confidence intervals.

In the caption of Figure 1, we included the following line:

The color bar denotes the mean partial correlation for each grid cells, computed from the partial correlations across individual bootstrapping samples.

Further in the line 269, we added:

To ensure that the observed patterns of difference of partial correlation between SSM and TWS are not the artifacts arising from the computation of differences based on mean partial correlation, we compared the 95% confidence intervals obtained through bootstrapping. Our results indicate that, for the majority of the considered grid cells, the entire confidence intervals of the correlation (NIRv ~ TWS) fall outside the bounds of the correlation (NIRv ~ SSM) which indicates that the correlations differences are significant, thus enhancing the robustness and confidence in our findings (Figure S4).

Please refer to the line number from the revised manuscript.

Line 64: I suggest tempering slight with "providing valuable insights into at least some of the root zone soil moisture"

The corresponding line is changed accordingly.

Line 79: Optional, but could add this reference as well:

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, doi:10.1029/2000WR900306, 2001.

We added the reference to the line 87.

While soil moisture fluctuations represent the largest variation of TWS (Rodell and Famiglietti, 2001), it is essential to note that certain regions exhibit notable short term fluctuations in lake and groundwater due to human management (Strassberg et al., 2007; Cooley et al., 2021).

Line 302: "not always the case" (sometimes true)

We change the line (line 326 in the revised manuscript):

This, however, is probably not the case and simply a reflection of reduced variability in surface soil moisture**.**

to

This, however, is probably not the case and might simply be a reflection of reduced variability in surface soil moisture.

Line 344-359: References to figure 4 panels here appear incorrect

Corrected within the paragraph.

Reviewer 2

The authors have done a good job responding to the reviewer comments, including adding substantial additional analyses, appropriate limitations statements, and stronger links to the existing literature. The additional analyses reinforce the previous findings and provide new insights to the data. I have only a few minor remaining comments, which I have added to my initial comments and the authors' initial response below. After these have been addressed, I believe the manuscript is suitable for publication and will make an excellent contribution to the literature.

We appreciate the positive feedback and constructive comments from the reviewer on our manuscript. Thank you for your valuable input in helping us improve our work.

Initial reviewer comment:

Lines 77-79: Assumptions 2 and 3 should be further justified and referenced. Assumption 2 is particularly concerning in grid cells with highly manipulated hydrologic systems, such as where irrigation results in substantial drawdown of groundwater and reservoir levels throughout the growing season.

Author response:

We will update this in the text in the same line with justifications.

Follow up reviewer comment:

Perhaps I am missing it, but I do not see any text changes in response to this comment. Could you please 1) add a citation to support the assumption that canopy water storage is much smaller than soil water storage (across all biomes), and 2) add further justification, a citation, and/or a caveat pertaining to the assumption that water storage in lakes (including reservoirs) and groundwater is negligible? Short-term fluctuations in reservoir and volumes can be very large in some areas due to human management of water, as was recently comprehensively quantified by Cooley et al. Similarly, seasonal and sub-seasonal changes in groundwater storage can be very large in areas with intensive groundwater irrigation; for example, the papers by Strassberg et al. and Breña-Naranjo et al. listed below report short-term changes in groundwater storage in the High Plains Aquifer. I understand if this assumption is not avoidable, but it should acknowledged as a limitation and appropriately referenced.

Cooley, S. W., Ryan, J. C., & Smith, L. C. (2021). Human alteration of global surface water storage variability. Nature, 591(7848), 78-81.

Breña-Naranjo, J. A., Kendall, A. D., & Hyndman, D. W. (2014). Improved methods for satellite-based groundwater storage estimates: A decade of monitoring the high plains aquifer from space and ground observations. Geophysical Research Letters, 41(17), 6167-6173.

Strassberg, G., Scanlon, B. R., & Rodell, M. (2007). Comparison of seasonal terrestrial water storage variations from GRACE with groundwater-level measurements from the High Plains Aquifer (USA). Geophysical Research Letters, 34(14).

We apologize that the previous comment of the reviewer slipped through. We agree with the reviewer and have added references (Cheng & Jia 2019, Stocker et al. 2023)to assumption 3 on negligible canopy water storage, and have emphasized the limitations associated with assumption 2, as suggested by the reviewer in line 85:

While soil moisture fluctuations represent the largest variation in Total Water Storage (TWS) (Rodell and Famiglietti, 2001), it is crucial to note that certain regions exhibit significant short-term fluctuations in lake and groundwater due to human management (Strassberg et al., 2007; Cooley et al., 2021).

Initial reviewer comment:

Figure 1c legend: Please specify what the white areas represent. Also, the color scale is variously described as blue and purple for positive correlations and red and orange for negative correlations in this legend and the legend for Figure S2; this should be standardized.

Author response:

We will specify in the caption that the white areas represent regions with no or insufficient number of data. Apart from this, the references to the figure's colors will be standardized.

Follow-up reviewer comment:

The caption is clearer, but the description of the legend colors is still inconsistent between Figs 1 and S2.

We thank the reviewer for catching this.

To make it more consistent,

We replaced "*The purple colour indicates that the correlation is positive while the red colour indicates a negative correlation between SSM and TWS*" in the caption of figure S2 with

"*The purple colour indicates the positive correlation of SSM with TWS while orange colour indicates the opposite."*

Initial reviewer comment:

Line 238 and Figure S7: This figure is referenced in text after Figure S2; the SI figures should be reordered to be sequential.

Author response:

We will take it into consideration and update our manuscript accordingly.

Follow-up reviewer comment:

Figures S7 and S8 still appear to be out of order.

We have changed the order of Figures in Supplementary section, matching it with the order in which it appears in the text of the manuscript.

Additional reviewer comment:

Finally, a number of typos remain (mostly related to spacing and punctuation). I did not record them all, but examples are present in lines 119, 127, 199, 325, 369, 393, 404 and 405, etc.

We tried to removed most of the typos error in the updated manuscript.