Dear Editor and Reviewers,

We would like to thank the Editor and three Reviewers for their efforts in handling the manuscript and for their valuable comments to improve the manuscript. We have revised our manuscript thoroughly according to the comments and provide a point-by-point response to the comments from three Reviewers below. The original comments from three Reviewers are in black font, and our responses are in blue font.

On behalf of all co-authors,
En Liu

Response to Reviewer #1's comments on the manuscript egusphere-2023-1597

RC1: 'Comment on egusphere-2023-1597', Anonymous Referee #1, 02 Feb 2024

While the information contained in the manuscript is quantitative, I ended up questioning the potential contribution of this paper to the readers. I have tried to address why I think While the revised manuscript has been improved in the sense that the objective can be better interpreted, I found the whole manuscript is still relatively complex to understand mainly due to 1) many awkward/incomplete sentences, 2) figures/tables are not well organized and indicated (e.g., labels, units, captions). I am afraid the readers might get confused due to missing information in the figures. Please check my specific comments below:

Response: The authors thank the Reviewer #1 for her/his constructive and insightful comments that help us improve the quality of the manuscript. The original comments from Reviewer #1 are in black font, and our responses are in blue font.

Title: Should the title include 'remote sensing RZSM'?

Response: Seven RZSM products evaluated in the study are derived from land surface models, they are numerical simulation products. And the SMOS L4 RZSM product is derived from the SMOS L3 3-day SSM product using a statistical exponential filter. These RZSM products are not closely related to remote sensing. Therefore, we did not use the term "remote sensing RZSM".

Figure (all): Please increase labels and units and add more explanation to the captions to help understand more correctly.

Response: We have revised all figures and captions.

Table 1: It is nice to have this organized table for the spatiotemporal resolution of each product. Can you add a plot showing how in-situ stations overlap with each RZSM product?

Response: Figure 3 shows how the time series of in-situ observations overlap with eight RZSM products.

Line 354: Didn't > did not (or change the sentence)

Response: Correction done.

Line 446-448: Richards' equation is just an equation; what would be the way for Richards's equation to produce different soil moisture dynamics? Do you mean retention curve? or changing boundary conditions?

Response: The LSMs generally produce higher soil moisture content than in situ observations due to the overestimated precipitation input. In addition, the inertia of remaining high soil moisture content after precipitation events could be weakened by modifying the initial soil moisture values, e.g. data assimilation of soil moisture observations or reducing the model default soil moisture values. Therefore, the Richard's equation could produce different soil moisture dynamics by modifying the soil water retention curve or changing the initial and boundary conditions.

Line 499: Are these soil properties data compared with observations? Is it overestimation (compared to soil properties in-situ) or just higher estimates?

Response: These soil properties data are compared with reference dataset developed by Shangguan et al. (2013), which integrates the physical and chemical properties of 8979 soil profiles along with the soil map of China. It shows overestimation compared to the soil property reference dataset.

Line 532: I do not expect significant spatial heterogeneity for precipitation in a watershed. And what do you mean by temporal heterogeneity? Accuracy for precipitation forcing has always been a major challenge.

Response: We agree, "spatiotemporal heterogeneity" was replaced by "spatial heterogeneity".

Line 536: Does not make any sense! "overestimation of in situ observations by LSM-based RZSM products" ???

Response: L536 was removed.

Fig. 10: So what is this figure for? Is this observation and what sites? Why didn't you plot this within Fig. 9?

Response: Fig 10 shows the soil stratification in Huai river basin. The soil stratification affects the water transfer from the surface layer to the root zone layer in land surface models.

Line 618-621: The sentence needs to be restructured.

Response: L618-621 was deleted.

Line 622: Why is line 618-621 the specific reason for the difficulty of vegetation parameterization? That is a very general theory and not appropriate to be selected as a specific reason as it does not stem from the model structure.

Response: L618-622 was deleted.

Line 666-670: The whole sentence needs to be rephrased.

Response: L666-670 is replaced by:

"The statistical scores for spatial-average validation are generally better than that for point-grid validation, which are shown in Tables 3 and S1, respectively. For the point-grid validation, the spatial representativeness of *in situ* soil moisture observations at the grid scale is insufficient due to the heterogeneity of the underlying surface and precipitation forcing."

While I didn't compile the whole list of awkward/unclear sentences, I do express my concern about the clarity of this paper due to the components that I tried to address here. I do encourage authors to review and revise the paper thoroughly to make sure the paper conveys what they want to convey.

Response: We have revised the manuscript thoroughly and make it more clear and concise.

Response to Reviewer #2's comments on the manuscript egusphere-2023-1597

RC2: 'Comment on egusphere-2023-1597', Anonymous Referee #2, 12 Jan 2024

Thanks for the considerable revisions made. I still have two small comments:

Response: The authors thank the Reviewer #2 for her/his constructive and insightful comments that help us improve the quality of the manuscript. The original comments from Reviewer #2 are in black font, and our responses are in blue font.

- L104: 2.84m should be 2.89m

Response: Correction done.

- Both the conclusion and the abstract would benefit from a sentence on the outlook/implications of this research as now it is mostly summarizing the results. Here the previous comment 'Lastly, how easily can we extrapolate these results to other regions?' could help.

Response: The following text has been added in the Abstract (L39-42):

"The study provides some insights into how to improve the ability of land surface models to simulate the land surface states and fluxes by taking into account the issues mentioned above. Finally, these results can be extrapolated to other regions located in the similar climate zone, as they share the similar precipitation patterns that dominate the terrestrial water cycle."

The following text has been added in the Conclusion (L720-723):

"The study could provide some insights into how to improve the ability of land surface models to perform the land surface analysis by addressing the above issues. Furthermore, these results can be extended to other regions to improve the numerical simulation capability of land surface models at global scale."

Response to Reviewer #3's comments on the manuscript egusphere-2023-1597

RC3: 'Comment on egusphere-2023-1597', Anonymous Referee #3, 31 Jan 2024

Review of the manuscript egusphere-2023-1597 Evaluation of root-zone soil moisture products over the Huai river basin by Liu et al.

SUMMARY

I believe that the Authors prepared a comprehensive rebuttal to the comments of the Reviewers and the manuscript has been improved accordingly. Still, I found some issues that merit in my opinion further attention. These issues are described in detail below followed by more specific edits. As soon as the Authors could also improve these parts, I think the manuscript should reach a good quality for possible publication.

The authors thank the Reviewer #3 for her/his constructive and insightful comments that help us improve the quality of the manuscript. The original comments from Reviewer #3 are in black font, and our responses are in blue font.

MAJOR ISSUES

Irrigation: at L141 it is stated that the main cover types in the HRB are rainfed croplands, followed by irrigated croplands. Later (L143) it is stated that 76% of the cultivated area is irrigated. Finally (L148) it is underlined that heavy irrigation in the HRB can explain the extra water available for evaporation. Based on these statements, I'm still confused by the role of irrigation and its effect on the present intercomparison. I understand that ground soil moisture observations are in rainfed areas. Thus, some LSM that do not consider irrigation can be compared accordingly. However, recently the use of remote sensing products to estimate irrigation has been widely promoted. So, a comparison of the ground measurements to remote sensing products or soil moisture products based on the assimilation of remote sensing into LSM is in my opinion misleading as soon as the area is irrigated and the remote sensing data capture to some extent this signal. I encourage the Authors to further clarify.

Response: Current land surface models (LSMs) traditionally don't take into account irrigation practices (Romaguera et al., 2012; Lievens et al., 2017; Brocca et al., 2018; Abolafia-Rosenzweig et al., 2019). Indeed, recent studies have shown that the remote sensing soil moisture products (e.g. SMAP, Sentinel-1, ASCAT and SMOS) can capture the irrigation signal to some extent. However, these remotely sensed SM retrievals alone are insufficient to assess spatiotemporally continuous estimates of irrigation and its effects on the water and energy cycles (Abolafia-Rosenzweig et al., 2019). The assimilation of remote sensing soil moisture into LSMs has been widely conducted and shows an overall superior performance than open loop in terms of soil moisture simulations. However, the assimilation of remote sensing data mainly improves the surface soil moisture simulations, and the impact on RZSM is less pronounced and mostly neutral (Lievens et al., 2017; Reichle et al., 2017). Besides, the satellite observations with coarse spatial resolution at kilometer level (e.g. SMAP and SMOS)

struggle to resolve the local irrigation practices at field scale and can't present accurate irrigation signal. For example, Abolafia-Rosenzweig et al. (2019) indicated that Sentinel-1 soil moisture observations (10 m spatial resolution) performs better than SMAP observations in resolving irrigation practices. Escorihuela andQuintana-Seguí (2016) compared SMOS L2 soil moisture and downscaled SMOScat at 1 km scale with SURFEX LSM simulations, respectively, the high resolution SMOScat product shows a lower correlation with modelled soil moisture compared to raw SMOS L2 soil moisture at a small heavily irrigated region. Among the eight RZSM products used in the study, SMAP L4 RZSM and SMOS L4 RZSM may contain irrigation signal, as SMAP L4 RZSM assimilates L1C brightness temperature which implicitly includes an irrigation signal, and SMOS L4 RZSM is derived from SMOS L3 surface soil moisture which can detect irrigation signal but with low skill. Though SMOS L4 RZSM may contain irrigation signal, it still underestimates the *in situ* soil moisture observations. Therefore, we think the irrigation signal captured by the remote sensing soil moisture has little impact on the intercomparison and the effect of irrigation signal contained in surface soil moisture on RZSM can be ignored.

Temporal resolution: soil moisture observations are collected at 8:00 am (L154). So, the aggregation to daily average soil moisture products is not consistent (L351 - 352). It would have been better to address the temporal mismatch by, e.g., selecting the consistent hour of the soil moisture products when possible or resampling the ground soil moisture time series. Please at least clarify this issue in the methods and during the discussion.

Response: We will clarify this issue in the methods and during the discussion.

In the methods (section 3.3), L353-358 are replaced by "In terms of the temporal resolution, GLDAS_CLSM and SMOS L4 products provide RZSM data at daily time intervals. NCEP CFSv2 and GLDAS_NOAH products provide RZSM data at 3-hourly and 6-hourly time interval, respectively, which don't have consistent hour of soil moisture data with *in situ* observations only available at 08:00 AM. To keep consistent, thus the other sub-daily RZSM datasets (hourly/3-hourly/6-hourly time steps, shown in Table 1) are aggregated to daily average values to match the daily sampling frequency of the *in situ* observations."

In the discussion (section 5.5), the following text are added in L689-692.

"Finally, the temporal mismatch between model-based RZSM values which are aggregated to daily average values and in situ observations available at 08:00 AM could also induce partial bias, but this type of bias is generally small due to the low variability of soil moisture during the day."

Homogenous area. The bold statement in the rebuttal (e.g., response to comment 2 of Reviewer #1) about having a homogenous area is arguable also in the light of the data and analysis presented and discussed. E.g., fig 3 shows with gray shaded area the standard deviation of the soil moisture observation which is not negligible in my opinion. Fig 9 e fig 10 shows strong variability of soil properties. Fig S2 and S3 show strong variability of the performance when point-to-grid comparison is performed. In Section 5.5 it is discussed that

(L669) results can be interpreted considering the high spatial variability resulting from different characteristics of the underlying surface and meteorological forcing. I encourage the Authors to further clarify. Please also consider moving the results (figure S2) and the discussion (section 5.5) into the results section.

Response: We agree. The soil moisture and soil properties show strong variability across different *in situ* stations. We average different datasets (e.g. soil moisture, soil properties) to weaken the spatial heterogeneity across *in situ* stations in this study. And investigate how different factors (forcing data, model parameters, soil properties, etc.) affect the soil moisture simulation from the perspective of physical process at regional scale rather than point scale. More importantly, the focus of this study is to investigate the sources of error of the different RZSM products, which could provide some insights about how to improve the ability of land surface models to simulate the land surface states and fluxes. Therefore, the heterogeneity across different stations is not the focus of the study.

Figure S2 and section 5.5 have been moved into the results section.

MINOR ISSUES

L18-19. The manuscript does not test the capability of soil moisture products to support drought monitoring and flood forecasting. Thus, the statement "and shows the potential for drought monitoring and flood forecast" is not supported. I would remove it.

Response: Correction done.

L24-27. The sentence is not clear. Please check English grammar.

Response: L24-27 is replaced by

"The ground microwave radiation signal captured by SMOS Level 3 (L3) SSM is more attenuated in the wet season due to a substantial increase in water vapor absorption and scattering than in the dry season, which is propagated to SMOS L4 RZSM. So SMOS L4 RZSM performs better in the dry season than in the wet season."

L165. Average coverage is 38%. Do you mean that for 38% of the area you have at least 1 station per cell? Please clarify.

Response: The number of meteorological stations divide the number of grid cells is 0.38. It means per cell has 0.38 station.

L332 – 333. Since you have added the section 3.3, I think you can now remove the sentence starting with "The time series…metrics".

Response: Correction done.

Figures. Some figures (fig. 4-5-6-10-11) are cut in the lower part and labels are not always visible. Please provide new figures.

Response: We have provided new figures.

Fig.10: Please be more precise in the caption legend. what do you mean by "different stations"?, all? L547. The fact that soil properties are time invariant depends on the spatial and temporal scale of the study. I agree that porosity is considered time-invariant in these LSM but I would consider other examples, e.g., texture.

Response: The caption legend of Fig. 10 is replaced by

"Fig. 11 Boxplot of soil properties for three soil layers at all *in situ* stations (Layer1 (0-16.6 cm): plough layer; Layer 2 (16.6-49.3 cm): black soil layer; Layer3 (49.3-138.3 cm): lime concretion layer)." The soil texture (clay and sand fractions) is shown in Fig 11a and b.

L748. The 5 th statement is not supported by the present study, i.e., you do not have results that support the fact that HWSD will contribute to improve the simulation. It is a hypothesis that should be tested but is also out of scope of the present study. Please reformulate.

Response: We have deleted the 5th statement in the conclusion.

Reference

Abolafia-Rosenzweig, R., Livneh, B., Small, E. E. and Kumar, S. V.: Soil Moisture Data Assimilation to Estimate Irrigation Water Use, J Adv Model Earth Syst, 11, 3670-3690, https://doi.org/10.1029/2019MS001797, 2019.

Brocca, L., Tarpanelli, A., Filippucci, P., Dorigo, W., Zaussinger, F., Gruber, A. and Fernández-Prieto, D.: How much water is used for irrigation? A new approach exploiting coarse resolution satellite soil moisture products, International Journal of Applied Earth Observation and Geoinformation, 73, 752-766, https://doi.org/10.1016/j.jag.2018.08.023, 2018.

Escorihuela, M. J. and Quintana-Seguí, P.: Comparison of remote sensing and simulated soil moisture datasets in Mediterranean landscapes, Remote Sensing of Environment, 180, 99-114, https://doi.org/10.1016/j.rse.2016.02.046, 2016.

Lievens, H., Reichle, R. H., Liu, Q., De Lannoy, G. J. M., Dunbar, R. S., Kim, S. B., Das, N. N., Cosh, M., Walker, J. P. and Wagner, W.: Joint Sentinel-1 and SMAP data assimilation to improve soil moisture estimates, Geophys Res Lett, 44, 6145-6153, https://doi.org/10.1002/2017GL073904, 2017.

Reichle, R. H., De Lannoy, M., G. J. and Liu, Q.: Assessment of the SMAP Level-4 Surface and Root-Zone Soil Moisture Product Using In Situ Measurements, J. Hydrometeorol., 18, 2621-2645, https://doi.org/10.1175/jhm-d-17-0063.1, 2017.

Romaguera, M., Krol, M. S., Salama, M. S., Hoekstra, A. Y. and Su, Z.: Determining irrigated areas and quantifying blue water use in Europe using remote sensing Meteosat Second Generation (MSG) products and Global Land Data Assimilation System (GLDAS) data, Photogrammetric Engineering and Remote Sensing, 78, 861–873, https://doi.org/10.14358/PERS.78.8.861, 2012.