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

RC2: 'Comment on egusphere-2023-1597', Anonymous Referee #2, 10 Nov 2023

This paper presents a study on the evaluation of different global root zone soil moisture products and local observations for the Huai River Basin in China. The authors present detailed information on the local conditions, and the different gridded products used. They comprehensively compare the different products with each other and with the observations. Also, the authors provide a discussion on the potential reasons for the differences found. In general, it is an interesting study with a lot of analyses and clear visualization of the results. Nevertheless, I have a few comments that must be addressed before the manuscript can be published.

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.

General comments:

The influence of land cover, vegetation and root representation:

The authors clearly discuss potential reasons for the mismatch between in situ RZSM observations and the global products, such as forcing data and soil texture maps. Also, shortly 'different model structures and parameterizations' (L89) are mentioned as potential cause for differences. I do think there is one more very important aspect that is missed here: the role of land cover and vegetation, vegetation roots, and soil evaporation and transpiration model representation. Vegetation is usually represented by land cover maps (that are usually prescribed similar to soil maps), which can be very different for the different models. Other relevant vegetation model properties could be Leaf Area Index (see for example Nogueira et al., 2020) or the root parameterization (e.g. Stevens et al., 2020 and Van Oorschot et al., 2021). Furthermore, transpiration of crops is very dependent on the growing season, which might be not represented by the global products. I think these issues should be specifically addressed in the introduction and discussion of the results.

Response: Thank you for rigorous consideration. We do agree with your idea. We will mention these issues in the introduction and discussion. Because the inconsistent vegetation type and parameterization schemes used in LSMs, it is difficult to compare them quantitatively.

The following text will be added in the introduction.

The text (Line 88-89) will be rephrased from "Finally, the accuracy of soil moisture simulations is also affected by different model structures and parameterisations"

To "Finally, the accuracy of soil moisture simulations is also affected by inadequate model structures and inaccurate parameterization schemes. Especially for vegetation parameterizations (e.g., canopy and root tissues), which show large uncertainties in simulating the water and heat fluxes in different LSMs (Nogueira et al., 2020; Stevens et al., 2020; van Oorschot et al., 2021). For example, van Oorschot et al. (2021) proposed a climate-controlled root zone storage capacity by calculating a time-varying total soil depth based on a moisture depth model instead of using a constant of 2.84 m in the original HTESSEL land model and improving water flux simulations.”

The following text will be added in the discussion.

“Vegetation also plays a crucial role in the water and carbon exchange between atmosphere and land surface through transpiration and photosynthesis, which has significant effect on the simulation of soil moisture by LSMs, especially for RZSM. On the one hand, different land cover maps are employed in LSMs to participate in the terrestrial water and carbon cycles. For example, GLDAS_NOAH uses modified IGBP MODIS (Moderate Resolution Imaging Spectroradiometer) 20-category vegetation classification, and GLDAS_CLSM uses the University of Maryland (UMD) land cover classification based on AVHRR (Advanced Very High Resolution Radiometer) land cover map (Rui et al., 2021), MERRA-2 uses the global land cover characteristics database, version 2.0 (Reichle et al., 2017). On the other hand, the parameterization schemes for vegetation canopy (e.g., Leaf Area Index high-and low-vegetation fraction, type and density, Nogueira et al. (2020)) and root tissues (root distribution, rooting depth, root density and root zone water storage, Stevens et al. (2020) and van Oorschot et al. (2021)) vary considerably across different LSMs. Therefore, it is difficult to depict consistently and accurately the dynamic evolution of vegetation for different LSMs. Furthermore, transpiration of crops is very dependent on the growing season, which might be not well represented in the LSMs.”

Introduction
L49-67: I think this paragraph is intended to describe the state-of-the-art of global surface soil moisture, and root zone soil moisture products. The authors mention many long names of different products, which shows the detailed literature review done for this study. However, for the reader it would be more clear if the paragraphs gives a more general overview, rather than all the specific products, by answering questions such as: Why do we only have SSM direct retrievals, and not RZSM? What is available for global RZSM? How are the RZSM products generated in general?

Response: The text L49-67 will be rephrased from “Recent satellite soil moisture…exponential filter model (Albergel et al., 2008; Al Bitar and Mahmoodi, 2020).”

To “Recently, microwave-based satellite missions provide global surface soil moisture (SSM) retrievals with approximately 3-day temporal resolution, for example, SMAP and SMOS SSM is retrieved from the brightness temperature of the passive microwave radiometer. ASCAT SSM is retrieved from the backscatter coefficient of the active microwave scatterometer. However, the soil moisture is limited to the top few centimeters (0-5 cm for L-band) due to the limitations of microwave penetration depth (Kerr et al., 2001; Reichle et al., 2017b). Therefore, various approaches have been developed to estimate the RZSM and are roughly divided into three categories (Liu et al., 2023). Including, (1) statistics-based methods, such as linear regression (Zhang et al., 2017) and cumulative distribution function (Gao et al., 2019), (2) data-driven machine learning methods, such as rand forest (Carranza et al., 2021) and artificial neural network (Kornelsen et al., 2014), (3) physically based methods, such as data assimilation of satellite-derived observations into LSMs (Albergel et al., 2017; Bonan et al., 2020). Among them, the assimilation of satellite-derived observations into LSMs is considered as the most accurate method to estimate RZM due to the explicit physical mechanism, while requiring large amounts of input data (air temperature, surface pressure, wind speed, solid and liquid precipitation, incoming shortwave and longwave radiation). To date, several RZSM products have been developed for broader global applications, such as the Global Land Data Assimilation System (GLDAS_NOAH and GLDAS_CLSM) (Rodell et al., 2004), the China Land Data Assimilation System (CLDAS) (Shi et al., 2014) and the Soil Moisture Active Passive (SMAP) Level 4 (L4) (Reichle et al., 2012; Reichle et al., 2017a), the European Centre for Medium-Range Weather Forecasts (ECMWF) fifth generation reanalysis (ERA5) (Hersbach et al., 2020), the Modern-Era Retrospective Analysis for Research and Applications version 2 (MERRA-2) (Gelaro et al., 2017), and the National Centers for Environmental Prediction Climate Forecast System version 2 (NCEP CFSv2) (Saha et al., 2014). These RZSM products are generated by combining LSMs driven by meteorological forcing fields from atmospheric general circulation model (AGCM) and satellite-derived data using different data assimilation techniques (Calvet and Noilhan, 2000; Rodell et al., 2004) and provide optimal land surface states and fluxes. In addition, the Soil Moisture and Ocean Salinity (SMOS) Centre Aval de Traitement des Données (CATDS) provides SMOS L4 RZSM products, which are derived from SMOS Level 3 (L3) 3-day SSM retrievals using a statistical exponential filter model (Albergel et al., 2008; Al Bitar and Mahmoodi, 2020).

L76-94: I think this should go before L68-75.

Response: We will move Line 76-94 before Line 68-95 in the revised manuscript.

Scale issue:
The authors mention that the scale mismatch is a relevant aspect for the differences between RZSM observed in this study (L75 and Sect. 5.4). However, it is not clear from the methodology how the gridded products are aggregated to match the in situ measurements, and how the scale of the different products compare to the size of the HRB selected area. Moreover, how heterogeneous is the selected area in terms of precipitation, temperature, vegetation, and soil moisture observations? Since the results are mostly based on averages of the area, does the heterogeneity play a role?

Response: The following text (section 3.3 Validation strategies) will be added in chapter 3 Methods.

“In terms of the temporal resolution, except for the RZSM products (e.g., GLDAS_CLSM, SMOS L4) provided on daily time steps, the other sub-daily RZSM datasets (hourly/3-hourly/6-hourly time steps, shown in Table 1) are aggregated to daily average values. Therefore, the aggregated RZSM products could match the observations at daily time intervals. In terms of spatial resolution, we didn’t change the spatial resolution of any RZSM products and used the original grid resolution. Two validation strategies were used in the study. The first is to compare the RZSM time series averaged over all in situ stations with the RZSM time series averaged over all model grids where the stations are located. The second one is the single point-grid validation, the measurements at each station are compared directly with the grid values where the station is located, if there is more than one station, the measurements of these stations are averaged. The point-grid validation has been provided in the supplement (Fig. S2 and S3).”

Most of the in situ stations are located in the Huaibei Plain, which is a major grain production area. According to the land cover map of the Huai River basin (Figure 1c), most of them (56 of 58) is located in the cropland regions, we will add it in the supplement (see Table S1 below). In terms of soil properties, the lime concretion black soil is the main soil type in the Huaibei plain, which is shown in line 482. Since most of the stations in the Huaibei Plain are located in a semi-humid region, which shares the similar meteorological conditions and topography. Therefore, we believe that soil properties, vegetation, precipitation and temperature are homogeneous among different in situ stations. Moreover, the point-grid validation draws the same conclusion as the station-averaged validation. Therefore, the heterogeneity may have little effect on moisture.

Table S1 Overview of in situ stations in Huai River Basin

Station Name	Longitude (E)	Latitude (N)	Elevation (m)	Land cover
Taolaoba	117.16	32.18	48	Irrigated Crop
Chahua	116.02	33.03	39	Rainfed Crop
Hanting	116.32	33.02	28	Rainfed Crop
Songji	115.27	32.82	39	Rainfed Crop
Funan	115.57	32.64	33	Rainfed Crop
Santa	115.70	32.81	33	Rainfed Crop
Yaoli	116.17	31.82	58	Irrigated Crop
Guanting	116.85	31.80	51	Irrigated Crop
Zhuangmu	117.11	32.36	27	Irrigated Crop
Guiji	116.62	32.78	23	Irrigated Crop
Xiaji	116.54	32.65	25	Rainfed Crop
Shuangfu	115.57	33.34	37	Rainfed Crop
Fentai	115.73	33.45	35	Rainfed Crop
Santang	115.83	33.31	32	Rainfed Crop
Lixin	116.21	33.14	28	Rainfed Crop
Jieshou	115.36	33.27	42	Rainfed Crop
Yangqiao	115.39	33.02	28	Rainfed Crop
Guangwu	115.33	33.37	42	Rainfed Crop
Huangling	115.13	33.04	37	Rainfed Crop
Quanyang	115.44	33.11	35	Rainfed Crop
Kanheliu	115.85	33.10	33	Rainfed Crop
Kouziji	116.09	32.84	26	Rainfed Crop
Sanshilipu	116.11	32.70	27	Rainfed Crop
Xiaqiao	116.38	32.64	26	Rainfed Crop
Hengpaitou	116.36	31.59	72	Woodland
Xianghongdianxia	116.18	31.58	116	Woodland
Wangchenggang	116.53	31.74	76	Irrigated Crop
Lumiao	115.80	34.00	39	Rainfed Crop
Dasi	115.87	33.80	42	Rainfed Crop
Youhe	115.79	33.63	38	Rainfed Crop

Huagou	116.06	33.51	33	Rainfed Crop
Dahu	116.35	33.52	31	Rainfed Crop
Chenqiao	116.56	33.09	25	Rainfed Crop
Heliu	116.97	33.03	25	Rainfed Crop
Linhuanzha	116.57	33.67	29	Rainfed Crop
Guzhenzha	117.33	33.30	18	Rainfed Crop
Wudaogou	117.34	33.16	21	Rainfed Crop
Hexiangzha	117.18	33.00	18	Rainfed Crop
Tancheng	116.56	33.44	29	Rainfed Crop
Xibakou	117.87	33.15	11	Rainfed Crop
Xulouzha	116.75	33.92	30	Rainfed Crop
Suxianzha	117.08	33.67	28	Rainfed Crop
Gukouzha	116.45	34.27	39	Rainfed Crop
Kuaitanggou	117.55	33.75	20	Rainfed Crop
Yanglou	116.78	34.32	39	Rainfed Crop
Langanji	117.23	33.93	25	Rainfed Crop
Dulou	116.85	34.20	37	Rainfed Crop
Xiangyang	117.58	33.47	24	Rainfed Crop
Shuangdui	116.90	33.42	25	Rainfed Crop
Shuoli	116.90	34.03	32	Rainfed Crop
Huangmiao	117.65	33.08	19	Rainfed Crop
Baoji	117.11	33.16	22	Rainfed Crop
Dinghouying	117.34	33.46	24	Rainfed Crop
Xuanmiao	116.27	34.52	54	Rainfed Crop
Longhai	116.35	34.40	45	Rainfed Crop
Zhangzhuangzhai	116.60	34.12	37	Rainfed Crop
Sixian	117.92	33.43	16	Rainfed Crop
Dazhuang	117.87	33.67	20	Rainfed Crop

Discussion:

The authors explain potential causes for the mismatches between the satellite products and the observations by using specific analyses of the precipitation, temperature and soil type. Many performance metrics have been used throughout the analyses, but I think the use of these different metrics could be exploited more in the discussion. Different metrics represent different aspects of the timeseries, which could explain different processes. The authors could relate the causes in section 5.1 and 5.2 more specifically to the different metrics used. Here, also the vegetation/land cover aspect should be included as mentioned before. Lastly, how easily can we extrapolate these results to other regions?

Response: The section 5.1 and 5.2 will be rephrased as suggested. We will use more metrics and link them to different processes.

Actually, caution is required when extrapolating these results to other regions. Firstly, the uncertainty of precipitation derived from the AGCM varies considerably across different regions. For example, the large-scale stratiform precipitation is better resolved than the small-scale convective precipitation processes in the AGCM. Therefore, the precipitation simulation in the extratropical zone generally performs better than in the tropical zone, and performs better in winter than in summer (Beck et al., 2019; Lavers et al., 2022). In addition, the underlying surface conditions (e.g., land cover, soil properties) are strongly dependent on the local climate conditions. Regarding the soil properties, there are no observed soil profiles incorporated into the global soil datasets (e.g., HWSD) for some regions.

Irrigation
The role of irrigation in this study is confusing, due to the following statements:
- L119: '76% is irrigated'
- L135: 'Stations are located in areas without irrigation'
- L562: 'heavily irrigated HRB in China'
- L452: 'a signature of irrigation'
- L573: 'indirectly account for irrigation'

I understand that the entire HRB is heavily irrigated, but in this study we only look at the Huaibei Plain which has only rainfed crops as indicated in Fig. 1. It remains unclear to me which area is used for the gridded products, the entire HRB or only the Huaibai Plain? This is not entire clear from the methods. If it is the Huaibai plain (which would make more sense), then irrigation is not an issue in this paper, and should not be emphasized.

Response: The Huaibei Plain is used for the gridded products in this study. According to the land cover map (Figure 1c), the Huaibei Plain has only the rainfed crops (e.g. winter wheat, corn, Soybean, sorghum, sesame, etc.). The irrigated crops in this study mainly refer to rice fields. However, it should be noted that the Huaibei Plain still requires large amounts of irrigation, because the mean annual precipitation is less than mean annual evaporation demand. Therefore, the Huaibei Plain is prone to agricultural drought (Gou et al., 2022). Natural precipitation is the main source of water for the rainfed crops in the Huaibei Plain, supplemented by irrigation when there is no precipitation for a long time.

We completely agree with the comment that the irrigation factor is irrelevant and should not be emphasized. We will delete related statements about irrigation in the revised manuscript. L135 "Stations are located in areas without irrigation" and L452-453 "The overestimation of RZSM by ERA5 (Fig. 3) could be a signature of irrigation because the in situ RZSM observations do not capture irrigation" will be deleted in the revised manuscript.

Specific comments:
- L20,21: What are L4 and L3 here?

Response: L4 and L3 refer to Level 4 and Level 3, respectively. We will replace them with Level 4 (L4) and Level 3 (L3) in the revised manuscript.

- L59: I think ERA5, MERRA2 and NCEP CFSv2 are not the only existing global products, it might be good to emphasize this with for example 'amongst others'.

Response: Agree. L49-67 has been reworded.

"To date, several RZSM products have been developed for broader global applications, such as the Global Land Data Assimilation System (GLDAS_NOAH and GLDAS_CLSM) (Rodell et al., 2004), the China Land Data Assimilation System (CLDAS) (Shi et al., 2014) and the Soil Moisture Active Passive (SMAP) Level 4 (L4) (Reichle et al., 2012; Reichle et al., 2017a), the European Centre for Medium-Range Weather Forecasts (ECMWF) fifth generation reanalysis (ERA5) (Hersbach et al., 2020), the Modern-Era Retrospective Analysis for Research and Applications version 2 (MERRA-2) (Gelaro et al., 2017), and the National Centers for Environmental Prediction Climate Forecast System version 2 (NCEP CFSv2) (Saha et al., 2014)."

- L75: etc. is not very scientific

Response: We will delete "etc." in the revised manuscript.

- L99: it is maybe not 'difficult', but 'out of scope'

Response: L99-100 "As it is difficult…in this study" will be removed in the revised manuscript.

- L113: Fig. 1 instead of Figure 1

Response: The Figure 1 will be replaced by Fig. 1.

- Figure 1: There are two red lines in the figure, so the legend is not entirely clear

Response: We will revise the Figure 1.

- L144-145: performance metrics of P and T with respect to ground observations? This is not clear.

Response: The text (L142-143) will be rephrased from "The dataset has been extensively validated and is of high quality."

To "The dataset has been extensively validated against ground observations and is of high quality."

- L200: 'Saha and coauthors, 2011' is not a valid reference

Response: We will revise this reference, the "Saha and coauthors, 2011" was replaced by "Saha et al. 2011" in the revised manuscript.

- Chapter 2 Datasets: The authors use many different units for scale, for example 0.5°x0.5° (L137); 1:5 million (L150); 30x30 arcseconds (L163). It would be helpful for the reader to include for each scale metric a rough comparison to for instance kilometre to easily compare the resolution of the different products.

Response: L137 will be replaced by "with a spatial resolution of 0.5° (approximately 55.6 km)".

L150 will be replaced by "Soil databases used in many global LSMs have traditionally relied on the FAO/UNESCO 1:5 million scale World Soil Map with a spatial resolution of 5 arc minutes (approximately 10 km)".

L156 will be replaced by "with a resolution of 30 arcseconds (approximately 1 km)".

L163 will be replaced by "The dataset provides information on soil properties for eight layers

(0-2.3 m) at a spatial resolution of 30×30 arcseconds (approximately 1 km)".

- Section 2.1 'The HRB study area': for the reader the full name of HRB would be more clear

Response: We will use "The Huai River Basin study area" instead of "The HRB study area".

- Section 2.3 and 2.4: the authors describe a lot of different products, but it is not directly clear from these sections what is actually used for this study. Both sections could be much more concise when only referring to the relevant information for this study. For example, it is not directly relevant that ERA5 'covers the period from January 1940 to present … ocean waves' (L170). Table 1 gives a very concise overview, and could be valued more and referred to more often in the text.

Response: The section 2.3 and 2.4 will be rephrased for a more concise description in the revised manuscript.

At the end of section 2.3, the following text will be added in the revised manuscript

"The FAO/UNESCO and HWSD V1.2 soil datasets are employed in different LSMs, respectively. A China soil dataset developed by Shangguan et al., (2013) is used as a reference to evaluate the soil properties of FAO/UNESCO and HWSD V1.2 datasets.

- Table 1: a reference would be more informative than the 'data access' column

Response: The data access will be replaced by the following reference.

GLDAS_NOAH:

Beaudoing, H. and M. Rodell, NASA/GSFC/HSL (2020), GLDAS Noah Land Surface Model L4 3 hourly 0.25 x 0.25 degree V2.1, Greenbelt, Maryland, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: [21 September 2021], 10.5067/E7TYRXPJKWOQ.

GLDAS_CLSM:

Li, B., H. Beaudoing, and M. Rodell, NASA/GSFC/HSL (2020), GLDAS Catchment Land Surface Model L4 daily 0.25 x 0.25 degree GRACE-DA1 V2.2, Greenbelt, Maryland, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: [22 September 2021], 10.5067/TXBMLX370XX8.

ERA5:

Hersbach, H. and Coauthors: ERA5 hourly data on single levels from 1979 to present, Copernicus Climate Change Service (C3S) Climate Data Store (CDS), Accessed: 22 September 2021, https://doi.org/10.24381/cds.adbb2d47, 2018.

MERRA-2:

Global Modeling and Assimilation Office (GMAO) (2015), MERRA-2 tavg1_2d_lnd_Nx: 2d,1-Hourly, Time-Averaged, Single-Level, Assimilation, Land Surface Diagnostics V5.12.4, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: 26 September 2021, 10.5067/RKPHT8KC1Y1T.

NCEP CFSv2:

Saha, S., et al. 2011, updated monthly. NCEP Climate Forecast System Version 2 (CFSv2) Selected Hourly Time-Series Products. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. https://doi.org/10.5065/D6N877VB. Accessed 28 October 2021.

SMAP L4:

Reichle, R., G. De Lannoy, R. D. Koster, W. T. Crow, J. S. Kimball, and Q. Liu. (2020). SMAP L4 Global 3-hourly 9 km EASE-Grid Surface and Root Zone Soil Moisture Geophysical Data, Version 5 [Data Set]. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. https://doi.org/10.5067/9LNYIYOBNBR5. Date Accessed 06-04-2021.

SMOS

CATDS (2021), CATDS-PDC L4SM RZSM – 1 day global map of root zone soil moisture values from SMOS satellite. CATDS (CNES, IFREMER, CESBIO). http://dx.doi.org/10.12770/316e77af-cb72-4312-96a3-3011cc5068d4. Date Accessed 17-09-2021.

CLDAS:

CMA (2020), The near-real-time product dataset of the China Meteorological Administration Land Data Assimilation System (CLDAS-V2.0). Available at http://data.cma.cn/en/?r=search/uSearch&keywords=cldas. Date Accessed 16-11-2021.

- Section 4.1: the first paragraph is not easily readable for the author due to all the numbers. All the numbers are also presented in Table 3, so it suffices to only mention the highly relevant numbers in the text here. Table 3 could also be combined with Fig. 2, same for Fig. 8 and Table 4.

Response: The first paragraph will be replaced by the following text:

“Figure 2 shows scatterplots of RZSM products against the in situ measurements averaged across all in situ stations over the HRB, from 1 April 2015 to 31 March 2020. Regarding the bias, except for the underestimation by SMOS L4 (-0.047 $m^3\ m^{-3}$), all the other products overestimate the RZSM observations by 0.030 $m^3\ m^{-3}$ to 0.117 $m^3\ m^{-3}$ (SMAP L4 and ERA5, respectively). ERA5 and CLDAS have the largest RMSE values among all the RZMS products due to the relatively large bias. Regarding correlation and ubRMSE, GLDAS_CLSM ($R = 0.69$, ubRMSE = 0.018 $m^3\ m^{-3}$) outperforms the other RZSM products, followed by MERRA-2, ERA5, CLDAS, SMAP L4, GLDAS_NOAH, NCEP CFSv2 and SMOS L4. Overall, GLDAS_CLSM performs best among the eight RZSM products in terms of R, ubRMSE and bias values, while SMAP L4 presents the lowest RMSE and the lowest bias. SMOS L4 presents the worst performance with the lowest R value. The detailed statistics are shown in Table 3.

The description about Fig. 8 and Table 4 (L394-398) will be replaced by “

The daily air temperature data derived from ERA5, MERRA-2, NCEP CFSv2, GLDAS_CLSM, CLDAS, GLDAS_NOAH and SMAP L4 are validated against in situ observations of daily air temperature after aggregating all sub-daily products to daily time steps. Figures 8 and S4 shows that the modelled air temperature captures the observed temporal variation well, with R values above 0.96. However, all of them show slight underestimation, indicated by negative bias values ranging from -4.0 to -5.2 K. In terms of the comprehensive scores of the four statistical metrics, GLDAS_NOAH air temperature outperforms the other datasets and SMAP L4 shows the worst performance. Detailed statistics are shown in Table 4.

- Figure 2 and 8: to improve visualization I would recommend to include density of points with colours (for example https://stackoverflow.com/questions/20105364/how-can-i-make-a-scatter-plot-colored-by-density)

Response: Figures 2 and 8 will be revised to use the same layout as Figure 4.

- Figure 4:
  - o The labels are not readable because of the small font

    Response: We will revise Fig. 4.

  - o What are the lines? A fit through the data points? Why do the authors use a different lay-out for a scatterplot than in Fig. 2 and 8?

    Response: The line in each subplot is a fit through the data points. Figures 2 and 8 will be revised to use the same layout as Figure 4. And we will add the explanation in the legend of Fig. 2, 4 and 8.

- Figure 8 and Table 4: I think it is more convenient to use degree Celsius than Kelvin for temperatures.

Response: The Kelvin is the international system (SI) of unit for thermodynamic temperature. The SI units should be used according to the requirement of HESS submission.

- L381: ‘good agreement’ is a subjective statement, I think it is questionable if a R>0.4 is ‘good’.

Response: This L381 will be replaced by “Overall, the R values between precipitation products and the observed precipitation is higher than 0.4 (left panel of Fig. 7)”

- Figure 9: This figure implies that also the models differentiate soil moisture for layer 0-30cm and 30-100cm, while for most models this is not the case?

Response: This figure doesn’t mean that the models differentiate soil moisture for layer 0-30cm and 30-100cm. In most LSMs, the soil layers for soil moisture and temperature simulations are generally not for layer 0-30 cm and 30-100 cm, e.g., 0-10 cm, 10-40 cm, 40-100cm and 100-200cm. When the soil properties (top layer: 0-30 cm and subsurface layer 30-100 cm) provided by FAO/UNESCO and HWSD are used in different LSMs, they are processed differently in the LSMs. For example, GLDAS uses the top layer soil parameter data for all layers (see GLDAS Soil Land Surface | LDAS (nasa.gov)). Such as the top layer soil properties are used to represent four soil layers (0-10 cm, 10-40 cm, 40-100 cm and 100-200 cm) for GLDAS_NOAH and two soil layers (surface:0-2 cm, root zone: 0-100 cm) for GLDAS_CLSM in terms of soil moisture.

- Abstract: the statements about the explanations of the differences are quite strong, because we know there is many other factors that play a role.

Response: We will weaken the statements about the explanations of the differences. The abstract will be replaced by the following text in the revised manuscript.

"Root zone soil moisture (RZSM) is critical for water resource management, drought monitoring and sub-seasonal flood climate prediction. While RZSM is not directly observable from space, several RZSM products are available and widely used at global and continental scales. This study conducts a comprehensive quantitative evaluation of eight RZSM products over the Huai River Basin (HRB) in China. The assessment is performed using observations from 58 in situ soil moisture stations from 1 April 2015 to 31 March 2020. Attention is drawn to the potential factors that contribute to the uncertainties of model-based RZSM, including errors in atmospheric forcing (precipitation, air temperature), vegetation parameterizations, soil properties, and spatial scale mismatch, etc. The results show that the Global Land Data Assimilation System Catchment Land Surface Model (GLDAS_CLSM) outperforms other RZSM products with the highest correlation coefficient (R=0.69) and the lowest unbiased root mean square error (ubRMSE=0.018 m^3 m^{-3}), respectively. All RZSM products tend to overestimate in situ soil moisture values, except for the Soil Moisture and Ocean Salinity Level 4 (SMOS L4) product, which underestimates RZSM. The underestimation of Surface Soil Moisture (SSM) in SMOS Level 3 (L3), caused by underestimated physical surface temperature and overestimated ERA interim soil moisture, may contribute to the underestimation of RZSM in SMOS L4. The other model-based RZSM products show an overestimation of in situ observations, which could be associated with the overestimation of the precipitation amounts and precipitation events (drizzle effects) and the underestimation of air temperature. In addition, the biased soil texture (organic carbon, clay and sand fractions) and flawed vegetation parameterizations (e.g., canopy and root tissues) affect the hydrothermal transport processes represented in different LSMs, leading to inaccurate soil moisture. The intercomparison of the eight RZSM products shows that MERRA-2 and SMAP L4 RZSM have the highest correlation, which could be attributed to the fact that both products use the catchment land surface model and the atmospheric forcing provided by the Goddard Earth Observing System Model, version 5 (GEOS-5), although the versions differ slightly. This in situ validation shows that GLDAS_CLSM could be used for drought monitoring and flood forecast in Huaibei Plain. Moreover, the RSZM intercomparison indicates that the model should focus on increasing the frequency of dry soil moisture, decreasing the frequency of wet soil moisture and the ability to capture the frequency peak of soil moisture. The uncertainty analysis implies that the model-based RZSM can be improved by correcting precipitation, using more accurate soil properties and more prefect vegetation parameterization schemes, etc."

Reference

Beck, H. E., Pan, M., Roy, T., Weedon, G. P., Pappenberger, F., van Dijk, A. I. J. M., Huffman, G. J., Adler, R. F. and Wood, E. F.: Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the CONUS, Hydrology and Earth System Sciences, 23, 207-224, 10.5194/hess-23-207-2019, 2019.

Gou, Q., Zhu, Y., Lü, H., Horton, R., Yu, X., Zhang, H., Wang, X., Su, J., Liu, E., Ding, Z., Wang, Z. and Yuan, F.: Application of an improved spatio-temporal identification method of flash droughts, J. Hydro., 604, 127224, https://doi.org/10.1016/j.jhydrol.2021.127224, 2022.

Lavers, D. A., Simmons, A., Vamborg, F. and Rodwell, M. J.: An evaluation of ERA5 precipitation for climate monitoring, Quarterly Journal of the Royal Meteorological Society, 148, 3152-3165, https://doi.org/10.1002/qj.4351, 2022.

Nogueira, M., Albergel, C., Boussetta, S., Johannsen, F., Trigo, I. F., Ermida, S. L., Martins, J. P. A. and Dutra, E.: Role of vegetation in representing land surface temperature in the CHTESSEL (CY45R1) and SURFEX-ISBA (v8.1) land surface models: a case study over Iberia, Geoscientific Model Development, 13, 3975-3993, https://doi.org/10.5194/gmd-13-3975-2020, 2020.

Reichle, R. H., Draper, C. S., Liu, Q., Girotto, M., Mahanama, S. P. P., Koster, R. D. and De Lannoy, G. J. M.: Assessment of MERRA-2 Land Surface Hydrology Estimates, Journal of Climate, 30, 2937-2960, https://doi.org/10.1175/jcli-d-16-0720.1, 2017.

Rui, H., Beaudoing, H. and Loeser, C.: README Document for NASA GLDAS Version 2 Data Products, Available at https://hydro1.gesdisc.eosdis.nasa.gov/data/GLDAS/GLDAS_NOAH025_3H.2.1/doc/README_GLDAS2.pdf,2021.

Stevens, D., Miranda, P. M. A., Orth, R., Boussetta, S., Balsamo, G. and Dutra, E.: Sensitivity of Surface Fluxes in the ECMWF Land Surface Model to the Remotely Sensed Leaf Area Index and Root Distribution: Evaluation with Tower Flux Data, Atmosphere, 11, http://doi.org/10.3390/atmos11121362, 2020.

van Oorschot, F., van der Ent, R. J., Hrachowitz, M. and Alessandri, A.: Climate-controlled root zone parameters show potential to improve water flux simulations by land surface models, Earth System Dynamics, 12, 725-743, https://doi.org/10.5194/esd-12-725-2021, 2021.