Referee Comments#2:

General comments:

This interesting analysis used in-situ observations in China to evaluate several reanalysis- and RS-based SM products. While it is a nice self-contained study with seemingly comprehensive analyses, I found the study lacking sufficient physical explanations supporting several findings of their analyses. Also, some figures are not very well presented and need to be updated. Therefore, I'd suggest the authors go through moderate revisions before this paper can be publishable. Below are some suggestions to improve the paper:

Response: Thanks very much for your constructive suggestion. During this revision, the manuscript is improved by focusing on the following issues:

(1) Evaluation strategies have been improved by (i) using unbiased root mean square error (ubRMSE) to remove the bias error caused by the mismatch of spatial representativeness between in situ data and all SM products; (ii) removing all the product data (including remotely sensed and reanalysis) when in situ observation were missing. As a result, all the related figures have been refined and corrected.

(2) More physical explanations have been added in the Results and Discussion Section.

At the discussion section, uncertainties caused by comparing in situ observations with all products at different layers and grid mismatch are discussed.

Line 368-373: *ESA CCI SM product showed the top layer soil content at 5-cm depth or so. The in-situ measurement depth and model output are at the depth of 0-10cm, which were also treated as the top layer soil content. Such difference would also cause representativeness errors. Previous studies have found that there is a close relationship between surface SM and SM in the upper ten centimetres (i.e., Albergel et al., 2008; Dorigo et al., 2015), so the SM measurements at the depth of 10 cm were chosen as the reference to evaluate satellite-based and reanalysis products. Furthermore, introducing ubRMSE and conducting comparison at regional scale can remove the bias error caused by mismatch of grid cell to some extent.*

We further discuss why ESA CCI showed lower correlation with in situ observations.

Line 374-380: *The ESA CCI combined data generally increase the number of observations available for a time period but the correlation coefficients were not better than those of the best performing single dataset (Dorigo et al., 2015). Dorigo et al. also studied the possible reasons of input data, and found that the low correlation of combined product possibly due to the merging procedure, including the influence of vegetation (Taylor et al., 2012), the different original overpass time, and the scaling of high resolution ASCAT product to lower resolution reference products. Beck et al. (2021) found that ESA CCI SM performed better in eastern Europe in terms of high-frequency fluctuations, and speculated the overall performance of ESA CCI may be not so good due to incorporating ASCAT that performed less well.*

The physical explanations of spatio-temporal SM variation have also been added.

Line 387-392: *Precipitation and evaporation are found to be the most important determinant of soil moisture simulation performance, in which the evaporation is associated with temperature and radiation (Gottschalck et al., 2005; Mall et al., 2006; Chen & Yuan, 2020). SM value in the analysis is overestimated, partly due to the reason that the JJA precipitation over China is overestimated by models (e.g., Luo et al., 2013; Yun et al., 2020). The largest bias of precipitation overestimation using the hourly 31-km-resolution ERA5 reanalysis data is found over the Tibetan and Yun-Gui Plateaus, the North China Plain, and the southern mountains, which gives one the explanation why reanalysis products represent the worst performance over the NC region.*

Insufficient explanations/supports:

1: ESA-CCI seems to not represent seasonality well. Why? It seems no variation there. I think this explanation on "which may be because of snow or frozen soil during these periods" is too thin. To me this still does not explain well on why worst seasonality are there.

Response: In the revised manuscript, monthly SM data in cold seasons (frozen and snowing) were deleted, so we deleted this sentence.

Line 285-286: *ESA CCI yields the worst seasonal cycle results considering the changing*

tendency, which may be because of lack of available data by conditional constraints of satellite sensors.

We also add the explanation in the Discussion section.

Line 374-380: *The ESA CCI combined data generally increase the number of observations available for a time period but the correlation coefficients were not better than those of the best performing single dataset (Dorigo et al., 2015). Dorigo et al. also studied the possible reasons of input data, and found that the low correlation of combined product possibly due to the merging procedure, including the influence of vegetation (Taylor et al., 2012), the different original overpass time, and the scaling of high resolution ASCAT product to lower resolution reference products. Beck et al. (2021) found that ESA CCI SM performed better in eastern Europe in terms of high-frequency fluctuations, and speculated the overall performance of ESA CCI may be not so good due to incorporating ASCAT that performed less well.*

2: it seems discouraging that none of the products available captures the anomalies well especially in NC. Can the author provide some feasible explanations on why this is the case, and discuss how this could influence applications in those regions and what are the potential future directions for improvements?

Response: What is worthy to say is that, Figure 8 showed the interannual anomalies of JJA SM. Surface SM is a variable associate with precipitation and evaporation, both of which fluctuate greatly with time in the JJA seasons. To improve the quality of SM, all reanalysis data would improve their performance in representing precipitation and evaporation, especially during extreme events.

We also added some discussion in **Line390-392:**

The largest bias of precipitation overestimation using the hourly 31-km-resolution ERA5 reanalysis data is found over the Tibetan and Yun-Gui Plateaus, the North China Plain, and the southern mountains, which gives one the explanation why reanalysis products represent the worst performance over the NC region.

3.Line 301: I think “which is partly due to the combined influence of longwave and shortwave radiation” does not sufficiently explain why low correlation there. Please

expand what you mean exactly. Also, if separation of LW and SW radiation helps, would it be possible to use LW and SW data to re-draw this scatter plot?

Response: Figure 12 was improved by adding more information about the MAM and SON seasons, and the description was also refined as follows.

Line 358-366: *Previous studies have showed that soil moisture is influenced by the combination of precipitation and evaporation, in which land surface evaporation is linked with temperature and surface net radiation (Jasper et al., 2006; Harmsen et al., 2009). Figure 12 shows scatter plots of (a, d, g) precipitation, (b, e, h) temperature, and (c, f, i) net radiation anomalies versus observed SM anomalies over different regions in (left column) MAM, (middle column) JJA, and (right column) SON seasons. Obvious positive correlations are found between precipitation and SM in the YH regions during MAM and SON seasons, and in the NE and NC regions during JJA season. Temperature and net radiation show negative correlation with in the NE, NC, and YH regions. The correlation coefficient is low for all meteorological variables in the NW region, which may be attributed to the special soil type there. Soil moisture in the NE and NC regions tends to be influenced by temperature during cold seasons. SM in the YH region tend to be influenced by radiation during warm seasons, due to the large evaporation there.*

4. 12 & L298-L302: overall I think it's an interesting figure. However, authors fail to explain in more detail on the underlying physical mechanisms responsible for these correlations and why they wanted to perform these analyses. This paragraph is too thin. In addition, It seems these plots are more driven by the availability of data, instead of driven by hypothesis testing needs. It would be helpful for the authors to put more thoughts on this figure and provide readers with more insights on why they chose to do the analysis and what's new after doing the analysis.

Response: The aim of this figure is to study the soil moisture memory in different seasons. We have provided more insights as follows:

Line 306-308: *The information of soil moisture autocorrelation gives hint for the assimilation of surface soil moisture into land surface models (Crow and Van den Berg, 2010), in which during summer and winter, some other related meteorological elements*

should be considered.

Related references:

Crow, W., and Van den Berg, M.: An improved approach for estimating observation and model error parameters in soil moisture data assimilation, Water Resources Research, 46(12), doi:10.1029/2010WR00940, 2010.

Figure presentation problems:

5. It is very difficult to distinguish in-situ line in Fig. 6 as it can be confused with ERA-5. I'd suggest to use thicker black line to denote in-situ observations in Fig. 6. Also, be better to use consistent legend with Fig. 4 & Fig. 8.

Response: As suggested, thicker black line has been used to denote in-situ observations in Fig. 6. Furthermore, the legends in Fig. 4 and Fig.8 have been unified.

6. : I think it would be very difficult for readers to directly extract useful information from this figure, partly because of the color bar used, which makes it all red (plus there are so many panels). I'd suggest to use more continuous colors, with more contrasting from 0-1, such that differences in the correlations are better presented. Since only very few locations show negative correlations, you can cap the lower bound at 0, and just mention "limited negative correlation" in the caption. This way, 0-1 can be better contrasted (using blue to red) to support your interpretation on the figure in the main text.

Response: Thanks for your suggestion, and the color bar has been modified in the revised manuscript. Furthermore, more discussion has been added in the main text as answered in Response#4.

7. 10 & Fig. 11: the caption is incomplete and misleading. It did not mention which skill metrics is plotted here. Please mention it explicitly in the caption. Also, draw a reference line on 0 such that readers know where to expect good performance.

Response: Fig.10 and Fig.11 were changed by showing the rBias and ubRMSE of remote sensing and reanalysis SM against in situ observations under extreme (and severe) dry or extreme (and severe) wet conditions, and the figure caption has been improved.

Minor:

8. I do not think the literature review is comprehensive. Beck et al. (2021; HESS) presented a much more comprehensive study performed at the global scale. It should be included in the Introduction and discussions on relevance to your study needs to be mentioned. I disagree with the claim in L62 that no long-term SM products have been compared with ESA-CCI. Please revise accordingly.

Response: Thanks for your recommendation. We have added Beck et al. (2021, HESS) in the Introduction section, and include it in the discussion as the support of the physical mechanisms especially in the Discussion section.

Line 378-380: *Beck et al. (2021) found that ESA CCI SM performed better in eastern Europe in terms of high-frequency fluctuations, and speculated the overall performance of ESA CCI may be not so good due to incorporating ASCAT that performed less well.*

Line 401-403: *Beck et al. (2021) concluded that assimilating satellite soil moisture estimate maybe not improve more than increasing model resolution or improving soil moisture simulation ability, which is in line with our results. This suggest that improving model simulation performance of SM is beneficial especially at long-term scales.*

Furthermore, "no long-term SM products have been compared with ESA-CCI" has been adjusted as follows:

Line 64-65: … *few studies on long-term SM products over 30 years have been compared with the ESA CCI product using in situ measurements in East China*.

Overall comment:

9. In fact, I like the study very well because it is self-contained, with comprehensive analysis, and the writing is good too. However, I am thinking what could be more useful to the community, is perhaps for the authors to share their in-situ soil moisture observations through posting the data via figshare or other publicly accessible data portal. It seems to me that this study is only unique because of its observations, which are generally not shared with the public. If the data can be shared properly with the whole community, people may find more innovative ways of using the data for other research purposes such as drought

monitoring. Is this something that the authors are considering? It could be helpful to at least comment on or discuss this issue in an academic paper.

Response: Thanks for your suggestion, and the in situ data can be obtained by requesting from the International Soil Moisture Network website (https://ismn.geo.tuwien.ac.at/en/), and National Meteorological Information Center of China (NMIC, http://data.cma.cn/site/index.html). We have added the above information in the Data Availability section.

Line 422-425: *The updated Chinese soil moisture presented as volumetric soil moisture (θv, unit=m^3 m^{-3}) for 1981 to 1999 was downloaded from the International Soil Moisture Network website (https://ismn.geo.tuwien.ac.at/en/). The in situ SM measurements are obtained by requesting from the website of National Meteorological Information Center of China (NMIC, http://data.cma.cn/site/index.html).*