Referee Comments#1:

General comments:

Soil moisture, as one of the essential climate variables, has attracted more and more attention from climate research. However, there is still a long way to go for the recently widely used soil moisture products, including reanalyses based on models and retrievals from remotely sensed data, to be comparable with observations. To further develop and properly use them, it is necessary to compare with in situ observations to reveal their uncertainties. In this manuscript, the five satellite-based and reanalysis soil moisture products were evaluated in China with in situ observations for top soil layer (0-10 cm). By now the manuscript still needs to further discuss the uncertainties of in situ observations of soil moisture data, the influence of sparse data samples, and thus the unfair to compare grid products using point-scale measurements. In particular, the author pointed out that the bias term controlled the deviations of soil moisture products from the observed values. This partly stems from the spatial mismatches in the comparisons of the soil moisture measured at a point with model grid means. So, it requires more discussion about its implications. In addition, the method part needs to provide more details, for example, how the monthly means were estimated using 3-sample observations per month.

Response: We appreciate your comments, which are helpful for us to further improve this paper. In the revised manuscript, we have focused on the following issues.

(1) More detailed information has been added in the revised manuscript, such as the combination of active and passive product (see response#2), improvement of ERA5 to ERAI (see response#3), how monthly means were estimated using 3-sample observations per month (see response#4), and so on.

(2) In order to remove the bias error caused by the mismatch of spatial representativeness between in situ data and all SM products, the unbiased root mean square error (ubRMSE) was introduced to evaluate temporal dynamic variability (see Figure 3, Figure 11 and Table 3). Furthermore, the comparison was conducted at regional scales by calculating the reginal average of monthly value for all SM products, which can reduce the uncertainty caused by grid mismatch to some extent.

(3) A more in-depth discussion especially focusing on the physical explanations has

been added in the revised manuscript.

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.*

Specific comments and suggestions:

1. "mainland China" is NOT a right term, you can use: the Chinese Mainland, Mainland of China or China's Mainland.

Response: Thanks for your suggestion, and all the "mainland China" have been changed into "the Chinese Mainland" as suggested.

2. 2.1.1 ESA CCI SM, how the various retrievals of the passive and active sensors combined should be detailed a bit more, for example, using land surface model products?

Response: Thanks for your advice, and we have added more detail information about how the various retrievals of the passive and active sensors combined in the revised manuscript, as showed in Section 2.1.1. we also added the related reference.

Line 84-96: *The ESA CCI SM v04.4 combined product is employed in this study, which provides SM data starting from November 1978 until June 2018 with a spatial resolution of 0.25°. The project of ESA CCI is to use C-band microwave scatterometers (Aqua satellite and the Advance Scatterometer, ASCAT) and multi-channel microwave radiometers (SMMR, SSM/I, TMI, AMSR-E, WindSat, AMSR2) to produce a long-term reliable time series of SM (Chakravorty et al., 2016). The ESA CCI SM v4 is better at detecting SM changes (Balenzano et al., 2011) than previous versions, as it merges all active and passive Level 2 products directly to generate the combined product, rather than creating active and passive products separately and then merging together (ESA, 2018; Gruber et al., 2019). Global Land Data Assimilation System Noah (GLDAS 2.1) was used as a scaling reference in the combined product to obtain a consistent climatology, flagging of high vegetation optical depth (VOD) for Soil Moisture and Ocean Salinity (ESA SMOS) and AMSR-2 method changed (Dorigo et al., 2017; Pasik et al., 2020). A polynomial SNR-VOD regression and the p-value based mask was used to fill spatial gaps in TC-based SNR estimates, and exclude unreliable input dataset in the combined product, respectively. Here, we evaluate all the products over the period from 1981 to 2013 (the same as below), during which in situ measurements are also available. The top layer of ESA CCI SM data at the depth of 2~5 cm depth are estimated.*

Related references:

Balenzano, A., Mattia, F., Satalino, G., and Davidson, M. W. J.: Dense temporal series of C- and L-band SAR data for soil moisture retrieval over agricultural crops, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., 4, 439–450, https://doi.org/10.1109/jstars.2010.2052916, 2011.

Chakravorty, A., Chahar, B. R., Sharma, O. P., and Dhanya, C. T.: A regional scale performance evaluation of SMOS and ESA-CCI soil moisture products over India with simulated soil moisture from MERRA-Land, Remote Sens. Environ., 186, 514–527, https://doi.org/10.1016/j.rse.2016.09.011, 2016

Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber, A., Haas, E., Hamer, P. D., Hirschi, M., Ikonen, J., de Jeu, R., Kidd, R., Lahoz, W., Liu, Y. Y., Miralles, D., Mistelbauer, T., Nicolai-Shaw, N., Parinussa, R., Pratola, C., Reimer, C., van der Schalie, R., Seneviratne, S. I., Smolander, T., and Lecomte, P.: ESA CCI soil moisture for improved earth system understanding: state-of-the art and future directions, Remote Sens. Environ., 203, 185–215, https://doi.org/10.1016/j.rse.2017.07.001, 2017.

Gruber, A., Scanlon, T., van der Schalie, R., Wagner, W., and Dorigo, W.: Evolution of the ESA CCI Soil Moisture climate data records and their underlying merging methodology, Earth Syst. Sci. Data, 11, 717–739, https://doi.org/10.5194/essd-11-717-2019, 2019.

Pasik, A., Scanlon, T., Dorigo, W., de Jeu, R.A.M, Hahn, S., van der Schalie, R., Wagner, W., Kidd, R., Gruber, A., Moesinger, L., Preimesberger, W.: ESA Climate Change Initiative Plus - Soil Moisture: Algorithm Theoretical Baseline Document (ATBD) Supporting Product Version v05.2, Earth Observation Data Centre for Water Resources Monitoring (EODC) GmbH, TU Wien, VanderSat, CESBIO and ETH Zürich, pp: 71, 2020.

3. 2.1.5 ERA5 SM, the improvements of land processes in ERA5 against ERAI are helpful to understanding of the results with respect to in situ soil moisture in these two reanalysis.

Response: The main improvements of ERA5 against ERAI are listed as follows: (1) uses a new version of the ECMWF assimilation system IFS (IFS Cycle 41R2); (2) combines vast amounts of historical observations, including ozone, aircraft and surface pressure data, as well as various newly reprocessed datasets and recent instruments that could not be ingested in ERA-Interim; (3) includes more model input, for example, the World Climate Research Programme (WCRP) Coupled Model Intercomparison Project (CMIP) greenhouse gases, volcanic eruptions, sea surface temperature (SST), and sea-ice cover, which are appropriate for climate; and (4) has much higher spatial and temporal resolution. And the above information was added in Section 2.1.5.

Line 117-124: *ERA5 is the latest reanalysis product produced by ECMWF, covering the period from 1979 to present. The product uses a new version of the ECMWF assimilation system IFS (IFS Cycle 41R2), and combines vast amounts of historical observations, including ozone, aircraft and surface pressure data, as well as various newly reprocessed datasets and recent instruments that could not be ingested in ERA-Interim (C3S, 2017). The ERA5 model input includes the World Climate Research Programme (WCRP) Coupled Model Intercomparison Project (CMIP) greenhouse gases, volcanic eruptions, sea surface temperature (SST), and sea-ice cover, which are appropriate for climate. Furthermore, the spatial (31 km globally) and temporal (hourly) resolutions of ERA5 are rather high compared to ERAI. ERA5 will eventually cover the period from 1950 to the present, and one of its key improvements is better SM (Komma et al., 2008).*

4. 2.2 In Situ SM and Preprocessing of Datasets, the in situ observations were took from three datasets, so details about difference in the operation of measurements and the means of quality control for the datasets are necessary to assess the credibility of in situ data.

Response: The detailed information about the operation of measurements has been added in the revised manuscript, and also listed in Table 1.1.

Line 132-135: *The ISMN provides a global in-situ soil moisture database, which has been widely used for validation of satellite products and model simulation (e.g. Albergel et al., 2012). The SM data at the depth of 0~5 cm and 5~10 cm was obtained and averaged as the value at the depth of 0~10 cm.*

Line 140-141: *The SM data was observed at the depth of 10 cm, 20 cm, 50 cm, 70 cm, and 100 cm using drying methods, with the data at 10-cm depth utilized.*

Line 153-155: *The SM mass percent was measured at 11 levels with the depth of 0~5* *cm, 5~10 cm, 10~20 cm, 20~30 cm, 30~40 cm, 40~50 cm, 50~60 cm, 60~70 cm, 70~80 cm, 80~90 cm, and 90~100 cm, in which the value at 10 cm depth are calculated as the average of the values at the depth of 5~10 cm and 10~20 cm.*

Line 156-159: *Considering that the field capacity and the dry bulk density are not measured at all stations, data from 119 stations are selected from 1981 to 2013. Not all in situ data were suitable for evaluation given instrumental error and observational conditions, for example, the available measurement period, installation depth and sensor placement. Therefore the evaluation was conducted in unfrozen and snow-free seasons, such as June-July-August (JJA).*

5. The 'CN05.1' should be defined before its first citation.

Response: The 'CN05.1' was defined as 'the station observational meteorology dataset (CN05.1)' in the revised manuscript (**Line 168-169**).

6. Line 155, 'different drought/well conditions', 'well' is a typing error?

Response: Sorry for the typo, and the "well conditions" has been changed into "wet conditions" **(Line 173)**.

7. More detailed information on the decomposition of MSEs and the test methods is necessary for potential readers.

Response: The evaluated metrics (including bias, relative bias, the Pearson correlation coefficient, root mean square difference (RMSD), and the unbiased relative root mean square error (ubRMSE) has been added in Section 2.4.1 (see Equation (3) to (7)).

Line 179-194: *The comparisons were conducted through the statistical metrics, such as the Bias, relative Bias (rBias), Pearson correlation coefficient (R), root mean square difference (RMSD), and the unbiased root mean square error (ubRMSE) using the following formulas:*

$$Bias = \frac{\sum_{t=1}^{n}(x_{p,t} - x_{obs,t})}{n} \tag{3}$$

$$rBias = \frac{Bias}{Mean(Observation)} \tag{4}$$

$$R = \frac{\sum_{t=1}^{n}(x_{obs,t} - \mu_{obs})(x_{p,t} - \mu_p)}{\sqrt{\sum_{t=1}^{n}(x_{obs,t} - \mu_{obs})^2}\sqrt{\sum_{i=1}^{n}(x_{p,t} - \mu_p)^2}} \tag{5}$$

$$RMSD = \sqrt{\frac{\sum_{t=1}^{n}(x_{p,t} - x_{obs,t})^2}{n}} \quad (6)$$

$$ubRMSE = \sqrt{RMSD^2 - Bias^2} \quad (7)$$

in which n is the total number of time steps, $x_{p,t}$ *and* $x_{obs,t}$ *is the value of SM products (including remote sensing and reanalysis) and observation at time-step t,* μ_{obs} *and* μ_p *are the mean of the in situ observed values and all SM products, Mean(observation) is the average of observation. The metrics of rBias was used to study the performance of various regions under different drought or wet conditions. The ubRMSE is introduced to evaluate temporal dynamic variability to get rid of the bias error caused by the mismatch of spatial representativeness between the in situ data and all SM products (Jackson et al., 2010, 2012; Entekhabi et al., 2014). What is worthy to say, the in situ observation were not considered as 'true' value because of instrumental errors and representativeness, so the RMSD terminology was used in this study.*

Related references:

Entekhabi, D., et al. (2014), SMAP Handbook Soil Moisture Active Passive, Mapping Soil Moisture Freeze/Thaw From Space, 180 pp., Nat. Aeronaut. Space Admin., Jet Propul. Lab., Pasadena, Calif.

Jackson, T., M. Cosh, R. Bindlish, P. Starks, D. Bosch, M. Seyfried, D. Goodrich, S. Moran, and J. Du: Validation of Advanced Microwave Scanning Radiometer soil moisture products, IEEE Trans. Geosci. Remote Sens., 48(12), 4256–4272, doi:10.1109/TGRS.2010.2051035, 2010.

Jackson, T. J., et al.: Validation of Soil Moisture and Ocean Salinity (SMOS) soil moisture over watershed networks in the U.S., IEEE Trans. Geosci. Remote Sens., 50(5), 1530–1543, doi:10.1109/TGRS.2011.2168533, 2012.

The detailed information on the decomposition of MSEs was also added in Section 2.4.2 (see Equation (8) to (12)).

Line 196-213: *To better explain the disagreement between all the SM products and in situ observations, the mean square errors (MSEs, as defined in Eq.(8)) of each product in individual regions are utilized. To decompose the MSEs, the Nash-Sutcliffe efficiency (NSE, Nash and Sutcliffe, 1970) are utilized as defined in Eq.(9).*

$$\mathrm{MSE} = \frac{1}{n}\sum_{t=1}^{n}(x_{p,t} - x_{obs,t})^2 \tag{8}$$

$$\mathrm{NSE} = 1 - \frac{\sum_{t=1}^{n}(x_{p,t}-x_{obs,t})^2}{\sum_{t=1}^{n}(x_{obs,t}-\mu_{obs})^2} = 1 - \frac{MSE}{\sigma_{obs}^2} \tag{9}$$

NSE was decomposed as the correlation, the conditional bias, and the unconditional bias as showed in Eq.(9) (Murphy, 1988).

$$\mathrm{NSE} = \mathrm{A} - \mathrm{B} - \mathrm{C} \tag{10}$$

$$\mathrm{A} = \mathrm{R}^2$$

$$\mathrm{B} = [\mathrm{R} - (\sigma_p/\sigma_{obs})]^2$$

$$\mathrm{C} = [(\mu_p - \mu_{obs})/\sigma_{obs}]^2$$

in which R is the correlation coefficient of observations and products, σ_{obs} *and* σ_p *are the standard deviation of in situ data and all SM products. The Eq.(10) can be transformed as Eq.(11), representing the correlation, the bias and the variability.*

$$\mathrm{NSE} = 2 \cdot \alpha \cdot \mathrm{R} - \alpha^2 - \beta_n^2 \tag{11}$$

$$\alpha = \sigma_p/\sigma_{obs}$$

$$\beta = (\mu_p - \mu_{obs})/\sigma_{obs}$$

Finally, the Eq.(12) was obtained by substituting Eq.(11) into Eq.(9) as follows:

$$\mathrm{MSE} = 2\sigma_p\sigma_{obs}(1 - R) + (\sigma_p - \sigma_{obs})^2 + (\mu_p - \mu_{obs})^2 \tag{12}$$

Related references:

Murphy, A.: Skill scores based on the mean square error and their relationships to the correlation coefficient, Monthly Weather Review, 116, 2417–2424, 1988.

Nash, J.E., Sutcliffe, J.V.: River flow forecasting through. Part I. A conceptual models discussion of principles, Journal of Hydrology, 10, 282–290, 1970.

8. Fig. 2, the spatial pattern for ERA-Interim looks pretty different from that for ERA5 and others, especially across the arid northwest and regions along the coasts. Please doublecheck it, otherwise, give an explanation.

Response: We doublechecked Figure 2, and found that the figure legends of

NCEP/DOE R2 and ERA-Interim were wrong. The large difference between ERAI and ERA5 in the regions along the coasts attributes to the spatial resolution.

9. Line 195, the larger rRMSEs in the Yangtze-Huai basin may be associated with the irrigation influence on the in situ observations. However, it's hard to think of its direct links to monsoon precipitation.

Response: Thanks for your advice. In order to remove the bias error caused by the mismatch of spatial representativeness between in situ data and all SM products, the ubRMSE was introduced instead of relative RMSE to evaluate temporal dynamic variability as showed in Figure 3. The results were showed as follows:

Line 246-251: *The distribution of the ubRMSE for all stations is shown in Fig. 3 to evaluate temporal SM dynamical variability. By removing the bias, the NCEP product has the lowest ubRMSE with values between 0.01 and 0.03 m^3/m^3, indicating its better performance at capturing the temporal variation of in situ SM. Large ubRMSE are found for the ESA CCI with values large than 0.04 m^3/m^3, indicating that this remote sensing product needs to be improved at temporal variation. Spatially large ubRMSE are also found in the Yangtze-Huai region and in the south of Northeast China, which may be attributed to the high SM values. A possible explanation for poor performance in the NC region might be that this region is strongly influenced by irrigation.*

10. Fig. 4, the regionally averaged observations show higher soil moisture in NW than the other three regions. It is NOT consistent with the precipitation patterns in Fig. 1. The discrepancy should be discussed a little bit more.

Response: Generally, the northwest region was located in the semi-arid region of China, where the annual mean precipitation between 200-400 mm. The NW region selected in this study located in the east of the Northwest China, where is the transitional zone from semi-humid region to semi-arid regions. As showed in Figure 1, the NW region consists of the arid, semi-arid, and semi-humid regions. Besides, soil moisture is influenced not only by precipitation, but also by the evaporation (affected by temperature and wind speed, etc), soil types, and other related factors. The JJA soil moisture value is obviously influenced by the contribution of both precipitation and evaporation, and we have added more discussion in section 3.4.

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.*

11. 3.2.2 Seasonality, since the previous results talk about the summer (JJA) soil moisture comparisons with observations, how the seasonal soil moisture were selected in this section should be clarified further. Further, the soil moisture discussed in the manuscript focused on the top soil layer (0-10 cm), so I guess its seasonality connected closely to precipitation annual cycle. However, in Fig. 6, it looks not so, please discuss it further.

Response: Besides the JJA SM, we also calculate monthly SM from 1981-2013 during unfrozen and snow-free seasons to study the seasonal variation of SM. The above information has been clarified in the discussion section (Section 3.5).

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.*

Related references:

Jasper K, Calanca P, Fuhrer J. Changes in summertime soil water patterns in complex terrain due to climatic change. Journal of Hydrology, 2006, 327: 550-563.

Harmsen E W, Norman L M, Nicole J S, J E Gonzalez. Seasonal climate change impacts on evapotranspiration, precipitation deficit and crop yield in Puerto Rico. Agricultural Water Management, 2009, 96: 1085-1095.

12. Line 230, 'snow or frozen soil during these periods.' The frozen seasons should be excluded in the comparisons, otherwise the model soil moisture is virtually a different variable from the observed.

Response: In the former manuscript, we only discarded in situ soil moisture data during snow or frozen days. During this revision, the months with large percent of frozen and snow days were discarded for comparison. Furthermore, if the in situ observation were missing, all reanalysis data at the same period were also treated as missing value.

Line 156-159: *Considering that the field capacity and the dry bulk density are not measured at all stations, data from 119 stations are selected from 1981 to 2013. Not all in situ data were suitable for evaluation given instrumental error and observational conditions, for example, the available measurement period, installation depth and sensor placement. Therefore the evaluation was conducted in unfrozen and snow-free seasons, such as June-July-August (JJA).*

Line 227-231: *The comparisons were performed as follows: (i) make a correspondence between all soil moisture data sets and in situ SM by using the values at the nearest neighbor grids; (ii) compare all the SM products at regional scales by calculating the reginal average of monthly value of all SM products, which has been proved can reduce the uncertainty caused by grid mismatch to some extent (Nie et al., 2008); (iii) if the in situ observation were missing, all reanalysis data at the same period were also treated as missing value, which were not taking into account.*

Related references:

Nie, S., Luo, Y., Zhu, J.: Trends and scales of observed soil moisture variation in China, Advance in Atmosphere Science, 25, 43–58, 2008.

13. Line 286, 'The SC-PDSI is utilized (Wells et al., 2004).', for what is SC-PDSI used?

Response: The SC-PDSI is utilized here to define different dry/wet conditions. To be simplicity, this sentence and the following one has been combined into one sentence as:
Line 343: *Figure 10 shows the rBias under different humid/arid conditions by utilizing SC-PDSI (Wells et al., 2004)*

Related references:

Wells, N., Goddard, S., and Hayes, M. J.: A self-calibrating palmer drought severity index, J. Clim., 17, 2335–2351, https://doi.org/10.1175/1520-0442, 2004.