Dear editor,

Thank you for editing our manuscript. The first part of this document includes the point-by-point responses to the reviews (Reviewer 1, Reviewer 2, Reviewer 3). Comments of the referees are marked as e.g. << Reviewer 1 Major comment 1>> followed by the answer from the authors, which includes the changes made in the manuscript to fulfill the referees' suggestions. The section of responses to the referees is followed by a marked-up version of the manuscript.

Best regards

Yongzhe Chen, Xiaoming Feng and Bojie Fu

To Reviewer #1:

We thank referee#1 for the valuable comments that will help us in improving the quality and readably of the manuscript. We have carefully revised the MS following your comments and suggestions. We provide a detailed response to the Referee's comments in the Supplement.

Reviewer 1 General comment. The authors propose a global dataset of top (0-5cm) soil moisture with 10-day temporal and 0.1 spatial resolution, covering the period 2003-2018. The dataset was produced gradually, backward in time, through machine learning methods (neural networks) for 5 periods that correspond with the availability of 11 different passive and active satellite remote sensing soil moisture products. Besides satellite observations (starting with SMAP in step one), 9 environmental properties were fed to the neural network, and from step two on, previously modeled soil moisture was included to enable the expansion backward in time. The final product is evaluated with observations of the international soil moisture network and, in comparison to other merged products, rated superior, however, the potential for further improvement is also emphasized. Altogether, the work seems sound and the developed method and dataset appear valuable for further scientific studies and applications. Nevertheless, some of the steps in the processing chain need further clarification and the data structure needs to be improved before the manuscript can be considered to suffice for publication.

Response: Thank you for your careful reading and the positive comments on our work. We agree that some steps in the method was unclearly written, while the data structure is not easy for other researchers to use. We have added the missing important details in the revision for further clarification of the processing chain, and reuploaded the dataset with filename changed and table added, according to your valuable suggestions. Please see the details in the responses below.

Reviewer 1 Specific comment 1. I suggest to remove "new" from the title, since all dataset proposed in this journal are somewhat new. You may consider to name it "combined" or "improved" or "complete" or "optimal". Have you thought about giving the product an acronym? That improves recognizability and makes it easier to reuse it in other studies and publications.

Response: Thank you for the suggestion, we have changed it to 'improved'. We also named the product as 'RSSSM' since it is a remote sensing-based surface soil moisture. In other parts of the article, all the 'SIM' have been changed to 'RSSSM' as well, including those in figures and tables.

Reviewer 1 Specific comment 2. L14: more than 10**6 not correctly displayed in the online abstract (here it reads 106)

Response: We have revised it to '*more than one million*'.

Reviewer 1 Specific comment 3. L15: Please state also the temporal resolution (10 days)

Response: We have added this important information.

Reviewer 1 Specific comment 4. L32: resolution of ERA INTERIM is rather 0.75°

Response: Thank you for reminding us. We have corrected it accordingly.

Reviewer 1 Specific comment 5. L34-37: I agree that these products have many shortcomings, but other than the dataset provided by the authors, the models provide also information about the deeper soil layers. This important point should not be omitted here.

Response: We agree that models can simulate soil moisture at deeper layers, which is an important advantage. We have added this point in the revised manuscript: '*Apart from surface soil moisture that can be observed by satellites, the modeling way provides also the information on the moisture in deeper soil layers.*'

Reviewer 1 Specific comment 6. L75-76: "data averaging" - what type of averaging is meant, spatial or temporal? "can hardly unify the temporal variations." Please specify what the "temporal variations" refer to. Is it the temporal variations of the different soil moisture data products?

Response: Sorry for the unclear expressions. It's neither spatial nor temporal averaging. Instead, CCI product is achieved by rescaling the soil moisture data retrieved from each microwave sensor first and then averaging the rescaled soil moisture data products during the same period (i.e., the common period for two or more products) based on some criteria (e.g. the estimated error) (Dorigo et al., 2017; Gruber et al., 2017; Gruber et al., 2019; Liu et al., 2012). The ‘temporal variations’ in this sentence refers to the temporal variations of the different soil moisture data products. Following your question, we have revised the sentence to: ‘*Rescaling the soil moisture data retrieved from each sensor by using CDF matching followed by averaging the rescaled data during one common period, which is adopted in CCI, can hardly unify the temporal variations of different soil moisture products.*’.

Reviewer 1 Specific comment 7. Instead of Table 1 or in addition, it would be good to have a timeline figure from 2003-2018 that shows a bar for every dataset used in the process of creating the final product, including the 11 soil moisture products, the time-varying quality impact factors and the intermediate modeling products (SIM-1T, SIM-2T,...).

Response: Thank you for this nice suggestion. Following your comment, we have added a timeline figure showing the temporal coverages (including used data periods and unused data periods) of all 11 soil moisture products, the time-varying quality impact factor (i.e. three LAI products), and that of the intermediate products. The figure below is attached as Figure 1 in the revised manuscript. Table 1 is thus removed.

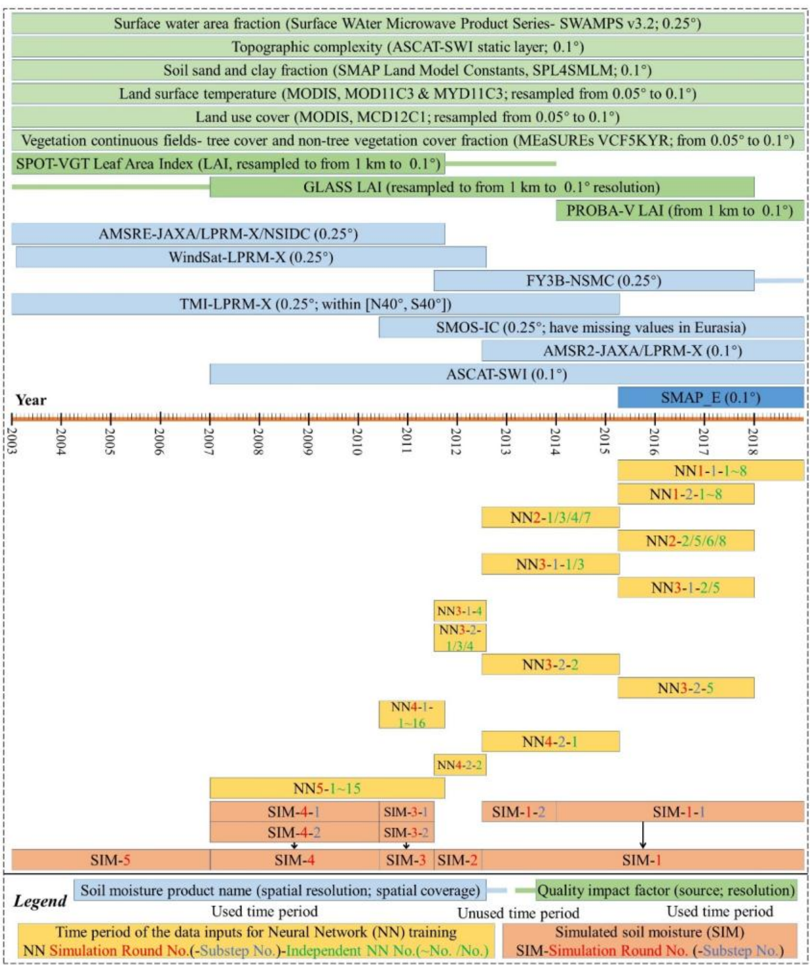


Figure R1: The timeline figure showing the time periods of the soil moisture datasets and the 'quality impact factor' products (e.g. LAI dataset) used in this study (listed above the timeline), as well as the periods of data applied for the trainings of 67 independent neural networks and the neural network simulation outputs (i.e. simulated soil moisture) of eight substeps (listed

带格式的: 居中

below the timeline).

Reviewer 1 Specific comment 8. L110: Specify why SMAP is mentioned as the "best product" here. Is it because of the spatial resolution, the algorithms or with respect to the in-situ observations? Can you add a citation to corroborate this statement?

Response: SMAP is the 'highest quality product', with respect to in-situ observations. In the Introduction part, this point has been stated: '*Although new sensors such as SMOS (Stillman and Zeng, 2018) and SMAP (Entekhabi et al., 2010), can produce significantly improved estimates because L-band microwaves (1~2 GHz) can better penetrate the vegetation canopy (Burgin et al., 2017; Chen et al., 2018; Karthikeyan et al., 2017; Kerr et al., 2016; Kim et al., 2018; Leroux et al., 2014; Stillman and Zeng, 2018), the applicability of both products is still limited. SMOS data have too much noise and too many missing values in Eurasia due to high radio frequency interference (RFI) (Oliva et al., 2012). While SMAP has the highest quality (the unbiased RMSE of the passive product can be close to its target of 0.04 m^3/m^3) and has filtered RFI (Chen et al., 2018; Colliander et al., 2017), …*' Following your advice, we revised this sentence and added two new references supporting the best performance of SMAP product. It reads: '*SMAP has currently the highest quality of all remote sensing-based soil moisture products (Al-Yaari et al., 2019; Liu et al., 2019)…*'.

Reviewer 1 Specific comment 9. L143: change to "reference coordinate system".

Response: We have changed it accordingly.

Reviewer 1 Specific comment 10. L182: "based on the correlation between soil dielectric conductivity" - do you mean soil dielectric permittivity or soil electric conductivity?

Response: Sorry for that mistake. It should be 'soil dielectric permittivity', or 'soil dielectric constant'. We have corrected it to '*soil dielectric constant*' in the revision.

Reviewer 1 Specific comment 11. L186-188: "Because ..." this sentence is unclear.

Response: Sorry for the unclear expression. This sentence is to explain that actual LST can determine the bias of every LST estimate, which is used in the corresponding soil moisture retrieval. Hence, the actual LST will influence the biases of different soil moisture products. We have revised the sentence as: '*Because in the retrievals of different soil moisture products, different LST estimates are used, while the bias of each LST estimate compared to the actual LST is influenced by actual LST, we suppose that the actual LST can determine the accuracy of every LST estimate, and finally the relative performances of various soil moisture products (Kim et al., 2015)*'. We hope it will be easier for readers to understand.

Reviewer 1 Specific comment 12. L205: Figure 1 is never referenced in the manuscript. This should be done here or later at L225.

Response: Thanks for reminding. We have added: *'The basic flow is shown in Figure 2.'* (note: Figure 2 is Figure 1 in the original manuscript) at the beginning of the section 2.2 in the revised manuscript.

Reviewer 1 Specific comment 13. L219: Do the 140x360 zones include water (ocean) areas?

Response: Yes, the 140×360 zones include water (ocean) areas. However, for zones with no land or very limited land, the number of valid pixels is lower than 100, so these ocean zones are not applicable for subnetworks, and are excluded.

Reviewer 1 Specific comment 14. L220: A subnetwork has 100 pixels, but ("for a 0.1pixel in a given 10-day period, if all the subnetwork inputs have valid..."), how can one pixel have more subnetworks? Please improve the formulation.

Response: Sorry for the confusing expression. We have rewritten the paragraph as: '*Therefore, we divided the global extent except the polar areas (80°N~60°S) into 140×360 zones. Here, for a 0.1° pixel during a specific 10-day period, if all the input data (soil moisture products and quality impact factors) have valid values, it can provide one valid data point. So, the maximal number of valid data points applied to train a subnetwork = 100 × the number of 10-day periods within the training period. The subnetworks with valid data points less than 100 (e.g. those in oceans) were dropped, leaving usually >15,000* *subnetworks included in an independent neural network*.'

Reviewer 1 Specific comment 15. L222: What is an "individual neural network"? Is it the collective of all zonal neural networks for one simulation (SIM-T1, SIM-T2, ...)? Is the maximum possible number of subnetworks 50.400 or less because of ocean cells?

Response: We have changed it to '*independent neural network*' to make it consistent with the expression in the abstract. An independent neural network is the collective of all zonal subnetworks. Several independent neural networks constitute a simulation substep (for example, NN-1-1, NN-1-2, …, NN-1-8 are applied in Round 1- Substep 1), while each substep is responsible for one simulation (there are eight simulations: SIM-1-1, SIM-1-2, SIM-2, SIM-3-1, SIM-3-2, SIM-4-1, SIM-4-2 and SIM-5, for example, SIM-1-1 is the output of Round 1- Substep 1).

The number of subnetworks in each independent neural network is far below 50400, not only because of ocean zones, but also because some soil moisture data is only available in a region (e.g. TMI is available within [-40°S~40°N]). The paragraph is revised as in the response to Reviewer 1 Specific comment 14.

(**Please also note that SIM-1T, SIM-2T, …, SIM-4T are only the postprocessing results that are intended to be used as secondary training targets, while SIM-1, SIM-2, …, SIM-5 are combined to constitute our soil moisture products.)

Reviewer 1 Specific comment 16. L223: For reproducibility, it is required to state exactly the MATLAB version and the toolbox version and method/function name that was used for training the neural network.

Response: We have added the MATLAB version 2016a accordingly.

Reviewer 1 Specific comment 17. L256: "we classified all pixels" -> "we classified all 0.1 pixels", I suggest to add the resolution information that it is clear which of the different grids is addressed.

Response: We have revised it accordingly.

Reviewer 1 Specific comment 18. L259: Again, I thought that a pixel is the smallest unit in the process (i.e. subnetwork). So how can a pixel have a subnetwork? Not clear to me.

Response: We are sorry for the unclear description. Actually, the subnetwork belongs to a 1° ×1° zone, not a pixel. We have revised it as: '*For data simulation in a 0.1° pixel, the most preferable independent neural network is expected to be trained using all the available soil moisture data sources in that pixel. However, in the 1° zone where it is located, the subnetwork belonging to that preferable independent neural network may not exist due to limited valid data points (see section 2.2.1). Then, an alternative subnetwork driven by the combination of fewer soil moisture data inputs should be applied instead.*'

Reviewer 1 Specific comment 19. L261-262: "Hence, it is a ..." sentence seems incorrect. I think you should better write "neural network collocation" or "neural network constitution" to make it more clear that these are neural network realizations with identical configuration but different ingredients.

Response: We have changed it to 'neural network collocation' accordingly.

Reviewer 1 Specific comment 20. L272,815 and other occurrences: it is not clear how the 10 day periods are defined and how they relate to the ordinal numbering. A month has between 29 and 31 days, so how are the periods split and how does that affect the last 3rd where the number of days is variable? How does this variable length averaging affect the results and what are the implications for validation?

Response: The first and second 10-day periods in a month both contain exactly 10 days, but the last 10-day period has variable number of days (9(8)~11). This, however, may not have substantial effect on our results and data quality. This is because it takes at least three days for a microwave sensor to cover the globe. Also, for each grid, the days with observations are not the same among different sensors. Therefore, this study only took the average of the available soil moisture retrievals during a 10-day period (we have added a paragraph in section 2.1.1 as: '*To reduce noises and fill the gaps between sensor observing tracks (it takes at least 3 days for a microwave sensor to cover the whole globe), for every soil moisture product, both the daytime and nighttime observations within each 10-day period are combined by data averaging (the relative superiority of* *daytime and nighttime retrievals is not considered). For example, for SMAP, 11% of global land surface has data for only 5 days or less within a 10-day period.*'). Moreover, surface soil moisture may vary significantly even in a day, due to rainfall events, but the observations are transient. Therefore, either the 10-day averaged microwave soil moisture products or the simulated soil moisture data in this study can only roughly indicate the overall soil moisture condition, not exactly equals to the mean soil moisture during 10 complete days. Hence, it doesn't matter whether the 'last 10-day period' in a month has exactly 10 days or not. In fact, this data format is exactly the same as the ASCAT-SWI soil moisture and many other products (e.g. LAI) developed by the ESA- Copernicus Land Monitoring Service (https://land.copernicus.eu). For the validation process based on ISMN measurements, the mean in-situ soil moisture in the 'last 10-day period' of a month was also calculated by averaging the records in either 10 days or 11 days (or 8~9 days in February), which was consistent with the 'nominal' simulation period. Following this comment, we added this information: '*The temporal resolution is approximately 10 days, or to be specific, there are 3 data records within a month, for days 1~10, 11~20 and from 21 to the last day of that month, respectively.*' to the abstract in the revision.

Reviewer 1 Specific comment 21. L270-292: also this section would greatly benefit from a timeline bar plot that shows all the soil moisture products and simulated models, so that the overlaps can be grasped Immediately.

Response: We have added Figure 1 which shows the timelines of the simulated soil moisture corresponding to 8 substeps and the periods of data inputs for the trainings of 67 independent neural networks. Please find the detail in response to Reviewer 1 Specific comment 7. We also added a sentence: '*The training period of each neural network and the simulation period of each substep are shown in Figure 1.*' in this paragraph.

Reviewer 1 Specific comment 22. L318: define how R**2 is computed (based on Spearman or Pearson).

Response: The R^2 is computed based on Pearson, we have added this information in the revision.

Reviewer 1 Specific comment 23. L321: lower case r should be used for the correlation coefficient (based on Pearson?). Why are you mixing r and R^2 and do not use R^2 for all analyses?

Response: Following this comment, we have changed R to '*r*' to represent the Pearson correlation coefficient, including those in figures and tables. In this way, we can better distinguish the correlation coefficient from R^2.

To evaluate the overall performance, we showed the scatter plot between the simulated soil moisture and the measured values. Here, instead of *r*, we used R^2 better reveal the differences among the performances of different soil moisture products. However, in the temporal and spatial validation, at some sites or during some specific 10-day periods within a climatic region, the simulations and measurements can be negatively correlated, as shown in the figures within manuscript, which are actually of very low quality. But if we only use R^2, these low-quality data will be overshadowed (for example, if the *r* is -0.6, the R^2 can be as high as 0.36). Therefore, it's wiser to use temporal correlation and spatial correlation. Previous studies also used '*r*' to evaluate the spatial and temporal accuracy of surface soil moisture products against ISMN measurements, for example (Karthikeyan et al., 2017).

Reviewer 1 Specific comment 24. L322: please provide formula for A.R computation

Response: 'A.R' in this study stands for the correlation coefficient (*r*) between the anomalies of simulated soil moisture and the anomalies of measured soil moisture at a specific ISMN station. Following this comment, we have added the equation blow to show how the anomalies of simulated or measured surface soil moisture were calculated in the revised manuscript.

$$\overline{SSM(k)} = \frac{\sum_{y=1}^{ny} SSM(y,k)}{ny} \quad (ny \geq 3)\ ;\ SSM\ is\ either\ estimated\ or\ measured$$

$$SSM: surface\ soil\ moisture; k: the\ ordinal\ of\ 10\ day\ period\ in\ a\ year;$$

$$y: a\ year\ with\ measured\ SSM\ in\ the\ k^{th}\ 10\ day\ period;\ ny: number\ of\ those\ years$$

$$SSM_{anom}(y,k) = SSM(y,k) - \overline{SSM(t)}$$

$$SSM_{anom}(y,t): the\ anomalies\ of\ surface\ soil\ moisture\ during\ the\ \ t^{th}\ 10\ day\ period\ in\ year\ y.$$

Reviewer 1 Specific comment 25. L326: "in all grids", grids or pixels (1 x 1 or 0.1 x 0.1)?

Response: Thanks for careful reading. We revised it as '*in all 0.1° grids*'.

Reviewer 1 Specific comment 26. L326: please provide formulas for spatial pattern validation (at least in the supplement)

Response: We have provided more details for spatial pattern validation. Now it reads: '*Finally, we performed spatial pattern validation. In detail, for every 10-day period, we compared all the soil moisture measurements that were upscaled to 0.1° during that period with the corresponding estimated values. The spatial pattern evaluation indexes include correlation coefficient (r), RMSE, bias and ubRMSE values (Eq. 2).*'

$$\overline{SSM_{est}} = \frac{\sum_{i=1}^{ng} SSM_{est,i}}{ng} \; ; \quad \overline{SSM_{act}} = \frac{\sum_{i=1}^{ng} SSM_{act,i}}{ng} \; (ng \geq 20)$$

$$i\text{: } a\ grid\ with\ upscaled\ surface\ soil\ moisture\ measurements\ during\ a\ specific\ 10\ day\ period;$$

$$ng\text{: } the\ number\ of\ those\ grids\ in\ the\ globe$$

$$ubRMSE_{spatial} = \sqrt{\sum_{i=1}^{ng} [\left(SSM_{est,i} - \overline{SSM_{est}}\right) - \left(SSM_{est,i} - \overline{SSM_{act}}\right)]^2 / ng} \qquad \text{(Eq. 2)}$$

Reviewer 1 Specific comment 27. Figure 3: Use identical labels for the x-axis, add missing lower frame.

Response: We uniformed the labels for the x-axis. The figures have been adjusted accordingly. The revised Figure 3 (Figure 5 in the revised manuscript) is shown below:

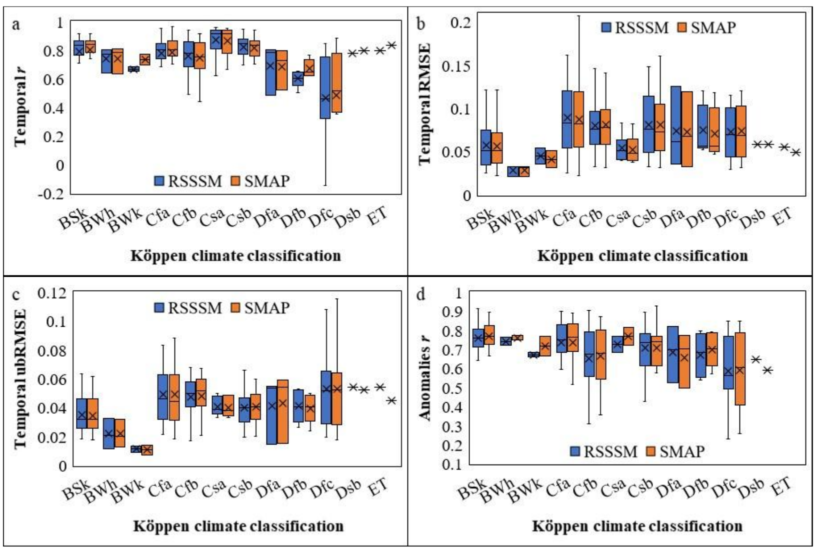


Figure R2: Comparison between the temporal accuracy of RSSSM and SMAP in regions with different Köppen-Geiger climate types. The four indexes are (a) r, (b) RMSE, (c) ubRMSE and (d) Anomalies r (A.R). The lengths of the error bars are 1.5 times that of the interquartile range, while the upper and lower boundaries and the central lines of the boxes indicate the 75th, 50th and 25th percentile values, with mean values marked by '×' (the forms of all the following boxplots are the same).

Reviewer 1 Specific comment 28. Figure 4: If the color key is put below the figure, the figure can be increased in the horizontal direction which leads to wider bars. You could even remove the x-axis labels and names and leave only the lowermost. By this you can increase the size of the bars and hence the readability (reduce redundancy).

Response: We moved the color key below the figure, and removed the x-axis names to increase the size of the bars (shown in Figure R3, Figure 6 in the revised manuscript). Other figures were also revised to be larger and clearer.

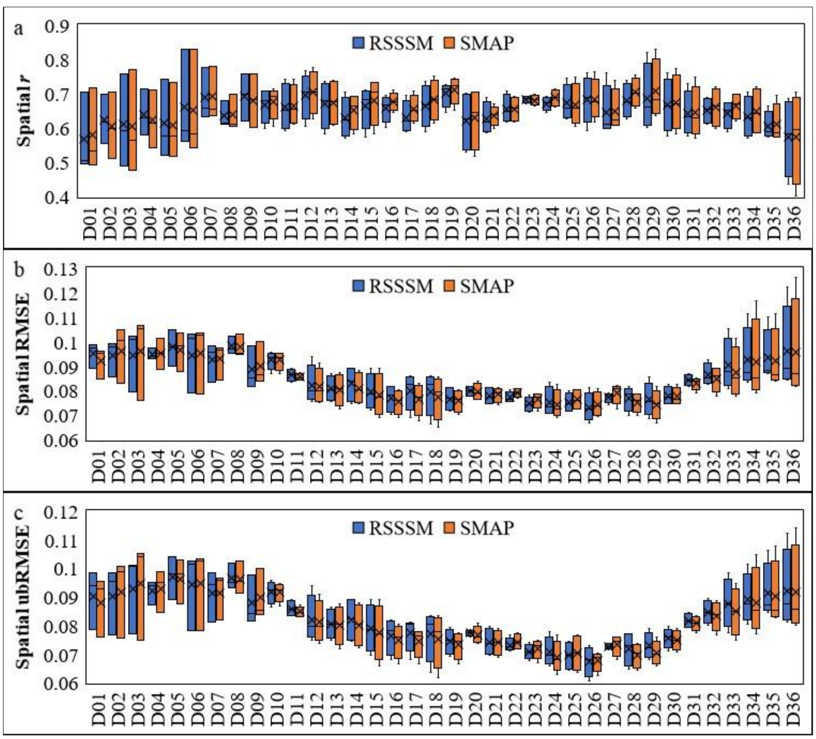


Figure R3: Comparison between the spatial pattern accuracy of RSSSM and SMAP in different 10-day periods during April 2015~2018. The three evaluation indexes are: (a) r, (b) RMSE and (c) ubRMSE. The length of each box/error bar is determined from the evaluation index values in three (January to March) or four (April to December) years.

Reviewer 1 Specific comment 29. L381: How is the performance of SIM if the SMAP training period is omitted, i.e. from 2003 until 2015D01, as compared to ASCAT-SWI?

Response: We have added this information in the manuscript accordingly. It reads as: '*If the data period of SMAP (2015D10~2018) is excluded, the overall R^2 and RMSE for RSSSM are 0.43 and 0.087, still better than ASCAT-SWI (R^2=0.33,* *RMSE=0.1).*'

Reviewer 1 Specific comment 30. Figure 7,10: As for fig. 4 place color key below the plots and increase the bars horizontally.

Response: We have revised all the figures accordingly.

Reviewer 1 Specific comment 31. Do you see any chance to improve the temporal resolution of the product in the future? If not, what are the constraints?

Response: Currently, it's probably not a good choice to further increase the temporal resolution of the long-term microwave surface soil moisture product. As illustrated in the response to Reviewer 1 Specific comment 20, for each grid, both the days and the hours of observations by different microwave sensors differ from each other. Because surface soil moisture has high variability in a short time period (even in a day) due to rainfall events, the actual soil moisture at the passing time of various sensors are not the same. However, this study used multiple sources of microwave surface soil moisture products as predictors of neural networks and SMAP soil moisture data as the training target, meaning that the neural network training should be based on the assumption that the products retrieved by different sensors contain exactly the same actual soil moisture information. As we can see, there exists a conflict. To solve it, for each sensor, we took the average of its available retrievals during a certain time period, and the larger number of data applied for averaging is, the better can the result represents the mean soil moisture during that period. Because 11% of global land only have 5 or less days with observations during a 10-day period, if the temporal resolution is improved, for example, to 5 days, there may be only 2~3 observations available. Considering the high temporal variability of surface moisture, the average of those limited data can hardly indicate the average soil moisture condition. This will lead to large uncertainties in the neural network training, and finally, our soil moisture simulation results.

This problem is an inherent constraint of microwave remote sensing data integration. So, to improve the temporal resolution, other data sources need to be incorporated. As we know, soil moisture retrieved from other remote sensing techniques (e.g. optical) have low quality over vegetated areas, and are heavily interfered by clouds. Therefore, model simulation may be the only solution to this problem. For example, assimilating the observational-based surface soil moisture into models such as GLEAM can achieve surface/root-zone soil moisture mapping at daily scale. Therefore, we have revised the last paragraph as: '*Another way to improve global surface soil moisture data accuracy as well as the temporal resolution is to combine satellite-based products with land surface models such as GLEAM. Remote sensing inversion can delineate more detailed spatial information on soil moisture, whereas the reanalysis-based models have advantages in characterizing temporal variations, and even on daily scale, except for…*'.

Reviewer 1 Specific comment 32. L499-500: Is SIM also superior to the other products if only the prior to SMAP period is considered (2003 until 2015D01)?

Response: Thank you for this advice. We have added the following comparisons to the manuscript.

1) if the data period of SMAP (2015D10~2018) is excluded, the overall R^2 and RMSE for RSSSM are 0.43 and 0.087, still better than ASCAT-SWI (R^2=0.33, RMSE=0.1).

2) when excluding the SMAP (training target) data period, the R^2 and RMSE for RSSSM are 0.41 and 0.089, also superior to those for GLDAS (R^2: 0.37; RMSE: 0.099).

3) without considering the SMAP period, the condition is the same (R^2 for RSSSM and ERA5-Land are 0.41 and 0.38; RMSE for these two products are 0.089 and 0.125).

4) when the SMAP data period is excluded, the R^2 and RMSE for CCI are 0.028 and 0.098, compared to 0.41 and 0.089 for RSSSM.

5) if the SMAP data period is excluded, RSSSM's R^2 and RMSE are 0.41 and 0.089, still better than both GLEAM v3.3a (R^2: 0.35; RMSE: 0.141) and GLEAM v3.3a (R^2: 0.34; RMSE: 0.128.

So, the comparisons above can prove that our product (RSSSM) is superior to the other products even if the SMAP period is excluded.

Reviewer 1 Specific comment 33. Are there plans to update the data-set on a regular basis?

Response: Yes. We plan to update the whole dataset when more advanced microwave sensors (e.g. P-band sensors) are launched and the global-scale higher-quality surface soil moisture data is available in future. We have added this promise in the discussion. Now it reads: '*Therefore, if microwave sensors with higher SNR or better penetration of vegetation canopy than SMAP are launched in future (for example, the upcoming P-band microwave sensors (Etminan et al., 2020; Ye et al., 2020)), by using the soil moisture or Tb retrieved from the new sensors as the reference, we can develop a temporally continuous soil moisture dataset since 2003, which is expected to have even higher accuracy than the SMAP product (we will update the complete RSSSM product then). In that sense, the data fusion algorithm proposed here will be very meaningful in future*.'

Reviewer 1 Specific comment 34. The dataset is organized as an archive of geotiff files. The problem with this structure is that the time identifier is only contained in the file name, but without practical formatting. If one wants to import a time series for a region or a single pixel, the data structure is quite unhandy. Also from the readme file and the metadata it is not quite clear what the 10 days ordinal numbering means exactly. Is it always the [1-9],[10-19],[20-29] or [1-10],[11-20],[21-30] periods? How are the months with variable length considered (28,29,30,31 days)? That's not clear also not from the manuscript. Further, I would suggest to add a table (csv) that links the different file names to their

specific period using ISO 8601 https://en.wikipedia.org/wiki/ISO_8601 notation:

e.g., a file named inventory.dat with a list like the following one:

Period, Filename 2003-01-01/2003-01-10, SMY2003DECA01.tif 2003-01-11/2003-01-20, SMY2003DECA02.tif ...

Response: Thank you for the suggestion on the naming and structure of our data. Actually, it is [1-10], [11-20], [21, the end of each month]. We have added a csv table named 'filename' linking the different file names to their specific period, following your instructions.

Reviewer 1 Specific comment 35. Also the numbering should be formatted as %02d so that, e.g., SMY2003DECA1.tif becomes SMY2003DECA01.tif. This is important if one wants create a chronological file list for looping over time. With the current scheme, the order would become SMY2003DECA1.tif SMY2003DECA10.tif, SMY2003DECA11.tif, ... This should be also applied to all tables in the manuscript (e.g., 2005D01 instead of 2005D1).

Response: We have changed the naming of the product as well as the abbreviations for each 10-day period, both in the manuscript and in the Supplement.

Reviewer 1 Specific comment 36. Figure S1: The figure and description is not completely clear. I assume that every number (yellow and blue frames) is one pixel (0.1x 0.1)? I think it would become more clear if you superimposed a light gray mesh for the pixels over the 1x1zones. But then, why are there 4 steps required to smooth the borders? It means that every boarder gets smoothed twice, and every corner point even four times.

Response: Sorry for the unclear description. The figure has been revised (see Figure R4), with light gray mesh superimposed for all the pixels over one 1°×1° zone. The four steps were used to process respectively the four borders of each 1°×1° zone (please note that in each step, only the border colored in blue is smoothed). We have added more details to make it clearer, which now reads as: '*The sketch of the four substeps in boundary fuzzification. The 1°×1° zones are separated by black solid lines (the 0.1°×0.1° pixels in one zone are superimposed by light grey mesh). For each substep (subfigures a~d), the soil moisture value within each pixel that is colored in blue are recalculated as the average of its original surface soil moisture and the original soil moisture value in its most adjacent yellow color pixel, weighted by the corresponding numbers labelled (i.e. 2 and 1). In this way, every border of a 1°×1° zone gets smoothed once (substeps 'a~d' are for four borders, respectively, where a~b are for the horizontal borders while c~d are for the vertical borders), but the four corners get smoothed twice (both horizontally and vertically).*'

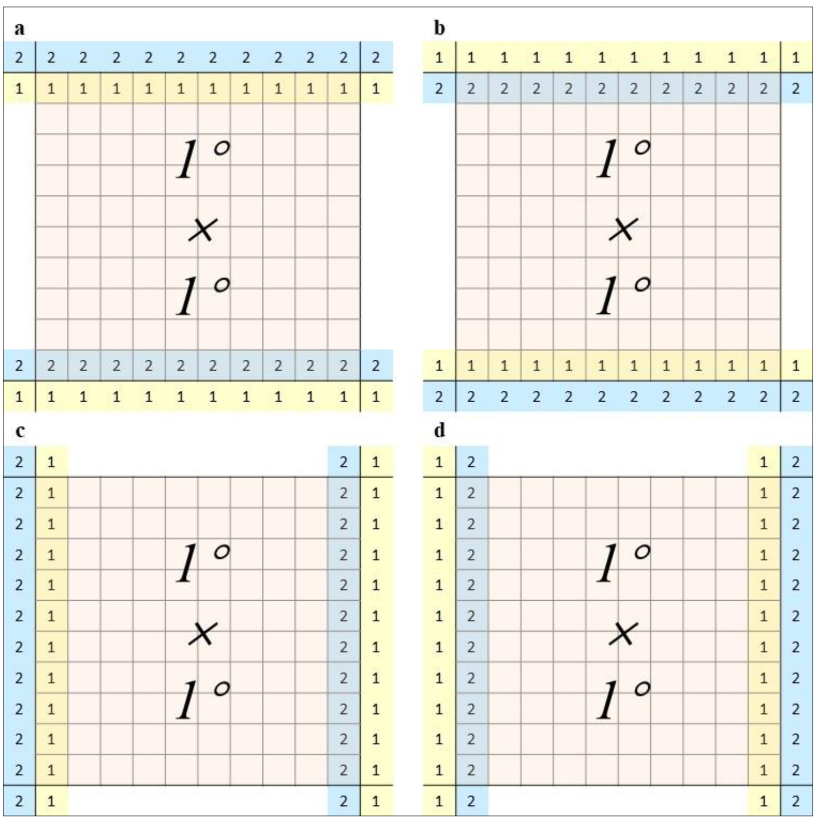


Figure R4. The sketch of the four substeps in boundary fuzzification. The 1°×1° zones are separated by black solid lines (the 0.1°×0.1° pixels in one zone are superimposed by light grey mesh). For each substep (subfigures a~d), the soil moisture value within each pixel that is colored in blue are recalculated as the average of its original surface soil moisture and the original soil moisture value in its most adjacent yellow color pixel, weighted by the corresponding numbers labelled (i.e. 2 and 1). In this way, every border of a 1°×1° zone gets smoothed once (substeps 'a~d' are for four borders, respectively, where a~b are for the horizontal borders while c~d are for the vertical borders), but the four corners get smoothed twice (both horizontally *and vertically).*

Reviewer 1 Specific comment 37. Figures S5, S8, S11: put the color-key to the bottom of the figure (a single key would be sufficient for all sub-figures), you could even remove the x-axis labels and names and leave only the lowermost. By this you can increase the size of the bars and hence the readability (reduce redundancy).

Response: We have revised these figures, making the size of the bars much larger now by adjusting the place of the color keys and removing the x-axis names.

Reviewer 1 Specific comment 38. There are often blanks missing between words. DOIs are completely missing in the reference list.

Response: Sorry for those mistakes. We have added the missing blanks and rewritten the reference list (add DOIs, abbreviate the journal names, correct wrong references) to make sure it meets the format requirement of ESSD. Thank you again for your careful reading and valuable suggestions.

References

Al-Yaari, A., Wigneron, J. P., Dorigo, W., Colliander, A., Pellarin, T., Hahn, S., Mialon, A., Richaume, P., Fernandez-Moran, R., Fan, L., Kerr, Y. H., and De Lannoy, G.: Assessment and inter-comparison of recently developed/reprocessed microwave satellite soil moisture products using ISMN ground-based measurements, Remote Sens. Environ., 224, 289-303, https://doi.org/10.1016/j.rse.2019.02.008, 2019

Burgin, M., Colliander, A., Njoku, E., Chan, S., Cabot, F., Kerr, H. Y., Bindlish, R., Jackson, T., Entekhabi, D., and Yueh, S.: A Comparative Study of the SMAP Passive Soil Moisture Product With Existing Satellite-Based Soil Moisture Products, IEEE Trans. Geosci. Remote Sensing, 55, 2959-2971, http://doi.org/10.1109/TGRS.2017.2656859, 2017

Chen, F., Crow, W. T., Bindlish, R., Colliander, A., Burgin, M. S., Asanuma, J., and Aida, K.: Global-scale evaluation of SMAP, SMOS and ASCAT soil moisture products using triple collocation, Remote Sens. Environ., 214, 1-13, https://doi.org/10.1016/j.rse.2018.05.008, 2018

Colliander, A., Jackson, T. J., Bindlish, R., Chan, S., Das, N., Kim, S. B., Cosh, M. H., Dunbar, R. S., Dang, L., Pashaian, L., Asanuma, J., Aida, K., Berg, A., Rowlandson, T., Bosch, D., Caldwell, T., Caylor, K., Goodrich, D., al Jassar, H., Lopez-Baeza, E., Martínez-Fernández, J., González-Zamora, A., Livingston, S., McNairn, H., Pacheco, A., Moghaddam, M., Montzka, C., Notarnicola, C., Niedrist, G., Pellarin, T., Prueger, J., Pulliainen, J., Rautiainen, K., Ramos, J., Seyfried, M., Starks, P., Su, Z., Zeng, Y., van der Velde, R., Thibeault, M., Dorigo, W., Vreugdenhil, M., Walker, J. P., Wu, X., Monerris, A., O'Neill, P. E., Entekhabi, D., Njoku, E. G., and Yueh, S.: Validation of SMAP surface soil moisture products with core validation sites, Remote Sens. Environ., 191, 215-231, https://doi.org/10.1016/j.rse.2017.01.021, 2017

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

Entekhabi, D., Njoku, E., O'neill, P., Kellogg, K. H., Crow, W., Edelstein, W. N., Entin, J. K., Goodman, S. D., Jackson, T., Johnson, F. M., Kimball, J., Piepmeier, J., Koster, R. D., Martin, E., McDonald, C. K., Moghaddam, M., Moran, M. S., Reichle, R., Shi, J. C., Spencer, D., Thurman, S. W., Tsang, L., and Zyl, J. V.: The Soil Moisture Active Passive (SMAP) Mission, Proc. IEEE, 98, 704-716, https://doi.org/10.1109/JPROC.2010.2043918, 2010

Etminan, A., Tabatabaeenejad, A., and Moghaddam, M.: Retrieving Root-Zone Soil Moisture Profile From P-Band Radar via Hybrid Global and Local Optimization, IEEE Trans. Geosci. Remote Sensing, doi: https://doi.org/10.1109/TGRS.2020.2965569, 2020. 1-9, https://doi.org/10.1109/TGRS.2020.2965569, 2020

Gruber, A., Dorigo, W. A., Crow, W., and Wagner, W.: Triple Collocation-Based Merging of Satellite Soil Moisture Retrievals, IEEE Trans. Geosci. Remote Sensing, 55, 6780-6792, https://doi.org/10.1109/TGRS.2017.2734070, 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

Karthikeyan, L., Pan, M., Wanders, N., Kumar, D. N., and Wood, E. F.: Four decades of microwave satellite soil moisture observations: Part 2. Product validation and inter-satellite comparisons, Adv. Water Resour., 109, 236-252, https://doi.org/10.1016/j.advwatres.2017.09.010, 2017

Kerr, Y. H., Al-Yaari, A., Rodriguez-Fernandez, N., Parrens, M., Molero, B., Leroux, D., Bircher, S., Mahmoodi, A., Mialon, A., Richaume, P., Delwart, S., Al Bitar, A., Pellarin, T., Bindlish, R., Jackson, T. J., Rüdiger, C., Waldteufel, P., Mecklenburg, S., and Wigneron, J. P.: Overview of SMOS performance in terms of global soil moisture monitoring after six years in operation, Remote Sens. Environ., 180, 40-63, https://doi.org/10.1016/j.rse.2016.02.042, 2016

Kim, H., Parinussa, R., Konings, A. G., Wagner, W., Cosh, M. H., Lakshmi, V., Zohaib, M., and Choi, M.: Global-scale assessment and combination of SMAP with ASCAT (active) and AMSR2 (passive) soil moisture products, Remote Sens. Environ., 204, 260-275, https://doi.org/10.1016/j.rse.2017.10.026, 2018

Kim, S., Liu, Y. Y., Johnson, F. M., Parinussa, R. M., and Sharma, A.: A global comparison of alternate AMSR2 soil moisture products: Why do they differ?, Remote Sens. Environ., 161, 43-62, https://doi.org/10.1016/j.rse.2015.02.002, 2015

Leroux, D. J., Kerr, Y. H., Bitar, A. A., Bindlish, R., Jackson, T. J., Berthelot, B., and Portet, G.: Comparison Between SMOS, VUA, ASCAT, and ECMWF Soil Moisture Products Over Four Watersheds in U.S, IEEE Trans. Geosci. Remote Sensing, 52, 1562-1571, https://doi.org/10.1109/TGRS.2013.2252468, 2014

Liu, J., Chai, L., Lu, Z., Liu, S., Qu, Y., Geng, D., Song, Y., Guan, Y., Guo, Z., Wang, J., and Zhu, Z.: Evaluation of SMAP, SMOS-IC, FY3B, JAXA, and LPRM Soil Moisture Products over the Qinghai-Tibet Plateau and Its Surrounding Areas, Remote Sens., 11, http://doi.org/10.3390/rs11070792, 2019

Liu, Y. Y., Dorigo, W. A., Parinussa, R. M., de Jeu, R. A. M., Wagner, W., McCabe, M. F., Evans, J. P., and van Dijk, A. I. J. M.: Trend-preserving blending of passive and active microwave soil moisture retrievals, Remote Sens. Environ., 123, 280-297, https://doi.org/10.1016/j.rse.2012.03.014, 2012

Oliva, R., Daganzo, E., Kerr, Y. H., Mecklenburg, S., Nieto, S., Richaume, P., and Gruhier, C.: SMOS Radio Frequency Interference Scenario: Status and Actions Taken to Improve the RFI Environment in the 1400–1427-MHz Passive Band, IEEE Trans. Geosci. Remote Sensing, 50, 1427-1439, https://doi.org/10.1109/TGRS.2012.2182775, 2012

Stillman, S. and Zeng, X.: Evaluation of SMAP Soil Moisture Relative to Five Other Satellite Products Using the Climate Reference Network Measurements Over USA, IEEE Trans. Geosci. Remote Sensing, 56, 6296-6305, https://doi.org/10.1109/TGRS.2018.2835316, 2018

Ye, N., Walker, J. P., Yeo, I., Jackson, T. J., Kerr, Y., Kim, E., McGrath, A., PopStefanija, I., Goodberlet, M., and Hills, J.: Toward P-Band Passive Microwave Sensing of Soil Moisture, IEEE Geosci. Remote Sens. Lett., doi: https://doi.org/10.1109/LGRS.2020.2976204, 2020. 1-5, https://doi.org/10.1109/LGRS.2020.2976204, 2020

To Reviewer #2:

We thank referee#2 for the valuable comments that will help us in improving the quality and readably of the manuscript. We have carefully revised the MS following your comments and suggestions. We provide a detailed response to the Referee's comments in the Supplement.

****

Reviewer 2 General comment. The authors tried to generate long-term surface soil moisture at a global scale, via data fusion of 11 microwave remote sensing-based soil moisture products since 2003 through neural network approach, and SMAP soil moisture products were used as the training target. The idea is very interesting and should be encouraged to explore further how much extent the machine learning can help in Earth Observation for delivering physically-consistent (or physic-aware) products. However, the way the current manuscript is written, organized is still far from clarity, structured for this reviewer to comprehend their contributions. I would suggest rejection and encourage the author to continue along this line of effort.

Response: We thank the reviewer for the positive comment on the idea of generating long-term surface soil moisture at a global scale in this study. We agree that the organization of this manuscript, especially the explanation of the method is not clear, and sorry for some confusing or wrong terminologies. These problems will make the readers and reviewers misunderstand several rather complex algorithms, which however, are also the major innovation points of this study, including the selection of quality impact factors as neural networks inputs and the design of five rounds of simulations (the organization structure of 67 independent neural networks). We have carefully followed your advice, revising the explanations of the key methods. In addition, for each of your doubt or query, we explained our design of the method in greater detail, and have clarified the related sentences in the manuscript, hoping that our real contributions could be comprehended by you and other readers. Please find the details in the response to each of the following comments.

Reviewer 2 Major comment 1. The author claimed that "This new dataset, once validated against the International Soil Moisture Network (ISMN) records, is supposed to be superior to the existing products (ASCAT-SWI, GLDAS Noah, ERA5-

Land, CCI/ECV and GLEAM), and is applicable to studying both the spatial and temporal patterns. " This assumption is too strong. On the other hand, it seems the author referred to the validation of the NN-based 10-d soil moisture products versus the 10-d averaged ISMN in-situ observations (as seen in Figure 5, Figure 8, Figure S3, S6, and S9). Is it true? In any case, it should be specified under what conditions the generated product is performing better than other products. "supposed to be superior" is really not a scientific statement.

Response: Thank you for this comment. In this study, for our product (named RSSSM hereinafter) and each existing product, by referring to all valid ISMN sites' surface soil moisture measurements, we carefully conducted overall validation (evaluation indexes are overall R^2 and RMSE values), temporal variation validation (evaluated by temporal correlation coefficient, temporal RMSE and unbiased RMSE, etc.) and spatial pattern validation (evaluated by spatial correlation coefficient, spatial RMSE and unbiased RMSE, etc.). Please see Method section 2.3 in the revised manuscript for details, while the accuracy comparison among all products are in Result section 3.2. The validation results indicate that our RSSSM product is more comparable to the site measurements, both in terms of R^2 and RMSE (see Figure 7, Figure 10, Figure S3, Figure S6, Figure S9 in the revised manuscript and Supplementary). For temporal variation accuracy, RSSSM is proven to be better than ASCAT-SWI, GLDAS and CCI, both in temporal correlation and RMSE, especially in arid regions, relatively cold areas (Figure 8, Figure 11, Figure S7 and Table 2). The temporal accuracy of RSSSM is similar to ERA5-Land and GLEAM v3.3 products (temporal correlation is somewhat lower, but the temporal RMSE value of our product is lower, see Figure S4, Figure S10 and Table 2). For spatial pattern accuracy, our RSSSM product is found superior to all other products (please refer to Figure 9, Figure 12, Figure S5, Figure S8, Figure S11 and Table 3), almost all year round, especially during the growing seasons. Based on these findings, we propose that our product (RSSSM) have better agreement with the site-measured surface soil moisture than the five existing soil moisture products (ASCAT-SWI, GLDAS Noah, ERA5-Land, CCI/ECV and GLEAM). Moreover, the observational-based soil moisture, CCI, has limited spatial coverage and significantly reduced data accuracy before 2012, while ASCAT-SWI is only available since 2003. These problems have been well solved by our estimation (the data quality is maintained during 2003~2018, see Figure 13). We agree with you that the phrase 'supposed to be superior' is not a scientific statement, and the claim is probably too strong and condescending. So, we corrected the sentences as: '*This* *new dataset, named RSSSM, is proved comparable to the in-situ surface soil moisture measurements at sites of the International Soil Moisture Network (overall R^2 and RMSE values of 0.42 and 0.087 m^3 m^{-3}), while the overall R^2 and RMSE values for the existing products (ASCAT-SWI, GLDAS Noah, ERA5-Land, CCI/ECV and GLEAM) are within the range of 0.31~0.41 and 0.095~0.142 m^3 m^{-3}, respectively. The advantage of RSSSM is especially obvious in arid or relatively cold areas, and during growing seasons. Moreover, the persistent high data quality as well as complete spatial coverage ensure the applicability of RSSSM to both the spatial and temporal pattern studies.*' We have also corrected all the relevant unclear statements (e.g. supposed superior, expected to be better, …) throughout the manuscript.

Reviewer 2 Major comment 2. There were some strange 'terminologies' the author used for discussion, for example: a. 'penetrability of microwave' (which is seldom found in the literature. A more widely used term is 'microwave penetration depth'); b. "Soil moisture retrieval from passive microwave sensors is based on the correlation between soil dielectric conductivity, that is influenced by soil moisture …'. Following the theoretical development of soil moisture retrievals from remote sensing, the relationship between soil moisture and dielectric constant is the fundamental (not soil dielectric conductivity).

Response: We are sorry for the unsuitable terminologies. we corrected the sentence 'the penetrability of microwaves is usually <5 cm of soil' to '*current satellite microwave sensors can only detect soil moisture within top 5 cm of soil*' following this comment and your Specific comment 5, and also corrected '*dielectric conductivity*' to '*dielectric constant*'.

Reviewer 2 Major comment 3. "However, this data is regional, with a large temporal gap, and cannot be seen as observational-based only since precipitation data is incorporated." This is a very strange argument. We all know there is a strong link between precipitation and soil moisture variation. Physically speaking, one used the antecedent precipitation index to understand how precipitation events drive the variation of soil moisture. This is like one of 'quality impact factors'. If the above argument is true, we can argue that the author's approach in this manuscript is also not 'observation-based', as they used LAI, land cover, LST, and many other factors.

Response: We checked the work by *Qu et al.* (Qu et al., 2019) carefully, and found that precipitation data was actually not applied as an input of the random forest in that study.

Furthermore, other recent studies focusing on long-term soil moisture mapping based on microwave remote sensing data did not incorporate precipitation as ancillary inputs of neural networks as well (Santi et al., 2016; Yao et al., 2019; Yao et al., 2017).

Hence, we conducted a research on the role of precipitation in neural network training to further explore the reasons.

Because it takes at least 3 days for a microwave sensor to cover the whole globe, for 11% of global land, there will be only 5 or less observations for random days within a 10-day period. By taking the average of these available data, this study only focuses on the mean soil moisture condition during that 10-day period. Then, to see how much can the incorporation of precipitation data improve the neural network training efficiency, we calculated 10-day averaged GPM Final-Run precipitation, which can well indicate the overall precipitation water availability (the antecedent precipitation index is not used, because it must be calculated on daily scale, and the attenuation coefficient is hard to determine at global scale (Kohler and Linsley, 1951)). Taking the first primary independent neural network, NN1-1-1, as an example, we performed contribution tests on all the input features at the global scale (not for each separate zone), including 9 'quality impact factors', 4 predictor soil moisture products and precipitation - a potential ancillary soil moisture indicator. For each predictor, we added a random error that is controlled within the standard deviation of the predictor, and then the increased MSE in neural network training can indicate the relative contribution of that variable. The results (see Figure R1a, that is Figure S1a in the revised Supplement) show that precipitation will only contribute to 1.7% of the training efficiency, which is much lower than the contribution of any soil moisture product (the total contribution fraction of the four soil moisture products is 61.2%), and is also lower than that of most 'quality impact factors'. This suggests that microwave soil moisture datasets together with several 'quality impact factors' of microwave soil moisture retrieval are enough to predict the training target- SMAP soil moisture, and there is no need to add precipitation as another ancillary index of soil moisture.

'Quality impact factors' are defined in this study as the variables that will have a significant impact on the retrieval errors of soil moisture by microwave remote sensing (section 2.1.2). Although the relative performances of different soil moisture products is related to surface moisture condition (Kim et al., 2015), it is found mainly due to the less vegetation in arid areas. After all, no explicit mechanism can support the idea that the retrieval errors of soil moisture are significantly influenced by water availability. Even if this is true, the soil water availability can be already indicated by the microwave soil moisture products. So, it is unreasonable to incorporate the precipitation variable as a ‘quality impact factor’. On the other hand, LAI, water area fraction, LST, land use cover, tree cover fraction, non-tree vegetation fraction, topographic complexity, and soil sand/clay fractions all have direct impacts on the microwave soil moisture retrieval errors, with solid physical mechanism (see section 2.1.2). Therefore, theoretically they should be added to the neural network, even though the land use cover type and soil sand fraction data prove to have limited contributions to NN training efficiency.

One may argue that if NARX (nonlinear autoregressive with external input) is applied instead, in which the soil moisture in the previous 10-day period is also incorporated as a predictor, precipitation data can be very beneficial to the neural network training. This is true, because precipitation directly contributes to soil moisture increases. However, NARX is not suitable for global-scale long-term continuous soil moisture mapping, because the base map (i.e., the soil moisture in the beginning of the simulation period) is hard to determine. Moreover, in mid to high latitudes, the lack of soil moisture retrievals over frozen ground in winters will lead to missing data there in summers when soil moisture data is otherwise available. So, if NARX is adopted, we can only estimate long-term soil moisture in tropics and sub-tropics with air temperature consistently higher than 0 ℃. Last but not least, if the soil moisture in the previous phase and the current precipitation amount are both incorporated, they will largely conceal the role of satellite-observed signals. As shown in Figure R1b (Figure S1b in the Supplement), the total contribution fraction of all four microwave soil moisture products is reduced to only 10.6%, while the roles of ASCAT, AMSR2-JAXA and AMSR2-LPRM are all negligible. Without taking full advantage of remote sensing, the simulations based on previous soil moisture and current precipitation will lead to errors over places where soil moisture gains are mostly driven by glacier melting, or in places with high levels of radiation-driven surface soil evaporation. The reliability of the derived soil moisture will be lower in irrigated croplands and afforestation/deforestation areas as well.

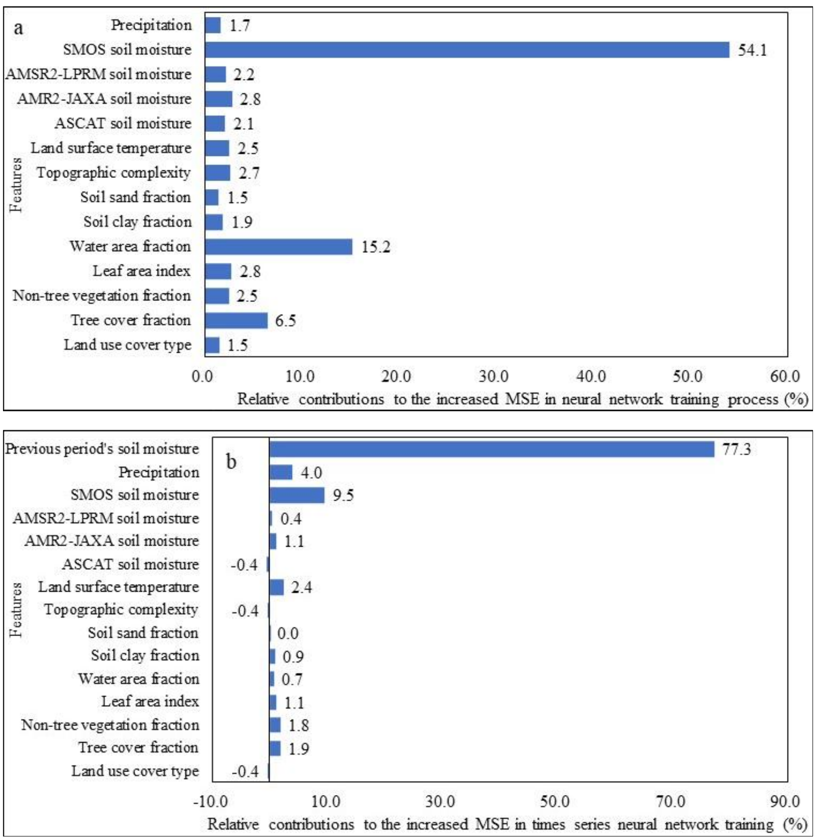


Figure R1: (Figure S1 in the revised Supplement): The roles of different input features in the soil moisture simulations based on BP neural networks and nonlinear autoregressive with external input (NARX) with microwave soil moisture products incorporated: (a) the contributions of different input features of a primary neural network: NN1-1-1, including 4 predictor soil moisture products, 9 quality impact factors of microwave soil moisture retrieval, plus 1 probable ancillary soil water indicator: 10-day averaged precipitation, to the neural network training efficiency indicated by the increased MSE; (b) the contributions of all the input features to the training efficiency, if NN1-1-1 is changed into a NARX, in which the SMAP soil *moisture for the previous period is also applied as a predictor.*

On account of all above, precipitation data is neither included as an ancillary soil moisture indicator, nor added as a 'quality impact factor' in this study.

Following this comment, we have added the above explanations to the Supplementary Data (Text S1), replacing the previous paragraph. We also added a sentence in the revised manuscript: '*The contribution analysis results (Figure S1) show that because various microwave soil moisture data have already been included, precipitation data is not an essential indicator of soil moisture, and is not utilized as a physically-based 'quality impact factor' either (see Text S1 for detailed explanations).*'

Actually, in Qu et al. (2019)'s study, the random forest's input features only include: microwave Tb products, DEM, IGBP global vegetation classification, latitude, longitude, and DOY. So, we feel sorry that the latter half of this sentence was not correct, which we have deleted. The sentences now reads: '*Another study rebuilt a soil moisture time series over the Tibetan Plateau by using SMAP data as the reference of a random forest (Qu et al., 2019). For environmental factors, while vegetation cover is not considered, elevation (DEM), IGBP land use cover type, grid location and the day of a year (DOY) are chosen as ancillary inputs. The training R^2 in this region reached 0.9, with a temporal accuracy higher than other products (temporal r=0.7; RMSE=0.07 in the unfrozen season). However, this data is regional (for Tibetan Plateau only), and with a temporal gap between AMSR-E and AMSR2 (October 2011~June 2012).*'.

Reviewer 2 Major comment 4. " are these factors used as direct spatial predictors of soil moisture or just because they are related to the errors of satellite soil moisture retrievals (i.e., the quality impact factors of soil moisture)? We insist on the latter, proposing two main reasons for the incorporation of environmental factors." This is very confusing and not necessarily correct, and not well grounded. We know the soil moisture retrieval from remote sensing is using a radiative transfer model to account for scattering and emissions from both soil and vegetation, which is conflicting with the author's statements.

Response: Sorry for the unsuitable sentences. We have re-written them as follows: '*Environmental factors, including DEM, LST and vegetation cover (indicated by NDVI, MVI, etc.), were used as ancillary neural network inputs for improved soil moisture simulation (Lu et al., 2015; Qu et al., 2019; Yao et al., 2017). According to these studies, these factors alone may not* *predict surface soil moisture well without the incorporation of any microwave remote sensing data, which can also be justified by the contribution analysis results (Figure S1a). This is because although they are somewhat related to soil moisture (e.g. soil moisture is limited in areas with low vegetation cover in general but high in forests (McColl et al., 2017)), the relationships are rather uncertain (e.g. at smaller scales, LAI may however have a negative influence on soil moisture due to the variation in evapotranspiration (Naithani et al., 2013), or without clear impacts (Zhao et al., 2010); also, soil moisture can be either high or low in summers when vegetation peaks (Baldocchi et al., 2006; Méndez-Barroso et al., 2009)). However, these factors are quite essential due to their direct impacts on soil moisture retrieval through radiative transfer model using microwave remote sensing data (Fan et al., 2020), and are retrieval quality impact factors. The detailed explanations are: 1) …*'. On the other hand, we agree that the precipitation and open water fraction can directly indicate the surface soil moisture (but because the microwave soil moisture products are already applied, precipitation data was not included as a predictor, see the detailed explanations in response to Reviewer 2 Major comment 3).

We understand that soil moisture retrieval from microwave remote sensing accounts for the scattering and emissions from both soil and vegetation, so that these factors can have direct impacts on the microwave soil moisture retrieval results, and are closely related to the retrieval errors, as we explained in the revision.

Reviewer 2 Major comment 5. 'Water Body' was used as one of the predictors (it should be predictor, rather than quality impact factors). This is very strange. As we know, water body map in either SMOS or SMAP soil moisture products were used to mark out those locations to avoid soil moisture retrievals over these water bodies (otherwise, it would be physically no sense, in terms of soil moisture). This is wrong and not physically sound to include water bodies as one of predictor for predicting surface soil moisture.

Response: We are sorry for the confusion. We agree that water is a direct indicator of surface soil moisture. However, 'water fraction', rather than 'water body', is used as both a quality impact factor and a potential indicator of surface soil moisture content. Water bodies (large lakes, oceans) are masked out in both existing (SMOS or SMAP) soil moisture products and our simulation product (RSSSM). However, in a grid with size of 0.1°×0.1° (approximately 120 km^2), there could be a small fraction of water, which may be rivers, streams, ponds, partly inundated wetlands or just paddy croplands. If we mask out all 0.1° resolution grids with even 1% of water (or less), there will be no data in many parts of the world, especially over humid areas. Previous soil moisture products also produce valid values in those grids. However, water can dramatically lower the brightness temperature (Tb), while different retrieval algorithms correct the impact of waters within grid differently, leading to different biases and relative accuracy of soil moisture estimates in grids with water (Ye et al., 2015). For example, we noted a strong underestimation of soil moisture by the NSIDC (National Snow and Ice Data Center) method (e.g., AMSRE-NSIDC) over rivers and small lakes in compared to the nearby lands). Moreover, the sensitivity of different microwave sensors to water fraction within grid may differ as well. Hence, the fraction of water (not only open water, but also inundated wetlands or croplands) in the grid can significantly influence the retrieval errors and relative reliability of various soil moisture products (Ye et al., 2015), which exactly meets the definition of 'quality impact factor'.

Therefore, this study uses 'water fraction' both as a quality impact factor and as an ancillary soil moisture indicator (please also note that the word 'predictor' in this manuscript only refers to the existing soil moisture products that are applied as the neural network inputs). This information has been added to the manuscript. For 'water fraction', the Surface WAter Microwave Product Series (SWAMPS) dataset (Schroeder et al., 2015) was applied because it is microwave based, including not only open water, but also partly-inundated lands. The contribution analysis on all the input features (see Figure R1) prove that the calculated water fraction plays a very important role in neural network training.

Following this comment, we have clarified the descriptions. It now reads: '*The second is the 'water fraction factor' (i.e., the fraction of water area in each pixel). Waters in land pixels dramatically lower the Tb, leading to overestimation of soil moisture there. Because there are different methods used for detection and correction of small area of water, either open water, wetlands or partly inundated wetlands and croplands (Entekhabi et al., 2010; Kerr et al., 2001; Mladenova et al., 2014; Njoku et al., 2003), microwave soil moisture data calibration and weight assignment based on the water fraction within land pixels make sense (Ye et al., 2015). In addition, water fraction is a direct indicator of surface soil moisture. In this study, daily water area fraction derived from the Surface WAter Microwave Product Series (SWAMPS) v3.2 dataset (Schroeder et al., 2015) was applied.*'. The confusing words 'water body' have been replaced by '*water area fraction*' in other parts of the manuscript as well.

Reviewer 2 Major comment 6. For ‘topographic complexity’ ‘soil texture’, the author used from different sources, one from ASCAT ancillary data and the other use SMAP ancillary data. This reviewer is wondering why such a choice? Why not making it consistent (i.e., get ancillary data from one single product, instead of two?

Response: The reason we used different sources of data for ‘topographic complexity’ and ‘soil texture’ is the data availability. SMAP only uses GMTED 2010 DEM data to derive quality flags for data retrieved in mountainous areas, while topographic complexity is only included as the ancillary data of ESA’s ASCAT-SWI product, which was calculated by normalizing the standard deviation of GTOPO30 elevation in each grid point to values from 0 to 100 (Scipal et al., 2005), and is closely related to errors of surface soil moisture retrieved from microwave remote sensing. On the other hand, soil texture data is not contained in the static layers of ASCAT-SWI product, so we have to obtain this important data from SMAP ancillary input collection. Moreover, because the data sources of topographic complexity and soil texture both have relatively high quality, we suggest that they could be used even they come from different soil moisture products. In the revised manuscript, we added this information: ‘*(topographic complexity data is not available from SMAP Constant; soil texture is not provided by ASCAT Constant)*’.

Reviewer 2 Major comment 7. ‘3σ denoise’. what is the effect of such a filter on identifying extreme years? For example, during 2003, 2010, 2018, 2019 there are extreme heat events in Europe and the soil moisture is so dry which can be beyond the 3 standard deviations.

Response: The ‘3σ denoise’ was conducted spatially, rather than temporally, to detect and delete the extreme values (usually salt and pepper noises in mountain areas) in each 1°×1° zone, during a certain 10-day period. So, if there are extreme heat or precipitation events, as you noted, the whole 1°×1° zone will see a sharp increase or decrease in soil moisture content in almost all the 0.1° grids within that zone (there are <100 grids in each zone). Therefore, due to the increase/decrease in the zonal mean soil moisture value, the extreme weather events will not be removed by this ‘spatial 3σ denoise’ step.

We are sorry for the confusion. We have added more detailed explanations. It reads as: '*After standardization of the original soil moisture data, to improve the neural network training efficiency, the potential salt and pepper noises are removed. For each map (a specific 10-day period), within each 1°×1° zone, the soil moisture values are filtered to the level of three standard deviations relative to the mean in that zone. This preprocessing step is thus called '3σ denoise' (note that the denoise is conducted spatially, rather than temporally, so that the extreme events will not be treated).*'

Reviewer 2 Major comment 8. NN design. SMAP is only available after 2015, so I am not sure what is the meaning of simulation period 2012D19-2013D36, but also 2014-2018. I guess this is constrained by the available data (PROBA-V and GLASS LAIs)? But in any case, it does not represent any physical meaning to predict 2015data with 2012-2013 data. At least, the NN design is not clear on why it is designed as such.

Response: Sorry for the confusion. In the first round of simulation, the division of the simulation period into two subperiods: 2012D19~2013 and 2014~2018, is due to the available data periods of PROBA-V and GLASS LAI. However, we did not predict data in 2015 by using the data in 2012~2013. In this study, the common period of predictor soil moisture products applied in each 'substep' of NN training always includes the corresponding soil moisture simulation period.

We agree that the design of 67 independent neural networks, which are embedded in 8 substeps applied for five rounds of simulations, is quite complex. However, it ensures the long-term continuous satellite-based soil moisture mapping, almost full spatial coverage at the global scale, as well as high data accuracy. Following this comment, we have revised this section in the manuscript to make the NN design clearer. It reads as:

'*Not only the 11 available microwave soil moisture data with different temporal spans are all incorporated, but they are also utilized as fully as possible through up to 5 rounds of neural network-based simulations, with at least four different soil moisture products retrieved from three sensors applied as predictors in each round (see details below). While increasing the sources of soil moisture data inputs can be beneficial to the training efficiency, the spatial coverage of the simulation output is sacrificed because the overlapping area of more soil moisture products is smaller. After all, most products have missing data in specific regions (e.g., mountains, wetlands and urban settlements), and some sensors are even unable to produce data* *at global scale (TMI is limited to [N40°, S40°]; SMOS lacks data in Asia). To solve that dilemma, we classified all 0.1° pixels according to the available predictor soil moisture products in it over a 10-day period (for example, if there are at most four predictor soil moisture data inputs in one round, there should be 4+3+2+1=10 combinations). However, to avoid soil moisture simulation under snow or ice cover (Section 2.2.2), not all combinations are considered. Then, corresponding to each selected combination, an independent neural network is trained. For data simulation in a 0.1° pixel, the most preferable independent neural network is expected to be trained using all the available soil moisture data sources in that pixel. However, in the 1° zone where it is located, the subnetwork belonging to that preferable independent neural network may not exist due to limited valid data points (see section 2.2.1). Then, an alternative subnetwork driven by the combination of fewer soil moisture data inputs should be applied instead. Hence, we should determine which neural network collocation is the best choice for every pixel. Apart from applicability, the relative priority order of different neural networks were obtained by comprehensively considering the number and quality of input soil moisture products, the variety of sensors, the quantity of training samples indicated by the number of 10-day periods, and the relative accuracy of training targets (the training target quality declines monotonously: SMAP>SIM-1T>SIM-2T>SIM-3T>SIM-4T). Sometimes, two most probable priority orders are given, with the simulation results of the corresponding two substeps integrated later. Specifically, when the LAI data source changes, the division of a single round into several substeps is also essential. Based on these principles, five rounds of neural networks are designed as follows, with 8 substeps containing a total of 67 independent neural networks. The training period of each neural network and the simulation period of each substep are shown in Figure 1 (below the timeline), and the details are as follows:*

For the first round's neural network (labelled as NN1), the potential training period is 2015D10~2018 ('D' is the ordinal of the 10-day period, so '2015D10' represents since the April 1st to 10th of 2015) because SMAP soil moisture data during that period is applied as the training target, while ASCAT-SWI10 (abbreviated as ASCAT), SMOS-IC (SMOS), AMSR2-JAXA and AMSR2-LPRM-X (AMSR2-LPRM) are the four predictor soil moisture products (details are in Table S1~S2). Because all the four predictors have data since 2012D19, the potential soil moisture simulation period is 2012D19~2018, which was further divided into two parts: one is 2014~2018 (substep1), for which the PROBA-V LAI data that starts from 2014, is applied, whereas the other is 2012D19~2013 (substep2), for which GLASS LAI data is used (note: because GLASS LAI covers from *the beginning of our study period till 2017, the training period for substep 2 is 2015D10~2017). Please refer to Table S1~S2 for details. The simulation results of the two substeps (SIM-1-1 and SIM-1-2) are combined as SIM-1, and then transformed into a secondary training target, denoted by SIM-1T. In the second round of simulation, the training target can either be SMAP or SIM-1T, while the input soil moisture data are ASCAT, SMOS, TMI-LPRM-X (TMI) and FY-3B-NSMC (FY). The simulation output, SIM-2, covers the period of 2011D20~2012D18, that is constrained by the common period of the four predictors (Table S3~S4). In the third round of neural network operation, the simulation period is 2010D16~2011D19. SMAP, SIM-1T and SIM-2T are combined to be the training targets (the training periods are within the range of 2011D20~2017D36), while the predictor soil moisture data are ASCAT, SMOS, TMI and WindSat-LPRM-X (WINDSAT). There are two substeps in round 3, distinguished by whether the priority order of the neural networks is determined mainly based on the training sample quantity and the training target quality (SIM-3-1), or by first considering the number of predictor soil moisture products (SIM-3-2, Table S5~S8). Because these two methods emphasize different aspects of neural network quality, in some pixels, SIM-3-1 will be advantageous, but in others, SIM-3-2 could be better. Hence, an algorithm is devised to combine the advantages of both simulations (SIM-3), which is described in Table S9. Next, the 4th round is for simulations during 2007D01~2010D15. With SIM-2T and SIM-3T combined to be the training target, ASCAT, WINDSAT, TMI, AMSRE-JAXA, AMSRE-LPRM-X (AMSRE-LPRM) and AMSRE-NSIDC are all applied as predictors (LAI data now comes from SPOT-VGT). Two substeps are also needed. In the first substep, neural networks are sorted by paying more attention to the number of soil moisture inputs and the sensors they are derived from, while the training sample size and training target quality are prioritized to make an alternative estimate (Tables S10~S13). Afterwards, SIM-4 is obtained by reasonably integrating these two results. In the final round, the soil moisture simulation is extended to as early as 2003. SIM-2T, SIM-3T and SIM-4T together are the training target, while the predictor soil moisture data entering the neural networks consist of WINDSAT, TMI, AMSRE-JAXA, AMSRE-LPRM and AMSRE-NSIDC (Table S14~S15).* ’

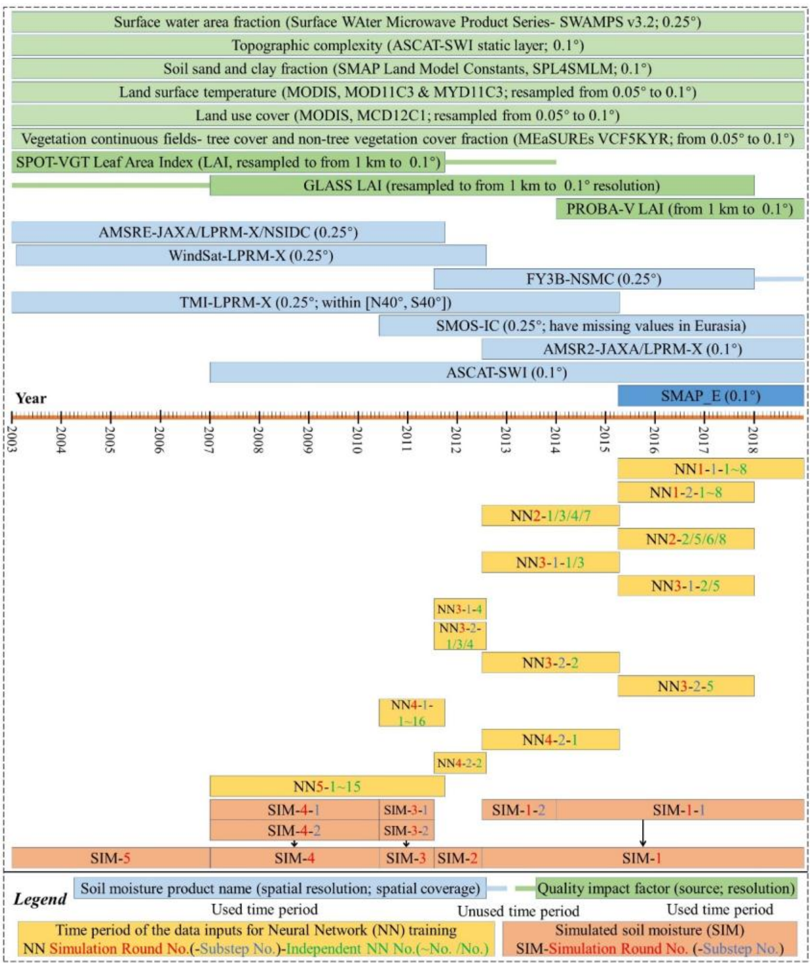


Figure R2 (Figure 1 in the revised manuscript): The timeline figure showing the time periods of the soil moisture datasets and the 'quality impact factor' products (e.g. LAI dataset) used in this study (listed above the timeline), as well as the periods of data applied for the trainings of 67 independent neural networks and the neural network simulation outputs (i.e. simulated *soil moisture) of eight substeps (listed below the timeline).*

The following is a plain language description, which you may choose to read if would like to understand the NN design deeper. First, considering that the temporal spans of different microwave sensors are all limited (see Figure R2 for details), we designed five rounds of neural networks to achieve long-term continuous soil moisture mapping, while ensuring that as many microwave soil moisture products as possible are applied as predictors of each round of NN. In detail, SMAP soil moisture data is used as the training target of the first round NN (labeled as NN1), with ASCAT-SWI, SMOS, AMSR2-JAXA and AMSR2-LPRM-X applied as predictor soil moisture products. The potential training period of NN1 is the time period of SMAP (2015D10~2018, Table S1). Because the four soil moisture predictors all have data since 2012D19, the potential soil moisture simulation period is 2012D19~2018. However, because PROBA-V LAI (quality impact factor) starts in 2014, the neural networks trained using PROBA-V LAI can only be used for the simulation during 2014~2018. For the remaining period (2012D19~2013), the applicable neural networks should be trained based on another LAI dataset- GLASS LAI, which covers from the beginning of our study period until 2017. Therefore, NN1 should be divided into two substeps. For substep 1 (marked by NN1-1), PROBA-V LAI is used, and the training period is 2015D10~2018D36 (Table S1), while the simulation period is 2014~2018 (Table S2). For substep 2 (denoted by NN1-2), GLASS LAI is applied instead, and the training period is 2015D10~2018D36, while the simulation period is 2012D19~2013. Because each predictor soil moisture product has missing values in some specific areas (e.g. SMOS-IC do not have values in Eurasia), there are 1~4 predictor soil moisture products available in every 0.1° grid. While the maximum number of combinations are 4+3+2+1=10, 8 of them are valid since the soil moisture retrievals over snow or ice is not recommended (Table S2). Corresponding to these 8 combinations, 8 independent neural networks are trained, each with a combination of predictor soil moisture products applied as neural network inputs (labeled as NN1-1(2)-1 ~ NN1-1(2)-8; for example, NN1-1(2)-1 is trained using all four predictor soil moisture products, and is the most preferable NN). However, even for a 0.1° grid with all four predictor soil moisture data available, we may not be able to simulate soil moisture there using NN1-1(2)-1. That is because the corresponding neural network, NN1-1(2)-1, may not exist in the 1°×1° zone where the grid is located, due to limited valid data points available for zonal subnetwork training (please refer to revised Method section 2.2.1 for details on the localized neural networks). Under this condition, the other less preferable independent neural networks should be applied instead (the relative priority order of all independent neural networks within a substep is determined by comprehensively considering the number and quality of input soil moisture products, the variety of sensors, the quantity of training samples indicated by the number of 10-day periods, and the relative accuracy of training targets). After simulation, we combined the results for substep 1 (NN1-1-1~8), which is denoted by SIM-1-1 and the results for substep 2 (NN1-2-1~8): SIM-1-2 to obtain SIM-1. After further processing steps (section 2.2.2), we convert SIM-1 into the secondary training target, SIM-1T. For the second round of NN, the training target can be either SMAP, (rimary training target), while the training period is 2015D10~2017 (GLASS LAI is used, ASCAT-SWI, SMOS and FY data products are applied as predictors), or SIM-1T (secondary training target), while the training period is 2012D19~2015D10 (ASCAT-SWI, SMOS, FY and TMI products can all be applied). There are 8 independent neural networks included in round 2 NN (see Table S3), while the corresponding simulation output is SIM-2, covering the period of 2011D20~2012D18 since FY data product is available since 2011D20 (see Table S4). The 3rd to 5th round of neural network training and simulations are even more complex (for example, in the 3rd round, the priority order of independent neural networks is not definite. Two probable orders are provided, leading to two substeps, and the simulation results of which are combined by taking the relative accuracy in each grid into account), but the basic principles are similar to those explained above (see Table S5~S15).

Reviewer 2 Specific comment 1. In the abstract, change ‘elaborate’ to ‘elaborated’, delete ‘various’, change ‘simulation’ to ‘simulations’

Response: we have corrected them accordingly.

Reviewer 2 Specific comment 2. In the abstract, ‘This new dataset, once validated against the International Soil Moisture Network (ISMN) records, is supposed to be superior to the existing products (ASCAT-SWI, GLDAS Noah, ERA5-Land, CCI/ECV and GLEAM), and is applicable to studying both the spatial and temporal patterns.' This is a very strong assumption, and should be avoided. Otherwise, the corresponding results should be shown.

Response: We have corrected the unsuitable statement as: '*This new dataset, named RSSSM, is proved comparable to the in-situ surface soil moisture measurements at sites of the International Soil Moisture Network (overall R^2 and RMSE values of 0.42 and 0.087 m^3 m^{-3}), while the overall R^2 and RMSE values for the existing products (ASCAT-SWI, GLDAS Noah, ERA5-Land, CCI/ECV and GLEAM) are within the range of 0.31~0.41 and 0.095~0.142 m^3 m^{-3}, respectively. The advantage of RSSSM is especially obvious in arid or relatively cold areas, and during growing seasons. Moreover, the persistent high data quality as well as complete spatial coverage ensure the applicability of RSSSM to both the spatial and temporal pattern studies.*'. Please also find the details in the response to Major Comment 1.

Reviewer 2 Specific comment 3. Lines 27~30: 'It has been endorsed by the Global Climate Observing System (GCOS) as an essential climate variable (Bojinski et al., 2014), probably the best indicator of ecological droughts (Martínez-Fernández et al., 2016; Samaniego et al., 2018). However, due to the large uncertainty in global-scale soil moisture data, its applicability in global ecosystem models are currently limited (Hashimoto et al., 2015; Stocker et al., 2019).' What do you want to say here? What are points? It is suggested to shorten the sentence.

Response: Sorry for the too complicated sentences. We have shorten the sentence as: *'It has been endorsed by the Global Climate Observing System (GCOS) as an essential climate variable (Bojinski et al., 2014), probably the best indicator of ecological droughts (Martínez-Fernández et al., 2016; Samaniego et al., 2018). However, due to the large uncertainty in global-scale soil moisture data, its applicability in global ecosystem models are currently limited (Hashimoto et al., 2015; Stocker et al., 2019).*'

'

Reviewer 2 Specific comment 4. Lines 31~33: 'The reanalysis land surface model products (e.g., the Global Land Data Assimilation System (GLDAS, with spatial resolution of 0.25°) (Rodell et al., 2004), ECMWF ERA-interim (0.25°) (Balsamo et al., 2015) and its newly-published successors: ERA5 (0.25°) and the land product, ERA5-Land (0.1°)(Hoffmann et al., 2019)) are the most frequently used.’ The sentence seems not completed.

Response: We have revised this complicated sentence as: ‘*The reanalysis-based land surface model products are the most frequently used, mainly including the Global Land Data Assimilation System (GLDAS, with 0.25° resolution) (Rodell et al., 2004), European Reanalysis (ERA)-interim (0.75°) (Balsamo et al., 2015) and its successors- ERA5 (0.25°) and ERA5-Land (0.1°) (Hoffmann et al., 2019)).*’

Reviewer 2 Specific comment 5. Lines 40: ‘the penetrability of microwaves is usually <5 cm of soil’. This is particularly not true for L-band passive microwave like SMOS, SMAP, which are dedicated to soil moisture monitoring. 'Penetrability' is usually called penetration depth.

Response: Following this comment, we have checked carefully to find that the L-band microwave can only be sensitive to the soil moisture within <5 cm of surface soil. For example, ‘L-band, the brightness temperature emission originates from the top ~5 cm of soil’ for SMAP (Entekhabi et al., 2010). ‘At L-band soil moisture in the first centimeters (typically 5 cm) impacts significantly on the emitted brightness temperature’ for SMOS (Kerr et al., 2001). This depth is however, larger than the observation depth of higher frequency microwave, which is 1 cm (C band) or less (Piles et al., 2018). We agree that the word ‘penetrability’ is not scientific, while ‘penetration depth’ are not very accurate as well. Therefore, we have changed the sentence to ‘*current satellite microwave sensors can only detect soil moisture within top 5 cm of soil*’ for make it clearer.

Reviewer 2 Specific comment 6. Line 45: change ‘frommicrowave’ to ‘from microwave’.

Response: We have made the revision accordingly.

Reviewer 2 Specific comment 7. Line 45: ‘Currently, the longest continuous record of global soil moisture retrieved frommicrowave remote sensing only is the ASCAT product’. Change ‘only is’ .

Response: Following this comment, we have corrected the sentence as: ‘*Currently, the longest continuous record of global surface soil moisture that is derived only from microwave remote sensing is the ASCAT product*’.

Reviewer 2 Specific comment 8. Line 63~64: 'Upon rescaling, the spatial patterns of the satellite products are almost replaced by those of GLDAS.' , Any citations?

Response: We have added the citations of (Gruber et al., 2019; Liu et al., 2012; Liu et al., 2011). Moreover, the statement has been changed into: '*Upon rescaling through CDF matching, the spatial patterns of the satellite products are generally replaced by those of GLDAS (Gruber et al., 2019; Liu et al., 2012; Liu et al., 2011)*'.

Reviewer 2 Specific comment 9. Line 66~67: 'Because the temporal variation in soil moisture is often better captured by model simulations than remote sensing inversions, CCI may undesirably combine the disadvantages of both.' How? Any proof?

Response: We have deleted the unsuitable description accordingly. There is no strong evidence for the claim that the temporal variation in soil moisture is better captured by model simulations than remote sensing inversions.

Reviewer 2 Specific comment 10. Change 'deviationsto' to 'deviations to'.

Response: We have checked the manuscript carefully and added all the missing blanks.

Reviewer 2 Specific comment 11. Lines 90~91: 'The training R^2 is only 0.45 (R=0.67)'. R^2 and R needs to specify.

Response: Thank you for this suggestion. We have revised it as: '*The training R-square value (R^2) is only 0.45 (or correlation coefficient, r, equals to 0.67)*'. Following this comment, we have also corrected all the 'R' into '*r*', to distinguish it from R^2 in the revised manuscript.

Reviewer 2 Specific comment 12. Lines 99~100: 'this data is regional, with a large temporal gap, and cannot be seen as observational-based only since precipitation data is incorporated.' Well, this is arguable.

Response: Accordingly, the sentence has been revised to '*However, this data is regional (for Tibetan Plateau only), and with a temporal gap between AMSR-E and AMSR2 (October 2011~June 2012).*' We have also re-written Text S1 to more clearly explain why precipitation was not included as an input feature of the neural networks in this study. Please also find the details in the responses to Major Comment 3.

Reviewer 2 Specific comment 13. Lines 148~150: 'are these factors used as direct spatial predictors of soil moisture or just because they are related to the errors of satellite soil moisture retrievals (i.e., the quality impact factors of soil moisture)?' What do you mean?

Response: We have revised the sentences to: '*Environmental factors, including DEM, LST and vegetation cover (indicated by NDVI, MVI, etc.), were used as ancillary neural network inputs for improved soil moisture simulation (Lu et al., 2015; Qu et al., 2019; Yao et al., 2017). According to these studies, these factors alone may not predict surface soil moisture well without the incorporation of any microwave remote sensing data, which can also be justified by the contribution analysis results (Figure S1a). This is because although they are somewhat related to soil moisture (e.g. soil moisture is limited in areas with low vegetation cover in general but high in forests (McColl et al., 2017)), the relationships are rather uncertain (e.g. at smaller scales, LAI may however have a negative influence on soil moisture due to the variation in evapotranspiration (Naithani et al., 2013), or without clear impacts (Zhao et al., 2010); also, soil moisture can be either high or low in summers when vegetation peaks (Baldocchi et al., 2006; Méndez-Barroso et al., 2009)). However, these factors are quite essential due to their direct impacts on soil moisture retrieval through radiative transfer model using microwave remote sensing data (Fan et al., 2020), and are retrieval quality impact factors. The detailed explanations are: 1) …*'.

Reviewer 2 Specific comment 14. Lines 178: 'Water bodies dramatically lower the Tb, leading to overestimation of soil moisture'. water bodies are marked/flagged out for soil moisture retrieval. The consideration of water bodies in your approach seems very strange.

Response: We have replaced 'water bodies' with 'water area fraction' for clarity. Please also refer to our response to Major Comment 5.

Reviewer 2 Specific comment 15. Line 182, are you sure it is soil dielectric conductivity, not soil dielectric constants?

Response: We have corrected 'soil dielectric conductivity' to '*soil dielectric constant*' accordingly.

Reviewer 2 Specific comment 16. Lines 199~200, 'For topographic complexity, the static layer of the Copernicus ASCAT-SWI product (hereinafter the ASCAT Constant) is adopted while for soil texture, the SMAP Constant is used.' Why not using static layers from the same satellite product?

Response: Thank you for the question. That is because topographic complexity and soil texture data cannot be obtained from one product. Also, the quality for the two data are satisfying (Reichle et al., 2018; Scipal et al., 2005). Please see the response to Major Comment 6.

Reviewer 2 Specific comment 17. Lines 227~228, 'For each map, soil moisture values are filteredto the level of three standard deviations relative to themean in each zone. This preprocessing step is thus called '3σ denoise'.' What is the effect of such filter on identifying extreme years? for example, during 2003, 2010, 2018, 2019 there are extreme heat events in Europe and the soil moisture is so dry which can be beyond the 3 standard deviations.

Response: This is a spatial (zonal) '3σ denoise', which helps to masks out the wrong retrievals in mountain areas (zonal extreme values), however will not mask out the extreme climatic events. which will not mask out the extreme climatic events. Please find the detailed explanation in the response to Major Comment 7. Following this comment, we have made the clarification as: "*After standardization of the original soil moisture data, to improve the neural network training efficiency, the potential salt and pepper noises are removed. For each map (a specific 10-day period), within each 1°×1° zone, the soil moisture values are filtered to the level of three standard deviations relative to the mean in that zone. This preprocessing step is thus called '3σ denoise' (note that the denoise is conducted spatially, rather than temporally, so that the extreme events will not be treated)*.'

Reviewer 2 Specific comment 18. Line 273: 'SMAP soil moisture is the training target'. SMAP is only available after 2015, so I am not sure what is the meaning of simulation period 2012D19~2013D36, but also 2014-2018. I guess this is constrained by the available data (PROBA-V and GLASS LAIs)? But in any case, it does not represent any physical meaning to predict 2015data with 2012-2013 data.

Response: Here, we did not predict data in 2015 by using the data in 2012~2013. In this study, the common data period of predictor soil moisture products in each substep always contains the period of the simulated soil moisture. There are two substeps in Round 1, which were separated due to the data period of different LAI products, while each substep was responsible for a simulation period (2012D19~2013D36 and 2014-2018, respectively). Following this comment, we have revised that section, for example, '*Because all the four predictors have data since 2012D19, the potential soil moisture simulation period is 2012D19~2018, which was further divided into two parts: one is 2014~2018 (substep1), for which the PROBA-V LAI data that starts from 2014, is applied, whereas the other is 2012D19~2013 (substep2), for which GLASS LAI data is used (note: because GLASS LAI covers from the beginning of our study period till 2017, the training period for substep 2 is 2015D10~2017).*'. For more details, please see response to Major Comment 8.

Reviewer 2 Specific comment 19. Line 278: '2011D20 to 2012D18 (Table S3~S4). In the third round (2010D16~2011D19)'. Why and how are these time spans defined?

Response: The time spans of the soil moisture simulation period corresponding to the different rounds of neural networks were determined based on the temporal coverages of the different microwave sensors which are utilized as predictors. For example, the simulation period (Note: not the training period) of the second round NN is constrained by the common period of ASCAT, SMOS, FY and TMI data. Following this comment, we have made some clarifications, such as: '*In the second round of simulation, the training target can either be SMAP or SIM-1T, while the input soil moisture data are ASCAT, SMOS, TMI-LPRM-X (TMI) and FY-3B-NSMC (FY). The simulation output, SIM-2, covers the period of 2011D20~2012D18, that is constrained by the common period of the four predictors (Table S3~S4).*' We have also attached the details for the design of five rounds of neural network operations in Tables S1~S15. Please see the details in the response to Major Comment 8.

Reviewer 2 Specific comment 20. Line 278, ‘In the third round (2010D16~2011D19), SMAP…’ Again, SMAP is only available after 2015.

Response: For the 3rd Round , the simulation period is 2010D16~2011D19, but the neural network training period could be 2015D10~2017D36, 2012D19~2015D10, or 2011D20~2012D21, depending on whether the training target is SMAP, or SIM-1T (SIM-1T is the postprocessed simulation output of the first round NN), or both SIM-1T and SIM-2T. The detailed information on NN training and soil moisture simulation in Round 3 were provided in Tables S5~S8. We have also revised the sentence in the manuscript as: ‘*In the third round of neural network operation, the simulation period is 2010D16~2011D19. SMAP, SIM-1T and SIM-2T are combined to be the training targets (the training periods are within the range of 2011D20~2017D36), while the predictor soil moisture data are ASCAT, SMOS, TMI and WindSat-LPRM-X (WINDSAT)….*’

References

Baldocchi, D., Tang, J., and Xu, L.: How switches and lags in biophysical regulators affect spatial-temporal variation of soil respiration in an oak-grass savanna, Journal of Geophysical Research: Biogeosciences, 111, https://doi.org/10.1029/2005JG000063, 2006

Balsamo, G., Albergel, C., Beljaars, A., Boussetta, S., Brun, E., Cloke, H., Dee, D., Dutra, E., Muñoz-Sabater, J., Pappenberger, F., de Rosnay, P., Stockdale, T., and Vitart, F.: ERA-Interim/Land: a global land surface reanalysis data set, Hydrol. Earth Syst. Sci., 19, 389-407, http://doi.org/10.5194/hess-19-389-2015, 2015

Bojinski, S., Verstraete, M., Peterson, T. C., Richter, C., Simmons, A., and Zemp, M.: The Concept of Essential Climate Variables in Support of Climate Research, Applications, and Policy, Bull. Amer. Meteorol. Soc., 95, 1431-1443, http://doi.org/10.1175/BAMS-D-13-00047.1, 2014

Entekhabi, D., Njoku, E., O'neill, P., Kellogg, K. H., Crow, W., Edelstein, W. N., Entin, J. K., Goodman, S. D., Jackson, T., Johnson, F. M., Kimball, J., Piepmeier, J., Koster, R. D., Martin, E., McDonald, C. K., Moghaddam, M., Moran, M. S., Reichle, R., Shi, J. C., Spencer, D., Thurman, S. W., Tsang, L., and Zyl, J. V.: The Soil Moisture Active Passive (SMAP) Mission, Proc. IEEE, 98, 704-716, https://doi.org/10.1109/JPROC.2010.2043918, 2010

Fan, X., Liu, Y., Gan, G., and Wu, G.: SMAP underestimates soil moisture in vegetation-disturbed areas primarily as a result of biased surface temperature data, Remote Sens. Environ., 247, 111914, https://doi.org/10.1016/j.rse.2020.111914, 2020

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

Hashimoto, S., Carvalhais, N., Ito, A., Migliavacca, M., Nishina, K., and Reichstein, M.: Global spatiotemporal distribution of soil respiration modeled using a global database, Biogeosciences, 12, 4121-4132, https://doi.org/10.5194/bg-12-4121-2015, 2015

Hoffmann, L., Günther, G., Li, D., Stein, O., Wu, X., Griessbach, S., Heng, Y., Konopka, P., Müller, R., Vogel, B., and Wright, J. S.: From ERA-Interim to ERA5: the considerable impact of ECMWF's next-generation reanalysis on Lagrangian transport simulations, Atmos. Chem. Phys., 19, 3097-3124, https://doi.org/10.5194/acp-19-3097-2019, 2019

Kerr, Y. H., Waldteufel, P., Wigneron, J., Martinuzzi, J., Font, J., and Berger, M.: Soil moisture retrieval from space: the Soil Moisture and Ocean Salinity (SMOS) mission, IEEE Trans. Geosci. Remote Sensing, 39, 1729-1735, https://doi.org/10.1109/36.942551, 2001

Kim, S., Liu, Y. Y., Johnson, F. M., Parinussa, R. M., and Sharma, A.: A global comparison of alternate AMSR2 soil moisture products: Why do they differ?, Remote Sens. Environ., 161, 43-62, https://doi.org/10.1016/j.rse.2015.02.002, 2015

Kohler, M. A. and Linsley, R. K.: Predicting the runoff from storm rainfall, US Department of Commerce, Weather Bureau, 1951.

Liu, Y. Y., Dorigo, W. A., Parinussa, R. M., de Jeu, R. A. M., Wagner, W., McCabe, M. F., Evans, J. P., and van Dijk, A.

I. J. M.: Trend-preserving blending of passive and active microwave soil moisture retrievals, Remote Sens. Environ., 123, 280-297, https://doi.org/10.1016/j.rse.2012.03.014, 2012

Liu, Y. Y., Parinussa, R. M., Dorigo, W. A., De Jeu, R. A. M., Wagner, W., van Dijk, A. I. J. M., McCabe, M. F., and Evans, J. P.: Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals, Hydrol. Earth Syst. Sci., 15, 425-436, http://doi.org/10.5194/hess-15-425-2011, 2011

Lu, Z., Chai, L., Ye, Q., and Zhang, T.: Reconstruction of time-series soil moisture from AMSR2 and SMOS data by using recurrent nonlinear autoregressive neural networks, 26-31 July 2015 2015, 980-983, https://doi.org/10.1109/IGARSS.2015.7325932.

Martínez-Fernández, J., González-Zamora, A., Sánchez, N., Gumuzzio, A., and Herrero-Jiménez, C. M.: Satellite soil moisture for agricultural drought monitoring: Assessment of the SMOS derived Soil Water Deficit Index, Remote Sens. Environ., 177, 277-286, https://doi.org/10.1016/j.rse.2016.02.064, 2016

McColl, K. A., Alemohammad, S. H., Akbar, R., Konings, A. G., Yueh, S., and Entekhabi, D.: The global distribution and dynamics of surface soil moisture, Nat. Geosci., 10, 100, https://doi.org/10.1038/ngeo2868, 2017

Méndez-Barroso, L. A., Vivoni, E. R., Watts, C. J., and Rodríguez, J. C.: Seasonal and interannual relations between precipitation, surface soil moisture and vegetation dynamics in the North American monsoon region, J. Hydrol., 377, 59-70, https://doi.org/10.1016/j.jhydrol.2009.08.009, 2009

Mladenova, I. E., Jackson, T. J., Njoku, E., Bindlish, R., Chan, S., Cosh, M. H., Holmes, T. R. H., de Jeu, R. A. M., Jones, L., Kimball, J., Paloscia, S., and Santi, E.: Remote monitoring of soil moisture using passive microwave-based techniques — Theoretical basis and overview of selected algorithms for AMSR-E, Remote Sens. Environ., 144, 197-213, https://doi.org/10.1016/j.rse.2014.01.013, 2014

Naithani, K. J., Baldwin, D. C., Gaines, K. P., Lin, H., and Eissenstat, D. M.: Spatial Distribution of Tree Species Governs the Spatio-Temporal Interaction of Leaf Area Index and Soil Moisture across a Forested Landscape, PLOS ONE, 8, e58704, https://doi.org/10.1371/journal.pone.0058704, 2013

Njoku, E. G., Jackson, T. J., Lakshmi, V., Chan, T. K., and Nghiem, S. V.: Soil moisture retrieval from AMSR-E, IEEE Trans. Geosci. Remote Sensing, 41, 215-229, https://doi.org/10.1109/TGRS.2002.808243, 2003

Piles, M., Schalie, R. v. d., Gruber, A., Muñoz-Marí, J., Camps-Valls, G., Mateo-Sanchis, A., Dorigo, W., and Jeu, R. d.: Global Estimation of Soil Moisture Persistence with L and C-Band Microwave Sensors, 22-27 July 2018 2018, 8259-8262, https://doi.org/10.1109/IGARSS.2018.8518161.

Qu, Y., Zhu, Z., Chai, L., Liu, S., Montzka, C., Liu, J., Yang, X., Lu, Z., Jin, R., Li, X., Guo, Z., and Zheng, J.: Rebuilding a Microwave Soil Moisture Product Using Random Forest Adopting AMSR-E/AMSR2 Brightness Temperature and SMAP over the Qinghai–Tibet Plateau, China, Remote Sens., 11, https://doi.org/10.3390/rs11060683, 2019

Reichle, R., Lannoy, G. D., Koster, R. D., Crow, W. T., Kimball, J. S., and Liu., Q.: SMAP L4 Global 9 km EASE-Grid Surface and Root Zone Soil Moisture Land Model Constants, Version 4. , NASA National Snow and Ice Data Center Distributed Active Archive Center, https://doi.org/10.5067/KGLC3UH4TMAQ, 2018

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C. J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D., and Toll, D.: The Global Land Data Assimilation System, Bull.

Amer. Meteorol. Soc., 85, 381-394, https://doi.org/10.1175/BAMS-85-3-381, 2004

Samaniego, L., Thober, S., Kumar, R., Wanders, N., Rakovec, O., Pan, M., Zink, M., Sheffield, J., Wood, E. F., and Marx, A.: Anthropogenic warming exacerbates European soil moisture droughts, Nat. Clim. Chang., 8, 421-426, https://doi.org/10.1038/s41558-018-0138-5, 2018

Santi, E., Paloscia, S., Pettinato, S., and Fontanelli, G.: Application of artificial neural networks for the soil moisture retrieval from active and passive microwave spaceborne sensors, Int. J. Appl. Earth Obs. Geoinf., 48, 61-73, https://doi.org/10.1016/j.jag.2015.08.002, 2016

Schroeder, R., McDonald, C. K., Chapman, D. B., Jensen, K., Podest, E., Tessler, D. Z., Bohn, J. T., and Zimmermann, R.: Development and Evaluation of a Multi-Year Fractional Surface Water Data Set Derived from Active/Passive Microwave Remote Sensing Data, Remote Sens., 7, https://doi.org/10.3390/rs71215843, 2015

Scipal, K., Naeimi, V., and Hasenauer, S.: Definition of Quality Flags, Vienna University of Technology, Vienna, 2005.

Stocker, B. D., Zscheischler, J., Keenan, T. F., Prentice, I. C., Seneviratne, S. I., and Peñuelas, J.: Drought impacts on terrestrial primary production underestimated by satellite monitoring, Nat. Geosci., 12, 264-270, https://doi.org/10.1038/s41561-019-0318-6, 2019

Yao, P., Lu, H., Yue, S., Yang, F., Lyu, H., Yang, K., McColl, K. A., Gianotti, D., and ENTekhabi, D.: Estimating Surface Soil Moisture from AMSR2 Tb with Artificial Neural Network Method and SMAP Products, IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 6998-7001, https://doi.org/10.1109/IGARSS.2019.8898152, 2019.

Yao, P., Shi, J., Zhao, T., Lu, H., and Al-Yaari, A.: Rebuilding Long Time Series Global Soil Moisture Products Using the Neural Network Adopting the Microwave Vegetation Index, Remote Sens., 9, https://doi.org/10.3390/rs9010035, 2017

Ye, N., Walker, J. P., Guerschman, J., Ryu, D., and Gurney, R. J.: Standing water effect on soil moisture retrieval from L-band passive microwave observations, Remote Sens. Environ., 169, 232-242, https://doi.org/10.1016/j.rse.2015.08.013, 2015

Zhao, Y., Peth, S., Wang, X. Y., Lin, H., and Horn, R.: Controls of surface soil moisture spatial patterns and their temporal stability in a semi-arid steppe, Hydrol. Process., 24, 2507-2519, https://doi.org/10.1002/hyp.7665, 2010

To Reviewer #3:

We thank referee#3 for the valuable comments that will help us in improving the quality and readably of the manuscript. We have carefully revised the MS following your comments and suggestions. We provide a detailed response to the Referee's comments in the Supplement.

****

Reviewer 3 General comment: The authors used an iterative neural network approach to produce a new satellite-based soil moisture dataset using 11 microwave soil moisture products, using SMAP data for training and ISMN database for validation. The approach is quite original and efficient resulting is a improvement in the accuracy of the spatio-temporal patterns at the global scale, and at a 0.1 degree resolution. However, the manuscript will need to be improved before acceptance, in its structure, clarity and tone.

Response: Thank you for your positive comments on our work. We have adjusted the article structure, revised the Method part to make it clearer, and modified the tone of the expressions in the comparison of our product against other products. In addition, we have revised the figures and tables following each of the comment. Please find the details in the responses to the following comments.

Reviewer 3 Major comment 1: The introduction would need to be improved. Several statements need to be supported by existing literature, others would need to be clarified. Finally the introduction would need to end with a brief description of the approach used in the study and how this approach will address the three major concerns raised from existing soil moisture products. See detailed comments below for details.

Response: We have carefully addressed all the raised problem, including adding some more references, clarifying the confusing phrases, and briefly introducing the approach on how it addressed the three major concerns. Please see our responses to Specific comment 3~11 for details. Thank you for these nice suggestions!

Reviewer 3 Major comment 2: The tone of the manuscript when referring to the new product and to past studies is not always appropriate. For instance, stating that the present product is “superior to the existing products” is useless, not informative and condescending. I would encourage the authors to rather explain how their product is an improvement to the global estimation of soil moisture, without necessarily condemn other products. In the result section, while nonlinearities between estimate and in-situ soil moisture measurements are identified for other products, it is not reported for the author’s product which I find quite biased.

Response: We are sorry for the inappropriate descriptions. We have corrected the descriptions in the abstract as: ‘*This new dataset, named RSSSM, is proved comparable to the in-situ surface soil moisture measurements at sites of the International Soil Moisture Network (overall R^2 and RMSE values of 0.42 and 0.087 m^3/m^3), while the overall R^2 and RMSE values for the existing products (ASCAT-SWI, GLDAS Noah, ERA5-Land, CCI/ECV and GLEAM) are within the range of 0.31~0.41 and 0.095~0.142 m^3/m^3, respectively. The advantage of RSSSM is especially obvious in arid or relatively cold areas, and during growing seasons. Moreover, the persistent high data quality as well as complete spatial coverage ensure the applicability of RSSSM to both the spatial and temporal pattern studies.*’ We have also reported the nonlinearity of the relationship between our data and in-situ measurements, following: ‘‘*However, RSSSM overestimates soil moisture when it is low, which is a problem inherited from SMAP product (Figure 4), and is a bit nonlinearly correlated with the measured values (Figure 7a).*’

Reviewer 3 Major comment 3. The validation approach is based on site specific comparison. However, soil moisture being so spatially variable within a 0.1 degree pixel, validation based on single site observations within 0.1 degree pixels can be quite meaningless. This might be particularly true when one considers that site selection for in-situ measurement is rarely motivated by representativity of the surrounding landscape, but by specific ecological reasons.

Response: We agree that although only the ‘good’ quality data records were used, due to high spatial variability in surface soil moisture, it is not very reasonable to compare the 0.1° resolution soil moisture product against the ISMN site-scale measurements. However, the currently available global-scale soil moisture products are all in coarse resolution, usually about 0.25°. To evaluate these coarser resolution products, previous studies also have to rely on the site-measured soil moisture, especially the ISMN dataset, while the validation process and the evaluation indicators are almost the same as this study (Al-Yaari et al., 2019; Albergel et al., 2012; Dorigo et al., 2015; Fernandez-Moran et al., 2017; Gao et al., 2020; Karthikeyan et al., 2017; Kerr et al., 2016; Kim et al., 2015b; Kolassa et al., 2018; Lievens et al., 2017; Zhang et al., 2019). For 29 ISMN networks used for validation in this study, 19 are dense networks (usually with multiple stations within one 0.1° pixel (Dorigo et al., 2015)), including AMMA-CATCH (Cappelaere et al., 2009; De Rosnay et al., 2009; Lebel et al., 2009; Mougin et al., 2009; Pellarin et al., 2009), BIEBRZA_S-1 (http://www.igik.edu.pl/en), BNZ-LTER (Van Cleve et al., 2015) (http://www.lter.uaf.edu/), CTP_SMTMN (Yang et al., 2013), FLUXNET-AMERIFLUX (http://ameriflux.lbl.gov/), FR_Aqui (Al-Yaari et al., 2018), HiWATER_EHWSN (Jin et al., 2014; Kang et al., 2014), HOBE (Bircher et al., 2012), HYDROL-NET_PERUGIA (Morbidelli et al., 2014), iRON (Osenga et al., 2019), MAQU (Su et al., 2011), OZNET (Smith et al., 2012; Young et al., 2008), REMEDHUS (http://campus.usal.es/~hidrus/), SASMAS (Rüdiger et al., 2007), SKKU (Hyunglok et al., 2016), SOILSCAPE (Moghaddam et al., 2010; Moghaddam et al., 2016), SWEX_POLAND (Marczewski et al., 2010), VAS (http://nimbus.uv.es/) and WSMN (http://www.aber.ac.uk/wsmn). This information has been added to Text S2. Therefore, the average of the data obtained from two or more stations within a 0.1° pixel, which was calculated in this study, can better represent the grid-scale soil moisture conditions (Gruber et al., 2020).

In addition, to avoid the errors induced by the high spatial variability of soil moisture as much as possible, we excluded the pixels with nonnegligible open water, wetland or inundated fields. In Supplementary Text- Text S2, the related details now read: '*It has been acknowledged that the scale difference between the records at ISMN sites and the 0.1° pixel-scale soil moisture data may lead to incomparability, especially for pixels with open water and inundated land (Loew, 2008). If the measurement site is located on land, away from water, yet the corresponding pixel contains much water, the pixel-scale soil moisture can be significantly higher than the site-measured values. Conversely, if the site is in or close to the open water or inundated areas, but land also exists in the pixel, the soil moisture measured at the station will be much higher than the pixel average value. Not only the absolute values are unmatchable, the temporal variations cannot be directly compared as well, because the moisture conditions of riverside (or wetland) soil and the land soil may change with precipitation differently. Therefore, the sites located in the pixels with average annual maximal water area fraction greater than 5% according to SWAMPS data are excluded (for example, some sites in wetlands in Canada)*' . We also added more explanations in the manuscript, following: 1) '*After data screening and processing (for example, in case of high spatial variability of soil moisture, we excluded the pixels with average annual maximal water area fraction greater than 5%, see Text S2), ...*' ; 2) *More than 90% of the stations are located in relatively flat areas with topographic complexity lower than 10%;* and 3) '*Hence, to make full* *use of all the good quality records, and to reduce the problem caused by the scale difference between simulation and measurement, the site-scale 10-day averaged soil moisture data are further aggregated to 0.1° pixel-scale by averaging all the data (different stations or different sensors) within the pixel (Gruber et al., 2020).*'

Reviewer 3 Specific comment 1: "This new dataset, once validated against the International Soil Moisture Network (ISMN) records, is supposed to be superior to the existing products" .

Do you mean this validation hasn't been done yet? Superior in what way?

Response: Sorry for the confusion. Following this comment, we show the situation that our product is preferred, according to the validation against site measurements, instead of the general description that our product is superior to other products. We have made the clarification as: '*This new dataset, named RSSSM, is proved comparable to the in-situ surface soil moisture measurements at sites of the International Soil Moisture Network (overall R^2 and RMSE values of 0.42 and 0.087 m^3/m^3), while the overall R^2 and RMSE values for the existing products (ASCAT-SWI, GLDAS Noah, ERA5-Land, CCI/ECV and GLEAM) are within the range of 0.31~0.41 and 0.095~0.142 m^3/m^3, respectively. The advantage of RSSSM is especially obvious in arid or relatively cold areas, and during growing seasons. Moreover, the persistent high data quality as well as complete spatial coverage ensure the applicability of RSSSM to both the spatial and temporal pattern studies.*'

Reviewer 3 Specific comment 2: "reveals that the surface moisture decline on rainless days is highest in summers over the low-latitudes but highest in winters over most mid-latitude areas." Soil moisture being so spatially variable, I find the impact of this statement quite limited – e.g. low latitude regions range from tropical/equatorial rain forests to deserts and one would expect as much differences in the sensitivity of soil moisture to precipitation between a desert and a tropical forest than between a tropical forest and a temperate prairie.

Response: Sorry for the unclear sentences. In the calculation of soil moisture decline on consecutive rainless days, because '*In desert areas, the random noise of the surface soil moisture product can hide the signal of moisture changes, while in wet areas (e.g. rainforests), 20 days without effective precipitation seldom occurs, leading to no results over most areas.*', the acquired results can only represent the condition in regions excluding deserts and rainforests. Following this comment, we have revised the sentence as: '*It also reveals that without considering the deserts and rainforests, the surface moisture decline on consecutive rainless days is highest in summers over the low-latitudes (30°S~30°N) but highest in winters over most mid-latitude areas (30°N~60°N; 30°S~60°S)*'.

Reviewer 3 Specific comment 3: "L47: "due to various disturbances": what type of disturbances?

Response: Following your comment, we have revised it as '*due to various disturbances from for example, high vegetation cover, open water fraction and complex topography (Draper et al., 2012; Fan et al., 2020; Ye et al., 2015).*' for better clarity.

Reviewer 3 Specific comment 4: "L49: " Although new sensors, SMOS : : :." -> "Although new sensors such as SMOS: : :"

Response: We have revised it accordingly.

Reviewer 3 Specific comment 5: L 50: "better penetrability" -> please be more specific: what depth?

Response: Sorry for the unclear expression. It does not indicate the nominal soil depth of the microwave soil moisture. We have revised it to '*... because L-band microwaves (1~2 GHz) can better penetrate the vegetation canopy*', by referring to (Piles et al., 2018).

Reviewer 3 Specific comment 6: L66: "Because the temporal variation in soil moisture is often better captured by model simulations than remote sensing inversions": please include a reference that support this statement. L67: "CCI may undesirably combine the disadvantages of both." Be more specific here (low accuracy of temporal variations from remote sensing products and low spatial accuracy from model simulations – am I right?). And please include another reference here for this second statement.

Response: Sorry for the arbitrary statements. We agree that there is no strong evidence supporting the claim that the temporal variation in soil moisture is better captured by model simulations than remote sensing inversions. Following this comment, we have deleted these two sentences in the Introduction part of the manuscript. Thank you for reminding!

Reviewer 3 Specific comment 7: L70: “are assimilated instead”: instead of what? this sentence is not clear.

Response: Following this comment, we have revised the sentence as: ‘*Currently, anomalies of CCI soil moisture (the deviations to the seasonal climatology that indicate whether the soil moisture at a time point is more humid or drier than the multi-year average) are assimilated instead of the original CCI time series (Martens et al., 2017).*’

Reviewer 3 Specific comment 8: L85: “Among these three approaches, machine learning proves to be probably the best choice” based on what criteria – again, please be more specific

Response: Following this comment, we revised the sentence as: ‘*Among these three approaches, machine learning proves to be probably the best choice according to the connection between precipitation and the changes in soil moisture, evaluated through a data assimilation technique, and triple collocation analysis result (Van der Schalie et al., 2018).*’

Reviewer 3 Specific comment 9: “L102: “substantial success has not been achieved yet.” This is a rather strong and yet vague statement that denies the merits of a large body of research. Please remove this statement.

Response: Sorry for the confusion. We have revised the sentence as: ‘*To be concluded, while previous studies have focused on developing long-term satellite-based surface moisture using machine learning, there remain some major concerns that need to be solved. 1)…*’.

Reviewer 3 Specific comment 10: “L102/103:” the high-quality microwave observations are not fully utilized”: this is not clear from the literature review – please develop this point in earlier sections of the introduction (i.e. in what way high-quality microwave observations haven’t been fully utilized, and how the authors are proposing to utilize them more efficiently).

Response: To avoid potential misunderstanding, we have revised the sentence as: ‘*1) the microwave observations from only at most three sensors are utilized, leading to large temporal and spatial gaps, and limited training efficiency*’. This point has been illustrated in the earlier sections of the introduction: '*Global long-term observational-based soil moisture has been developed recently by building a neural network between the SMOS product and the Tb of AMSRE (2003~September 2011) and AMSR2 (July 2012~2015) (Yao et al., 2017)…. The gap between the temporal spans of AMSRE and AMSR2, and the lack of SMOS data in Asia resulted in large quantities of missing data.*'

Reviewer 3 Specific comment 11: "L106-107: This statement should be removed from the introduction section. - this is rather a concluding statement. Instead please describe your approach in a couple sentences and how this approach addresses the three major concerns identified.

Response: Thank you for your advice. We have re-written the paragraph as: '*… there remain some major concerns that need to be solved. 1) the microwave observations from only at most three sensors are utilized, leading to large temporal and spatial gaps, and limited training efficiency; 2) it remains unclear which environmental factors should be incorporated as ancillary inputs, and why; and 3) the training designed for soil moisture estimation at global scale ought to be more complex than that for only a specific region to ensure a satisfying training efficiency. In this study, 11 high-quality microwave soil moisture products since 2003 are incorporated into 5 rounds of neural networks to achieve a spatially and temporally continuous simulation for 2003~2018, with as many sources of microwave observational data as possible used as predictors of each neural network. The quality impact factors of microwave soil moisture retrievals are also determined and then utilized as ancillary inputs to improve the training efficiency. Moreover, we designed localized subnetworks instead of only one global-scale neural network to account for the regional differences in training rules.*'

Reviewer 3 Specific comment 12: "Section 2.1 L110: please add citation to literature supporting this statement.

Response: In Introduction part, the best overall quality of SMAP soil moisture has been stated as follows: '*Although new sensors such as SMOS (Stillman and Zeng, 2018) and SMAP (Entekhabi et al., 2010), can produce significantly improved estimates because L-band microwaves (1~2 GHz) can better penetrate the vegetation canopy (Burgin et al., 2017; Chen et al.,* *2018; Karthikeyan et al., 2017; Kerr et al., 2016; Kim et al., 2018; Leroux et al., 2014; Stillman and Zeng, 2018), the applicability of both products is still limited. SMOS data have too much noise and too many missing values in Eurasia due to high radio frequency interference (RFI) (Oliva et al., 2012). While SMAP has the highest quality (the unbiased RMSE of the passive product can be close to its target of 0.04* m^3/m^3*) and has filtered RFI (Chen et al., 2018; Colliander et al., 2017), …*'

Following your advice, we have revised this sentence and added two new references supporting the best performance of SMAP product. It reads: '*SMAP has currently the best quality of all remote sensing-based soil moisture products (Al-Yaari et al., 2019; Liu et al., 2019)…*'.

Reviewer 3 Specific comment 13: L115-118: This sentence is too long and too complex. Please split into shorter and clearer sentences.

Response: Sorry for the too complex sentence. We have revised it to: '*However, in this study, the well-acknowledged surface soil moisture products retrieved through mature algorithms (see Figure 1) are directly applied instead of Tb. This is because: 1) the primary goal of this study is to calibrate and then fuse the existing popular microwave soil moisture products; and 2) the Tb signals at multiple bands contain too much information that is not related to soil moisture, which may weaken the training efficiency and lead to over-fitting.*'

Reviewer 3 Specific comment 14: Section 2.2 L147/160: the purpose of this argumentation is quite unclear as the environmental predictors that are selected are also important drivers of soil moisture dynamic.

Response: Sorry for the arbitrary sentence. Here, we would like to express the idea that without the incorporation of any microwave remote sensing product, these factors (LAI, topographic complexity, LST, etc.) alone may not predict surface soil moisture very well. This is because although we agree that they are somewhat related to soil moisture (e.g. usually soil moisture is limited in areas with low vegetation cover), they can hardly be considered as direct indexes on surface soil moisture content (since the relationships are rather uncertain; for example, as found in this study, in Europe, the soil moisture is low in summers when vegetation peaks). On the other hand, however, we admit that water area fraction is a direct indicator of surface soil moisture, and have corrected by adding a sentence '*In addition, water fraction is a direct indicator of surface soil moisture.*'.

We removed the unclear argument following your advice. It now reads: '*Environmental factors, including DEM, LST and vegetation cover (indicated by NDVI, MVI, etc.), were used as ancillary neural network inputs for improved soil moisture simulation (Lu et al., 2015; Qu et al., 2019; Yao et al., 2017). According to these studies, these factors alone may not predict surface soil moisture well without the incorporation of any microwave remote sensing data, which can also be justified by the contribution analysis results (Figure S1a). This is because although they are somewhat related to soil moisture (e.g. soil moisture is limited in areas with low vegetation cover in general but high in forests (McColl et al., 2017)), the relationships are rather uncertain (e.g. at smaller scales, LAI may however have a negative influence on soil moisture due to the variation in evapotranspiration (Naithani et al., 2013), or without clear impacts (Zhao et al., 2010); also, soil moisture can be either high or low in summers when vegetation peaks (Baldocchi et al., 2006; Méndez-Barroso et al., 2009)). However, these factors are quite essential due to their direct impacts on soil moisture retrieval through radiative transfer model using microwave remote sensing data (Fan et al., 2020), and are retrieval quality impact factors. The detailed explanations are: 1) …*'.

Reviewer 3 Specific comment 15: L201: since precipitation is such an important driver of soil moisture, the reasons why this variable hasn't been included as a quality impact factor should be included in the main document.

Response: We have revised the sentence as: '*The contribution analysis results (Figure S1) show that because various microwave soil moisture data have already been included, precipitation data is not an essential indicator of soil moisture, and is not utilized as a physically-based 'quality impact factor' either (see Text S1 for detailed explanations).*'

The Text S1 is a bit long, so we did not move all the information to the manuscript, but we summarized the key reasons in the revised manuscript, following your advice. The Text S1 has been revised as follows:

'*Because it takes at least 3 days for a microwave sensor to cover the whole globe, for 11% of global land, there will be only 5 or less observations for random days within a 10-day period. By taking the average of these available data, this study only focuses on the mean soil moisture condition during that 10-day period. Then, to see how much can the incorporation of* *precipitation data improve the neural network training efficiency, we calculated 10-day averaged GPM Final-Run precipitation, which can well indicate the overall precipitation water availability (the antecedent precipitation index is not used, because it must be calculated on daily scale, and the attenuation coefficient is hard to determine at global scale (Kohler and Linsley, 1951)). Taking the first primary independent neural network, NN1-1-1, as an example, we performed contribution tests on all the input features at the global scale (not for each separate zone), including 9 'quality impact factors', 4 predictor soil moisture products and precipitation - a potential ancillary soil moisture indicator. For each predictor, we added a random error that is controlled within the standard deviation of the predictor, and then the increased MSE in neural network training can indicate the relative contribution of that variable. The results (Figure S1a) show that precipitation will only contribute to 1.7% of the training efficiency, which is much lower than the contribution of any soil moisture product (the total contribution fraction of the four soil moisture products is 61.2%), and is also lower than that of most 'quality impact factor'. This suggests that various microwave soil moisture datasets together with several 'quality impact factors' of microwave soil moisture retrieval are enough to predict the training target- SMAP soil moisture, and there is no need to add precipitation as another ancillary indicator of soil moisture.*

'Quality impact factors' are defined in this study as the variables that will have a significant impact on the retrieval errors of soil moisture by microwave remote sensing (section 2.1.2). Although the relative performances of different soil moisture products is related to surface moisture condition (Kim et al., 2015a), it is found mainly due to the less vegetation in arid areas. After all, no explicit mechanism can support the idea that the retrieval errors of soil moisture are significantly influenced by water availability. Even if this is true, the soil water availability can be already indicated by the microwave soil moisture products. So, it is unreasonable to incorporate the precipitation variable as a 'quality impact factor'. On the other hand, LAI, water area fraction, LST, land use cover, tree cover fraction, non-tree vegetation area fraction, topographic complexity, and soil sand/clay fractions all have direct impacts on the microwave soil moisture retrieval errors, with solid physical mechanism (see section 2.1.2). Therefore, theoretically they should be added to the neural network, even though the land use cover type and soil sand fraction data prove to have limited contributions to NN training efficiency.

One may argue that if NARX (nonlinear autoregressive with external input) is applied instead, in which the soil moisture in the previous 10-day period is also incorporated as a predictor, precipitation data can be very beneficial to the neural network training. This is true, because precipitation directly contributes to soil moisture increases. However, NARX is not suitable for global-scale long-term continuous soil moisture mapping, because the base map (i.e., soil moisture in the beginning of the *simulation period) is hard to determine. Moreover, in mid to high latitudes, the lack of soil moisture retrievals over frozen ground in winters will lead to missing data there in summers when soil moisture data is otherwise available. So, if NARX is adopted, we can only estimate long-term surface soil moisture in tropics and sub-tropics with air temperature consistently higher than 0 ℃. Last but not least, if the soil moisture in the previous phase and the current precipitation amount are both incorporated, they will largely conceal the role of satellite-observed signals. As shown in Figure S1b, the total contribution fraction of all four microwave soil moisture products is reduced to only 10.6%, while the roles of ASCAT, AMSR2-JAXA and AMSR2-LPRM are all negligible. Without taking full advantage of remote sensing, the simulations based on previous soil moisture and current precipitation will lead to errors over places where soil moisture gains are mostly driven by glacier melting, or in places with high levels of radiation-driven soil evaporation. The reliability of the derived soil moisture will be lower in irrigated croplands and afforestation/deforestation areas as well.*

On account of all above, precipitation data is neither included as an ancillary soil moisture indicator, nor added as a 'quality impact factor' in this study.'

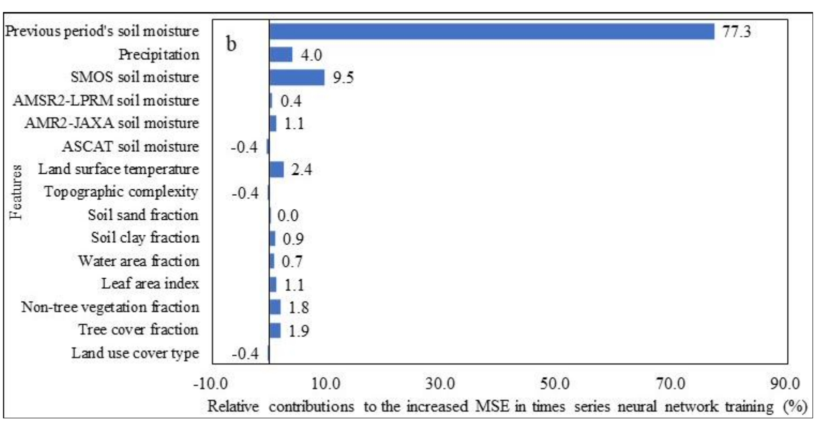


Figure R1 (Figure S1 in the revised Supplement): The roles of different input features in the soil moisture simulations based on BP neural networks and nonlinear autoregressive with external input (NARX) with microwave soil moisture products incorporated: (a) the contributions of different input features of a primary neural network: NN1-1-1, including 4 predictor soil moisture products, 9 quality impact factors of microwave soil moisture retrieval, plus 1 probable ancillary soil water *indicator: 10-day averaged precipitation, to the neural network training efficiency indicated by the increased MSE; (b) the contributions of all the input features to the training efficiency, if NN1-1-1 is changed into a NARX, in which the SMAP soil moisture for the previous period is also applied as a predictor.*

Reviewer 3 Specific comment 16: L202/203: This sentence should start this section, not end it. A table similar to table 1 but for the quality impact factors would be useful. The table would indicate the source of the data, the resolution and the temporal span for the dynamic factors.

Response: Following your suggestion, we have revised this section, putting this concluding sentence to the start. It now reads as:

'*In this study, 9 quality impact factors: LAI, water fraction, LST, land use cover, tree cover fraction, non-tree vegetation fraction, topographic complexity, and sand and clay fractions are selected and incorporated (see Figure 1). The reasons are as follows.*

Based on the two criteria above, the first environmental factor to be included is the 'vegetation factor' (i.e., vegetation water content, VWC). ... Because the leaf area index (LAI) stands for the total leaf area per unit land, which is closely related to VWC assuming a relatively stable leaf equivalent water thickness (Yilmaz et al., 2008), LAI is a suitable surrogate….'

We also added a timeline figure (Figure 1 in the revised manuscript) to show the temporal spans, sources and spatial resolution of all microwave soil moisture products, and the data for 9 quality impact factors. This could be clearer than a table.

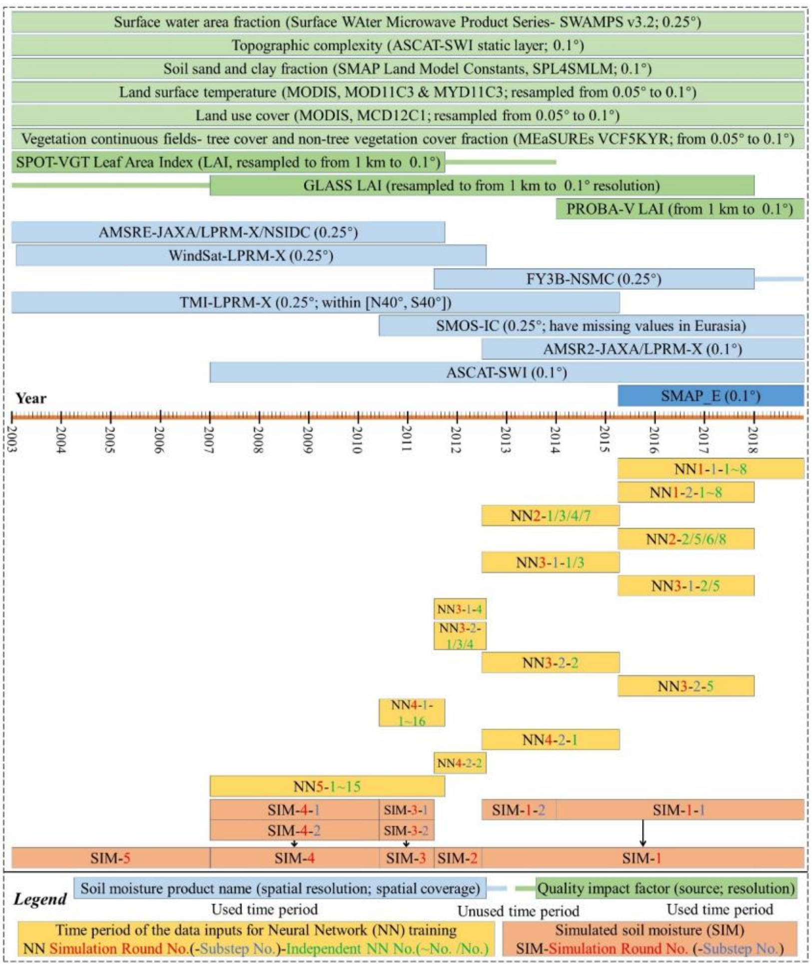


Figure R2 (Figure 1 in the revised manuscript): The timeline figure showing the time periods of the soil moisture datasets and the 'quality impact factor' products (e.g. LAI dataset) used in this study (listed above the timeline), as well as the periods of data applied for the trainings of 67 independent neural networks and the neural network simulation outputs (i.e. simulated *soil moisture) of eight substeps (listed below the timeline).*

Reviewer 3 Specific comment 17: Section 2.3 Sections 2.4, 2.5 and 2.6 should be included in section 2.3 as it details the different steps of the calculation flow. A clear justification on why a neural network approach was adopted should be included in this section.

Response: We have made the revision accordingly. The structure for Data and Method Section is now as follows:

'2 Data and Methods

2.1 Data for the production of global long-term surface soil moisture data

2.1.1 The satellite-based surface soil moisture data products

2.1.2 The quality impact factors of soil moisture retrievals

2.2 Methods for the production of global long-term surface soil moisture data

The global long-term surface soil moisture data production includes three basic parts, which are as follows. 1) Preprocessing: the production of high-quality neural network inputs; 2) neural network operation: the network training and soil moisture simulation and 3) postprocessing: the correction of potential errors or deficiencies in the soil moisture simulation outputs. …

2.2.1 Neural network design (1): localized neural networks

2.2.2 Preprocessing and postprocessing steps

2.2.3 Neural network design (2)- five rounds of simulations

2.3 Methods for the validation of surface soil moisture products

2.4 Methods for the intra-annual variation analysis of surface soil moisture'.

The justification of the neural network approach is now in the Introduction, '*Among these three approaches, machine learning proves to be probably the best choice according to the connection between precipitation and the changes in soil moisture, evaluated through a data assimilation technique, and triple collocation analysis result (Van der Schalie et al., 2018).*' In section 2.1.2, we have already mentioned the use of neural network approach in this study, '*… were used as ancillary neural network inputs for improved soil moisture simulation (Lu et al., 2015; Qu et al., 2019; Yao et al., 2017)*'.

Reviewer 3 Specific comment 18: L223: what is a hidden layer?

Response: In neural networks, hidden layer is located between the input and output layers. It is the result of nonlinear transformations of the input data through activation function, and can also be transformed into the output data. We have revised the sentence as: '*..., and the number of nodes in the hidden layer (between the input and output layers (Stinchcombe and White, 1989)) of each subnetwork is set to 7.*'

Reviewer 3 Specific comment 19: L242: reference required for the "suspicious value removal".

Response: Sorry for the confusing phrase. Actually, it is a detailed method invented by this study, to make sure only the most reliable estimates are applied as the training target of the next round of neural networks, in order to avoid significant error propagation along with the neural network round. Because multiple rounds of neural network is a characteristic of this study, this processing step was not found in previous researches (i.e., no reference). We have revised the sentence, removing the phrase '*suspicious value removal*'. It now reads: '*On the other hand, to avoid error propagation with training times by ensuring a high-quality training target for the next round's simulation, for every simulated result, we removed all the suspicious values. This preprocessing step is performed by first obtaining the maximal and minimum values of SMAP_E soil moisture in each pixel. If the simulated value is out of the range of the SMAP data during 2015~2018, the value is suspicious and is not used as the training target.*'.

Reviewer 3 Specific comment 20: L259: Here you are referring to a 1 degree pixel a presume? Please specify.

Response: We regret for the confusing sentences. We have revised the sentences to: '*For data simulation in a 0.1° pixel, the most preferable independent neural network is expected to be trained using all the available soil moisture data sources in that pixel. However, in the 1° zone where it is located, the subnetwork belonging to that preferable independent neural network may not exist due to limited valid data points (see section 2.2.1). Then, an alternative subnetwork driven by the combination of fewer soil moisture data inputs should be applied instead.*' In the revised manuscript, the word 'pixel' only stands for 0.1-

degree resolution, while the word ‘zone’ indicates 1° scale.

Reviewer 3 Specific comment 21: Please specify. L268: why 74 networks? Please explain.

Response: We have corrected it to ‘67 independent neural networks’ because the rest 7 networks are optional and not independent. There are multiple independent networks included in each round. That is because ‘*... While increasing the sources of soil moisture data inputs can be beneficial to the training efficiency, the spatial coverage of the simulation output is sacrificed because the overlapping area of more soil moisture products is smaller. After all, most products have missing data in specific regions ... To solve that dilemma, we classified all 0.1° pixels according to the available predictor soil moisture products in it over a 10-day period (for example, if there are at most four predictor soil moisture data inputs in one round, there should be 4+3+2+1=10 combinations). However, to avoid soil moisture simulation under snow or ice cover (Section 2.2.2), not all combinations are considered. Then, corresponding to each selected combination, an independent neural network is trained*.’ (Lines 269~277 in the revised manuscript). The revised explanations are clearer than the original version.

In addition, the training periods of 67 independent neural network are now shown in Figure R2 (that is Figure 1 in the revised manuscript). So, this sentence is now revised to: ‘*Based on these principles, five rounds of neural networks are designed as follows, with 8 substeps containing a total of 67 independent neural networks. The training period of each neural network and the simulation period of each substep are shown in Figure 1 (below the timeline), and the details are as follows: ...*’.

Reviewer 3 Specific comment 22: Section 2.7 Figure S2 should be included in the main manuscript as the spatial distribution of validation data is critically important to evaluate the overall strength of this validation. It is surprising that none of the Canadian sites of the Russian sites made it to the final set of sites for validation.

Response: Thank you for this advice. We have moved the figure showing the spatial distribution of validation sites to the main manuscript (shown in Figure 3).

The soil moisture measurement depth for the Russian sites are 0~10 cm (RUSWET-GRASS) or 0~20 cm (RUSWET-AGRO and RUSWET-VALDAI) (Robock et al., 2000), which do not match the nominal simulation depth (0~5 cm) of our soil moisture product (constrained by SMAP).

On the other hand, the Canadian sites are actually very limited (Dorigo et al., 2011; Dorigo et al., 2013) (https://www.geo.tuwien.ac.at/insitu/data_viewer/), and are often wetland sites with nonnegligible water fraction, which should be dropped. Previous studies also utilized very few sites in Canada and Russia (Kolassa et al., 2018; Lievens et al., 2017; Zhang et al., 2019).

In the revised manuscript, we added the information '*Accordingly, the measurements used for validation are limited to ≤ 5 cm in depth (e.g., the Russian networks were not applicable for this reason)*'.

Reviewer 3 Specific comment 23: L304/306: This is an important point. Soil moisture being so spatially variable, validation based on single site observations within 0.1 degree pixels can be quite meaningless. Especially when one considers that site selection for in situ measurement is rarely motivated by representativity of the surrounding landscape, but by specific ecological reasons.

Response: Thanks for careful thinking. Because grid-scale soil moisture measurements are unavailable, almost all the recent studies rely on the in-situ soil moisture records for the validation of remote sensing soil moisture products. Moreover, we have taken the average of the data at multiple stations in a 0.1° grid, and also excluded the grids with potentially high spatial heterogeneity of surface soil moisture. Please see the response to Major comment 3 for details.

Reviewer 3 Specific comment 24: Section 2.8 L335: "probably the best choice for periodic function fitting" : please support this statement by adequate reference to literature.

Response: We have revised it to '*Fourier functions can characterize the intra-annual variation well (Brooks et al., 2012; Hermance et al., 2007). Therefore, for the remaining areas …, we fitted the intra-annual cycle of soil moisture using the Fourier function*'.

Reviewer 3 Specific comment 25: L344/345: I don't understand this argument. Why restricting this analysis to 10 consecutive rainless days and not the whole range of 10-days sum of precipitation?

Response: This study would like to explore the impact of dry periods (20 consecutive days without effective rainfall) on the surface soil moisture in different areas. Moreover, without effective precipitation, the surface soil moisture changes are mainly driven by evaporation and deep percolation, and thus should be negative. This is simpler and can help us exclude the unreliable soil moisture values from the analysis. However, if we consider the sum of precipitation over 10 days, the surface soil moisture changes will be rather complex, and there would be much more erroneous data included, leading to unreliable analysis results. We have revised the sentence to: '*Because RSSSM indicates the average soil moisture condition during every 10-day period, we studied the surface soil moisture decline after 20 consecutive days (i.e., two adjacent 10-day periods) without effective precipitation to explore the impact of dry periods on surface soil moisture.*'.

Reviewer 3 Specific comment 26: Tables 3 to 8 could be synthesized into only two tables: one for temporal accuracy assessment and one for spatial accuracy assessment for the three products comparisons with SIM. Similarly, it would be nice to have figures 3, 4, 6, 7, 9 and 10 summarized in 2 figures where all four products appear (SIM, SMAP, GLDAS and ASCAT). This would facilitate comparison between products.

Response: Thank you for this advice. It should be noted that the periods corresponding to these comparisons differ from each other. SMAP has data since March of 2015, ASCAT is available from 2007, while SIM (named 'RSSSM' in the revised manuscript), GLDAS, ERA5 Land and GLEAM v3.3a all cover the whole study period (2003~2018), but with missing values in different areas. Although CCI (ECV) also covers the entire period, it lacks data in many places, especially before 2007 (Figure 13d). Therefore, the in-situ surface soil moisture measurements entering the comparison between RSSSM and SMAP are limited to March 2015~2018, whereas only the ISMN sites' data during 2007~2018 were applied for the accuracy comparison between RSSSM and ASCAT. Also, when comparing RSSSM against CCI (ECV), we only included the soil moisture records in the grids with both RSSSM and CCI data during the specific period. As we can see, the overall accuracy of RSSSM in Figure 4c (for the comparison with SMAP, during April 2015~2018) is R^2=0.46, RMSE=0.083, but in Figure 10 (for comparison against GLDAS, during 2003~2018), the overall R^2 of RSSSM is 0.42, and the RMSE is 0.087. In addition, the temporal accuracy of RSSSM in all climatic regions (Figure 5, Figure 8, Figure 11 in the revised paper), and its spatial accuracy during all seasons (Figure 6, Figure 9, Figure 12) are different among the comparisons against different soil moisture products. For that reason, if the 6 figures are combined into 2 figures, they will be too crowded, and the comparison will be not clear enough.

Following your comment, we have combined the tables as follows:

Table R1 (Table 2 in the revised manuscript): The mean and median values of the five evaluation indexes (correlation coefficient: r, RMSE, bias, unbiased RMSE (ubRMSE), and the anomalies r (A.R)) on the temporal accuracy of the surface soil moisture simulated in this study (RSSSM) and the other surface soil moisture products, when validated using the ISMN in-situ measurements. Note: 1) for the comparison of RSSSM against SMAP_E (SMAP) product, the validation period is from April 2015 to 2018; 2) for the comparison between RSSSM and ASCAT-SWI (ASCAT), the period is 2007~2018; 3) the comparison period for RSSSM and GLDAS Noah v2.1 (GLDAS), or ERA5-Land (ERA5-L), or CCI or GLEAM v3.3a (GLE-a) surface soil moisture product are 2003~2018; 4) the common period of RSSSM and GLEAM v3.3b (GLE-b) is from 2003 to September 2018.

Index	*r*		RMSE		bias		ubRMSE		A.R	
Product	RSSSM	SMAP	RSSSM	SMAP	RSSSM	SMAP	RSSSM	SMAP	RSSSM	SMAP
Mean	0.756	0.762	0.075	0.074	0.015	0.016	0.043	0.043	0.700	0.707
Median	0.795	0.798	0.067	0.066	0.009	0.013	0.043	0.043	0.720	0.744
Product	RSSSM	ASCAT	RSSSM	ASCAT	RSSSM	ASCAT	RSSSM	ASCAT	RSSSM	ASCAT
Mean	0.687	0.561	0.079	0.095	0.002	-0.007	0.047	0.062	0.627	0.554
Median	0.735	0.627	0.074	0.088	-0.001	-0.010	0.048	0.062	0.654	0.595
Product	RSSSM	GLDAS	RSSSM	GLDAS	RSSSM	GLDAS	RSSSM	GLDAS	RSSSM	GLDAS
Mean	0.689	0.613	0.080	0.091	0.001	0.028	0.047	0.051	0.620	0.519
Median	0.737	0.661	0.075	0.082	-0.002	0.029	0.048	0.049	0.661	0.567

Product	RSSSM	ERA5-L	RSSSM	ERA5-L	RSSSM	ERA5-L	RSSSM	ERA5-L	RSSSM	ERA5-L
Mean	0.689	0.734	0.080	0.112	0.001	0.082	0.047	0.050	0.620	0.648
Median	0.737	0.758	0.075	0.094	-0.002	0.073	0.048	0.049	0.661	0.672
Product	RSSSM	CCI	RSSSM	CCI	RSSSM	CCI	RSSSM	CCI	RSSSM	CCI
Mean	0.690	0.642	0.080	0.091	0.002	-0.002	0.047	0.049	0.620	0.530
Median	0.735	0.666	0.074	0.080	-0.002	0.006	0.049	0.047	0.658	0.552
Product	RSSSM	GLE-a	RSSSM	GLE-a	RSSSM	GLE-a	RSSSM	GLE-a	RSSSM	GLE-a
Mean	0.689	0.735	0.080	0.126	0.001	0.093	0.047	0.047	0.620	0.681
Median	0.737	0.771	0.075	0.119	-0.002	0.104	0.048	0.046	0.661	0.715
Product	RSSSM	GLE-b	RSSSM	GLE-b	RSSSM	GLE-b	RSSSM	GLE-b	RSSSM	GLE-b
Mean	0.688	0.729	0.080	0.117	0.001	0.077	0.047	0.046	0.618	0.670
Median	0.730	0.762	0.075	0.112	-0.002	0.091	0.048	0.045	0.659	0.705

Table R2 (Table 3 in the revised manuscript): The mean and median values of the four evaluation indexes (r, RMSE, bias and ubRMSE) on the spatial pattern accuracy of RSSSM and the other global long-term surface soil moisture products (SMAP_E, ASCAT-SWI, GLDAS Noah v2.1, ERA5-Land, CCI, GLEAM v3.3a and GLEAM v3.3b) in every 10-day period. For each pair of comparison, the evaluation indexes are for the common period of the two products, which is the same as Table 2. The abbreviations of the products are also the same as Table 2.

Index	*r*		RMSE		bias		ubRMSE	
Product	RSSSM	SMAP	RSSSM	SMAP	RSSSM	SMAP	RSSSM	SMAP
Mean	0.652	0.659	0.084	0.084	0.016	0.016	0.082	0.081
Median	0.655	0.664	0.082	0.081	0.019	0.019	0.080	0.078
Product	RSSSM	ASCAT	RSSSM	ASCAT	RSSSM	ASCAT	RSSSM	ASCAT

Mean	0.636	0.561	0.087	0.102	0.005	-0.010	0.085	0.097
Median	0.650	0.572	0.086	0.100	0.007	-0.009	0.085	0.095
Product	RSSSM	GLDAS	RSSSM	GLDAS	RSSSM	GLDAS	RSSSM	GLDAS
Mean	0.617	0.593	0.090	0.097	-0.005	0.035	0.086	0.087
Median	0.643	0.630	0.089	0.096	0.001	0.041	0.086	0.086
Product	RSSSM	ERA5-L	RSSSM	ERA5-L	RSSSM	ERA5-L	RSSSM	ERA5-L
Mean	0.616	0.575	0.090	0.125	-0.005	0.077	0.086	0.095
Median	0.641	0.633	0.089	0.125	0.001	0.082	0.086	0.092
Product	RSSSM	CCI	RSSSM	CCI	RSSSM	CCI	RSSSM	CCI
Mean	0.618	0.497	0.090	0.099	-0.004	0.003	0.086	0.093
Median	0.647	0.554	0.089	0.098	0.002	0.006	0.086	0.093
Product	RSSSM	GLE-a	RSSSM	GLE-a	RSSSM	GLE-a	RSSSM	GLE-a
Mean	0.617	0.576	0.090	0.139	-0.005	0.105	0.086	0.089
Median	0.643	0.616	0.089	0.142	0.001	0.112	0.086	0.088
Product	RSSSM	GLE-b	RSSSM	GLE-b	RSSSM	GLE-b	RSSSM	GLE-b
Mean	0.616	0.560	0.090	0.128	-0.005	0.088	0.086	0.090
Median	0.643	0.613	0.089	0.130	0.001	0.094	0.086	0.089

Reviewer 3 Specific comment 27: L380-383: However, it looks like the relationship between SIM estimates and in situ observations is nonlinear (Figure 5a). Furthermore, SIM seems to overestimate soil moisture in the lowest range (winter?) when a density of pixels is quite high. Please include these remarks in the results.

Response: Thank you for reminding. We have added a sentence accordingly. It reads: '*However, RSSSM overestimates soil moisture when it is low, which is a problem inherited from SMAP product (Figure 4), and is a bit nonlinearly correlated with the measured values (Figure 7a).*'

Reviewer 3 Specific comment 28: L393: the relationship between SIM and in situ measurements is also obviously nonlinear. Please include this remark in the text for fairness.

Response: We have revised the description accordingly. It now reads: '*While RSSSM is nonlinearly correlated with measured soil moisture, the relationship between GLDAS soil moisture and the measurements appears a bit more nonlinear, resulting in a smaller R^2 of 0.39 and higher RMSE of 0.097 for GLDAS product compared to RSSSM (R^2: 0.42; RMSE: 0.087, see Figure 10).*'

Reviewer 3 Specific comment 29: Why table S19 and Figure S7 do not appear in the main document like the other product comparison? Please move them to the main manuscript

Response: We have combined Table S17, Table S19, Table S21, Table S22 in the original Supplement to Table 2 in the revised manuscript (see Table R1). Table S18, Table S20, Table S23 and Table S24 are integrated into Table 3 as well (see Table R2). For the reasons described in the response to Reviewer 3 Specific comment 26 (the time periods and spatial extents for the comparisons of RSSSM in this study against different existing surface soil moisture products are different), the figures were not combined together to avoid over-crowdedness and unnecessary confusion. Thus, we retained Figure S7, Figure S8, etc. in the Supplement to prevent the main manuscript from too long and too complex to read.

Reviewer 3 Specific comment 30: L422-434: This belongs to the discussion section.

Response: We have moved the sentences belonging to the discussion section to the Discussion part. The revised sentences in the Result part now read: '*Next, we focus on the interannual change in data quality. According to Figure 13a~c, while the correlation coefficient for RSSSM does not vary significantly among different years, the RMSE and ubRMSE values in earlier periods are somewhat raised compared to those after 2012. Though the data quality of RSSSM can hardly be maintained as well, the degradation degree is much slighter than CCI. By comparing the spatial coverages of the 10-day scale RSSSM and CCI data (rainforests are excluded), it is shown that RSSSM covers all land surfaces except for permafrost, while coverages'* *interannual variation is also negligible throughout the entire period (the intra-annual cycles of data coverages result from the changes in frozen areas), which are preferable to CCI, whose data coverage before 2007 is limited (Figure 13d).*'

After the remove, the Discussion section has been revised as follows: '*In this study, an improved global long-term satellite-based surface soil moisture dataset, …. It is temporally-continuous during 2003~2018, and covers the whole globe except for frozen ground (CCI has limited spatial coverage before 2007, when ASCAT soil moisture is unavailable), ensuring its applicability in global long-term studies or ecosystem modelling…. The RMSE and ubRMSE values in earlier periods are somewhat higher than those after 2012, which is because: 1) five rounds of simulations were performed, with the output converted into the training target of the next round's neural networks, leading to a little error propagation as the simulation period extending to the past; and 2) the quality of microwave soil moisture data is generally lower in earlier times due to the relatively unadvanced microwave sensors with low signal-to-noise ratio (SNR). However, due to the elaborated design of the neural network set (localized networks, full use of 11 microwave soil moisture products, the determination of quality impact factors and the organization of 67 independent neural networks), high training efficiency is achieved, resulting in little amplification of noises and high maintenance of valid information during 16 years of simulation. This method turns out to be better than the simple CDF matching algorithm which may not efficiently calibrate the low-quality soil moisture data retrieved from earlier sensors.*'.

References

Al-Yaari, A., Dayau, S., Chipeaux, C., Aluome, C., Kruszewski, A., Loustau, D., and Wigneron, J. P.: The AQUI Soil Moisture Network for Satellite Microwave Remote Sensing Validation in South-Western France, Remote Sens., 10, http://doi.org/10.3390/rs10111839, 2018

Al-Yaari, A., Wigneron, J. P., Dorigo, W., Colliander, A., Pellarin, T., Hahn, S., Mialon, A., Richaume, P., Fernandez-Moran, R., Fan, L., Kerr, Y. H., and De Lannoy, G.: Assessment and inter-comparison of recently developed/reprocessed microwave satellite soil moisture products using ISMN ground-based measurements, Remote Sens. Environ., 224, 289-303, https://doi.org/10.1016/j.rse.2019.02.008, 2019

Albergel, C., de Rosnay, P., Gruhier, C., Muñoz-Sabater, J., Hasenauer, S., Isaksen, L., Kerr, Y., and Wagner, W.: Evaluation of remotely sensed and modelled soil moisture products using global ground-based in situ observations, Remote Sens. Environ., 118, 215-226, https://doi.org/10.1016/j.rse.2011.11.017, 2012

Baldocchi, D., Tang, J., and Xu, L.: How switches and lags in biophysical regulators affect spatial-temporal variation of soil respiration in an oak-grass savanna, Journal of Geophysical Research: Biogeosciences, 111, 10.1029/2005JG000063, 2006

Bircher, S., Skou, N., Jensen, K. H., Walker, J. P., and Rasmussen, L.: A soil moisture and temperature network for SMOS validation in Western Denmark, Hydrol. Earth Syst. Sci., 16, 1445-1463, http://doi.org/10.5194/hess-16-1445-2012, 2012

Brooks, E. B., Thomas, V. A., Wynne, R. H., and Coulston, J. W.: Fitting the Multitemporal Curve: A Fourier Series Approach to the Missing Data Problem in Remote Sensing Analysis, IEEE Trans. Geosci. Remote Sensing, 50, 3340-3353, http://doi.org/10.1109/TGRS.2012.2183137, 2012

Burgin, M., Colliander, A., Njoku, E., Chan, S., Cabot, F., Kerr, H. Y., Bindlish, R., Jackson, T., Entekhabi, D., and Yueh, S.: A Comparative Study of the SMAP Passive Soil Moisture Product With Existing Satellite-Based Soil Moisture Products, IEEE Trans. Geosci. Remote Sensing, 55, 2959-2971, http://doi.org/10.1109/TGRS.2017.2656859, 2017

Cappelaere, B., Descroix, L., Lebel, T., Boulain, N., Ramier, D., Laurent, J. P., Favreau, G., Boubkraoui, S., Boucher, M., Bouzou Moussa, I., Chaffard, V., Hiernaux, P., Issoufou, H. B. A., Le Breton, E., Mamadou, I., Nazoumou, Y., Oi, M., Ottlé, C., and Quantin, G.: The AMMA-CATCH experiment in the cultivated Sahelian area of south-west Niger – Investigating water cycle response to a fluctuating climate and changing environment, J. Hydrol., 375, 34-51, https://doi.org/10.1016/j.jhydrol.2009.06.021, 2009

Chen, F., Crow, W. T., Bindlish, R., Colliander, A., Burgin, M. S., Asanuma, J., and Aida, K.: Global-scale evaluation of SMAP, SMOS and ASCAT soil moisture products using triple collocation, Remote Sens. Environ., 214, 1-13, https://doi.org/10.1016/j.rse.2018.05.008, 2018

Colliander, A., Jackson, T. J., Bindlish, R., Chan, S., Das, N., Kim, S. B., Cosh, M. H., Dunbar, R. S., Dang, L., Pashaian, L., Asanuma, J., Aida, K., Berg, A., Rowlandson, T., Bosch, D., Caldwell, T., Caylor, K., Goodrich, D., al Jassar, H., Lopez-Baeza, E., Martínez-Fernández, J., González-Zamora, A., Livingston, S., McNairn, H., Pacheco, A., Moghaddam, M., Montzka, C., Notarnicola, C., Niedrist, G., Pellarin, T., Prueger, J., Pulliainen, J., Rautiainen, K., Ramos, J., Seyfried, M., Starks, P., Su, Z., Zeng, Y., van der Velde, R., Thibeault, M., Dorigo, W., Vreugdenhil, M., Walker, J. P., Wu, X., Monerris, A., O'Neill, P. E., Entekhabi, D., Njoku, E. G., and Yueh, S.: Validation of SMAP surface soil moisture products with core validation sites, Remote Sens. Environ., 191, 215-231, https://doi.org/10.1016/j.rse.2017.01.021, 2017

De Rosnay, P., Gruhier, C., Timouk, F., Baup, F., Mougin, E., Hiernaux, P., Kergoat, L., and LeDantec, V.: Multi-scale soil moisture measurements at the Gourma meso-scale site in Mali, J. Hydrol., 375, 241-252, https://doi.org/10.1016/j.jhydrol.2009.01.015, 2009

Dorigo, W. A., Gruber, A., De Jeu, R. A. M., Wagner, W., Stacke, T., Loew, A., Albergel, C., Brocca, L., Chung, D., Parinussa, R. M., and Kidd, R.: Evaluation of the ESA CCI soil moisture product using ground-based observations, Remote Sens. Environ., 162, 380-395, https://doi.org/10.1016/j.rse.2014.07.023, 2015

Dorigo, W. A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A., Drusch, M., Mecklenburg, S., van Oevelen, P., Robock, A., and Jackson, T.: The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements, Hydrol. Earth Syst. Sci., 15, 1675-1698, https://doi.org/10.5194/hess-15-1675-2011, 2011

Dorigo, W. A., Xaver, A., Vreugdenhil, M., Gruber, A., Hegyiová, A., Sanchis-Dufau, A. D., Zamojski, D., Cordes, C., Wagner, W., and Drusch, M.: Global Automated Quality Control of In Situ Soil Moisture Data from the International Soil Moisture Network, Vadose Zone J., 12, https://doi.org/10.2136/vzj2012.0097, 2013

Draper, C. S., Reichle, R. H., De Lannoy, G. J. M., and Liu, Q.: Assimilation of passive and active microwave soil moisture retrievals, Geophys. Res. Lett., 39, http://doi.org/10.1029/2011GL050655, 2012

Entekhabi, D., Njoku, E., O'neill, P., Kellogg, K. H., Crow, W., Edelstein, W. N., Entin, J. K., Goodman, S. D., Jackson, T., Johnson, F. M., Kimball, J., Piepmeier, J., Koster, R. D., Martin, E., McDonald, C. K., Moghaddam, M., Moran, M. S., Reichle, R., Shi, J. C., Spencer, D., Thurman, S. W., Tsang, L., and Zyl, J. V.: The Soil Moisture Active Passive (SMAP) Mission, Proc. IEEE, 98, 704-716, https://doi.org/10.1109/JPROC.2010.2043918, 2010

Fan, X., Liu, Y., Gan, G., and Wu, G.: SMAP underestimates soil moisture in vegetation-disturbed areas primarily as a result of biased surface temperature data, Remote Sens. Environ., 247, 111914, https://doi.org/10.1016/j.rse.2020.111914, 2020

Fernandez-Moran, R., Wigneron, J. P., De Lannoy, G., Lopez-Baeza, E., Parrens, M., Mialon, A., Mahmoodi, A., Al-Yaari, A., Bircher, S., Al Bitar, A., Richaume, P., and Kerr, Y.: A new calibration of the effective scattering albedo and soil roughness parameters in the SMOS SM retrieval algorithm, Int. J. Appl. Earth Obs. Geoinf., 62, 27-38, https://doi.org/10.1016/j.jag.2017.05.013, 2017

Gao, L., Sadeghi, M., and Ebtehaj, A.: Microwave retrievals of soil moisture and vegetation optical depth with improved resolution using a combined constrained inversion algorithm: Application for SMAP satellite, Remote Sens. Environ., 239, 111662, https://doi.org/10.1016/j.rse.2020.111662, 2020

Gruber, A., De Lannoy, G., Albergel, C., Al-Yaari, A., Brocca, L., Calvet, J. C., Colliander, A., Cosh, M., Crow, W., Dorigo, W., Draper, C., Hirschi, M., Kerr, Y., Konings, A., Lahoz, W., McColl, K., Montzka, C., Muñoz-Sabater, J., Peng, J., Reichle, R., Richaume, P., Rüdiger, C., Scanlon, T., van der Schalie, R., Wigneron, J. P., and Wagner, W.: Validation practices for satellite soil moisture retrievals: What are (the) errors?, Remote Sens. Environ., 244, 111806, https://doi.org/10.1016/j.rse.2020.111806, 2020

Hermance, J. F., Jacob, R. W., Bradley, B. A., and Mustard, J. F.: Extracting Phenological Signals From Multiyear AVHRR NDVI Time Series: Framework for Applying High-Order Annual Splines With Roughness Damping, IEEE Trans. Geosci. Remote Sensing, 45, 3264-3276, http://doi.org/10.1109/TGRS.2007.903044, 2007

Hyunglok, K., Wooyeon, S., Seongkyun, K., and Minha, C.: Construction and estimation of soil moisture site with FDR and COSMIC-ray (SM-FC) sensors for calibration/validation of satellite-based and COSMIC-ray soil moisture products in Sungkyunkwan university, South Korea, Journal of Korea Water Resources Association, 49, 133-144, http://doi.org/10.3741/JKWRA.2016.49.2.133, 2016

Jin, R., Li, X., Yan, B., Li, X., Luo, W., Ma, M., Guo, J., Kang, J., Zhu, Z., and Zhao, S.: A Nested Ecohydrological Wireless Sensor Network for Capturing the Surface Heterogeneity in the Midstream Areas of the Heihe River Basin, China, IEEE Geosci. Remote Sens. Lett., 11, 2015-2019, https://doi.org/10.1109/LGRS.2014.2319085, 2014

Kang, J., Li, X., Jin, R., Ge, Y., Wang, J., and Wang, J.: Hybrid Optimal Design of the Eco-Hydrological Wireless Sensor Network in the Middle Reach of the Heihe River Basin, China, Sensors, 14, https://doi.org/10.3390/s141019095, 2014

Karthikeyan, L., Pan, M., Wanders, N., Kumar, D. N., and Wood, E. F.: Four decades of microwave satellite soil moisture observations: Part 2. Product validation and inter-satellite comparisons, Adv. Water Resour., 109, 236-252, https://doi.org/10.1016/j.advwatres.2017.09.010, 2017

Kerr, Y. H., Al-Yaari, A., Rodriguez-Fernandez, N., Parrens, M., Molero, B., Leroux, D., Bircher, S., Mahmoodi, A., Mialon, A., Richaume, P., Delwart, S., Al Bitar, A., Pellarin, T., Bindlish, R., Jackson, T. J., Rüdiger, C., Waldteufel, P., Mecklenburg, S., and Wigneron, J. P.: Overview of SMOS performance in terms of global soil moisture monitoring after six years in operation, Remote Sens. Environ., 180, 40-63, https://doi.org/10.1016/j.rse.2016.02.042, 2016

Kim, H., Parinussa, R., Konings, A. G., Wagner, W., Cosh, M. H., Lakshmi, V., Zohaib, M., and Choi, M.: Global-scale assessment and combination of SMAP with ASCAT (active) and AMSR2 (passive) soil moisture products, Remote Sens. Environ., 204, 260-275, https://doi.org/10.1016/j.rse.2017.10.026, 2018

Kim, S., Liu, Y. Y., Johnson, F. M., Parinussa, R. M., and Sharma, A.: A global comparison of alternate AMSR2 soil moisture products: Why do they differ?, Remote Sens. Environ., 161, 43-62, https://doi.org/10.1016/j.rse.2015.02.002, 2015a

Kim, S., Parinussa, R. M., Liu, Y. Y., Johnson, F. M., and Sharma, A.: A framework for combining multiple soil moisture retrievals based on maximizing temporal correlation, Geophys. Res. Lett., 42, 6662-6670, http://doi.org/10.1002/2015GL064981, 2015b

Kohler, M. A. and Linsley, R. K.: Predicting the runoff from storm rainfall, US Department of Commerce, Weather Bureau, 1951.

Kolassa, J., Reichle, R. H., Liu, Q., Alemohammad, S. H., Gentine, P., Aida, K., Asanuma, J., Bircher, S., Caldwell, T., Colliander, A., Cosh, M., Holifield Collins, C., Jackson, T. J., Martínez-Fernández, J., McNairn, H., Pacheco, A., Thibeault, M., and Walker, J. P.: Estimating surface soil moisture from SMAP observations using a Neural Network technique, Remote Sens. Environ., 204, 43-59, https://doi.org/10.1016/j.rse.2017.10.045, 2018

Lebel, T., Cappelaere, B., Galle, S., Hanan, N., Kergoat, L., Levis, S., Vieux, B., Descroix, L., Gosset, M., Mougin, E., Peugeot, C., and Seguis, L.: AMMA-CATCH studies in the Sahelian region of West-Africa: An overview, J. Hydrol., 375, 3-13, https://doi.org/10.1016/j.jhydrol.2009.03.020, 2009

Leroux, D. J., Kerr, Y. H., Bitar, A. A., Bindlish, R., Jackson, T. J., Berthelot, B., and Portet, G.: Comparison Between SMOS, VUA, ASCAT, and ECMWF Soil Moisture Products Over Four Watersheds in U.S, IEEE Trans. Geosci. Remote Sensing, 52, 1562-1571, https://doi.org/10.1109/TGRS.2013.2252468, 2014

Lievens, H., Martens, B., Verhoest, N. E. C., Hahn, S., Reichle, R. H., and Miralles, D. G.: Assimilation of global radar backscatter and radiometer brightness temperature observations to improve soil moisture and land evaporation estimates, Remote Sens. Environ., 189, 194-210, https://doi.org/10.1016/j.rse.2016.11.022, 2017

Liu, J., Chai, L., Lu, Z., Liu, S., Qu, Y., Geng, D., Song, Y., Guan, Y., Guo, Z., Wang, J., and Zhu, Z.: Evaluation of SMAP, SMOS-IC, FY3B, JAXA, and LPRM Soil Moisture Products over the Qinghai-Tibet Plateau and Its Surrounding Areas, Remote Sens., 11, http://doi.org/10.3390/rs11070792, 2019

Loew, A.: Impact of surface heterogeneity on surface soil moisture retrievals from passive microwave data at the regional scale: The Upper Danube case, Remote Sens. Environ., 112, 231-248, https://doi.org/10.1016/j.rse.2007.04.009, 2008

Lu, Z., Chai, L., Ye, Q., and Zhang, T.: Reconstruction of time-series soil moisture from AMSR2 and SMOS data by using recurrent nonlinear autoregressive neural networks, 26-31 July 2015 2015, 980-983, https://doi.org/10.1109/IGARSS.2015.7325932.

Marczewski, W., Slominski, J., Slominska, E., Usowicz, B., Usowicz, J., Romanov, S., Maryskevych, O., Nastula, J., and Zawadzki, J.: Strategies for validating and directions for employing SMOS data, in the Cal-Val project SWEX (3275) for wetlands, Hydrol. Earth Syst. Sci., 2010, 7007-7057, http://doi.org/10.5194/hessd-7-7007-2010, 2010

Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A., and Verhoest, N. E. C.: GLEAM v3: satellite-based land evaporation and root-zone soil moisture, Geosci. Model Dev., 10, 1903-1925, https://doi.org/10.5194/gmd-10-1903-2017, 2017

McColl, K. A., Alemohammad, S. H., Akbar, R., Konings, A. G., Yueh, S., and Entekhabi, D.: The global distribution and dynamics of surface soil moisture, Nat. Geosci., 10, 100, https://doi.org/10.1038/ngeo2868, 2017

Méndez-Barroso, L. A., Vivoni, E. R., Watts, C. J., and Rodríguez, J. C.: Seasonal and interannual relations between precipitation, surface soil moisture and vegetation dynamics in the North American monsoon region, J. Hydrol., 377, 59-70, https://doi.org/10.1016/j.jhydrol.2009.08.009, 2009

Moghaddam, M., Entekhabi, D., Goykhman, Y., Li, K., Liu, M., Mahajan, A., Nayyar, A., Shuman, D., and Teneketzis, D.: A Wireless Soil Moisture Smart Sensor Web Using Physics-Based Optimal Control: Concept and Initial Demonstrations, IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens., 3, 522-535, https://doi.org/10.1109/JSTARS.2010.2052918, 2010

Moghaddam, M., Silva, A. R., Clewley, D., Akbar, R., Hussaini, S. A., Whitcomb, J., Devarakonda, R., Shrestha, R., Cook, R. B., Prakash, G., Santhana Vannan, S. K., and Boyer, A. G.: Soil Moisture Profiles and Temperature Data from SoilSCAPE Sites, USA., ORNL DAAC, Oak Ridge, Tennessee, USA., https://doi.org/10.3334/ORNLDAAC/1339, 2016

Morbidelli, R., Saltalippi, C., Flammini, A., Rossi, E., and Corradini, C.: Soil water content vertical profiles under natural conditions: matching of experiments and simulations by a conceptual model, Hydrol. Process., 28, 4732-4742, https://doi.org/10.1002/hyp.9973, 2014

Mougin, E., Hiernaux, P., Kergoat, L., Grippa, M., de Rosnay, P., Timouk, F., Le Dantec, V., Demarez, V., Lavenu, F., Arjounin, M., Lebel, T., Soumaguel, N., Ceschia, E., Mougenot, B., Baup, F., Frappart, F., Frison, P. L., Gardelle, J., Gruhier, C., Jarlan, L., Mangiarotti, S., Sanou, B., Tracol, Y., Guichard, F., Trichon, V., Diarra, L., Soumaré, A., Koité, M., Dembélé, F., Lloyd, C., Hanan, N. P., Damesin, C., Delon, C., Serça, D., Galy-Lacaux, C., Seghieri, J., Becerra, S., Dia, H., Gangneron, F., and Mazzega, P.: The AMMA-CATCH Gourma observatory site in Mali: Relating climatic variations to changes in vegetation, surface hydrology, fluxes and natural resources, J. Hydrol., 375, 14-33, https://doi.org/10.1016/j.jhydrol.2009.06.045, 2009

Naithani, K. J., Baldwin, D. C., Gaines, K. P., Lin, H., and Eissenstat, D. M.: Spatial Distribution of Tree Species Governs the Spatio-Temporal Interaction of Leaf Area Index and Soil Moisture across a Forested Landscape, PLOS ONE, 8, e58704, 10.1371/journal.pone.0058704, 2013

Oliva, R., Daganzo, E., Kerr, Y. H., Mecklenburg, S., Nieto, S., Richaume, P., and Gruhier, C.: SMOS Radio Frequency Interference Scenario: Status and Actions Taken to Improve the RFI Environment in the 1400–1427-MHz Passive Band, IEEE Trans. Geosci. Remote Sensing, 50, 1427-1439, https://doi.org/10.1109/TGRS.2012.2182775, 2012

Osenga, E. C., Arnott, J. C., Endsley, K. A., and Katzenberger, J. W.: Bioclimatic and Soil Moisture Monitoring Across Elevation in a Mountain Watershed: Opportunities for Research and Resource Management, Water Resour. Res., 55, 2493-2503, https://doi.org/10.1029/2018WR023653, 2019

Pellarin, T., Laurent, J. P., Cappelaere, B., Decharme, B., Descroix, L., and Ramier, D.: Hydrological modelling and associated microwave emission of a semi-arid region in South-western Niger, J. Hydrol., 375, 262-272, https://doi.org/10.1016/j.jhydrol.2008.12.003, 2009

Piles, M., Schalie, R. v. d., Gruber, A., Muñoz-Marí, J., Camps-Valls, G., Mateo-Sanchis, A., Dorigo, W., and Jeu, R. d.: Global Estimation of Soil Moisture Persistence with L and C-Band Microwave Sensors, 22-27 July 2018 2018, 8259-8262, 10.1109/IGARSS.2018.8518161.

Qu, Y., Zhu, Z., Chai, L., Liu, S., Montzka, C., Liu, J., Yang, X., Lu, Z., Jin, R., Li, X., Guo, Z., and Zheng, J.: Rebuilding a Microwave Soil Moisture Product Using Random Forest Adopting AMSR-E/AMSR2 Brightness Temperature and SMAP over the Qinghai–Tibet Plateau, China, Remote Sens., 11, https://doi.org/10.3390/rs11060683, 2019

Robock, A., Vinnikov, K. Y., Srinivasan, G., Entin, J. K., Hollinger, S. E., Speranskaya, N. A., Liu, S., and Namkhai, A.: The Global Soil Moisture Data Bank, Bull. Amer. Meteorol. Soc., 81, 1281-1300, http://doi.org/10.1175/1520-0477(2000)081<1281:TGSMDB>2.3.CO;2, 2000

Rüdiger, C., Hancock, G., Hemakumara, H. M., Jacobs, B., Kalma, J. D., Martinez, C., Thyer, M., Walker, J. P., Wells, T., and Willgoose, G. R.: Goulburn River experimental catchment data set, Water Resour. Res., 43, https://doi.org/10.1029/2006WR005837, 2007

Smith, A. B., Walker, J. P., Western, A. W., Young, R. I., Ellett, K. M., Pipunic, R. C., Grayson, R. B., Siriwardena, L., Chiew, F. H. S., and Richter, H.: The Murrumbidgee soil moisture monitoring network data set, Water Resour. Res., 48, https://doi.org/10.1029/2012WR011976, 2012

Stillman, S. and Zeng, X.: Evaluation of SMAP Soil Moisture Relative to Five Other Satellite Products Using the Climate Reference Network Measurements Over USA, IEEE Trans. Geosci. Remote Sensing, 56, 6296-6305, https://doi.org/10.1109/TGRS.2018.2835316, 2018

Stinchcombe and White: Universal approximation using feedforward networks with non-sigmoid hidden layer activation functions, 1989 1989, 613-617 vol.611, http://doi.org/10.1109/IJCNN.1989.118640.

Su, Z., Wen, J., Dente, L., van der Velde, R., Wang, L., Ma, Y., Yang, K., and Hu, Z.: The Tibetan Plateau observatory of plateau scale soil moisture and soil temperature (Tibet-Obs) for quantifying uncertainties in coarse resolution satellite and model products, Hydrol. Earth Syst. Sci., 15, 2303-2316, https://doi.org/10.5194/hess-15-2303-2011, 2011

Van Cleve, K., Chapin, F. S. S., and Ruess, R. W.: Bonanza Creek Long Term Ecological Research Project Climate Database - University of Alaska Fairbanks., 2015

Van der Schalie, R., De Jeu, R., Parinussa, R., Rodríguez-Fernández, N., Kerr, Y., Al-Yaari, A., Wigneron, J.-P., and Drusch, M.: The Effect of Three Different Data Fusion Approaches on the Quality of Soil Moisture Retrievals from Multiple Passive Microwave Sensors, Remote Sens., 10, https://doi.org/10.3390/rs10010107, 2018

Yang, K., Qin, J., Zhao, L., Chen, Y., Tang, W., Han, M., Lazhu, Chen, Z., Lv, N., Ding, B., Wu, H., and Lin, C.: A Multiscale Soil Moisture and Freeze–Thaw Monitoring Network on the Third Pole, Bull. Amer. Meteorol. Soc., 94, 1907-1916, https://doi.org/10.1175/BAMS-D-12-00203.1, 2013

Yao, P., Shi, J., Zhao, T., Lu, H., and Al-Yaari, A.: Rebuilding Long Time Series Global Soil Moisture Products Using the Neural Network Adopting the Microwave Vegetation Index, Remote Sens., 9, https://doi.org/10.3390/rs9010035, 2017

Ye, N., Walker, J. P., Guerschman, J., Ryu, D., and Gurney, R. J.: Standing water effect on soil moisture retrieval from L-band passive microwave observations, Remote Sens. Environ., 169, 232-242, https://doi.org/10.1016/j.rse.2015.08.013, 2015

Yilmaz, M. T., Hunt, E. R., and Jackson, T. J.: Remote sensing of vegetation water content from equivalent water thickness using satellite imagery, Remote Sens. Environ., 112, 2514-2522, https://doi.org/10.1016/j.rse.2007.11.014, 2008

Young, R., Walker, J., Yeoh, N., Smith, A., Ellett, K., Merlin, O., and Western, A.: Soil Moisture and Meteorological Observations From the Murrumbidgee Catchment, 2008.

Zhang, R., Kim, S., and Sharma, A.: A comprehensive validation of the SMAP Enhanced Level-3 Soil Moisture product using ground measurements over varied climates and landscapes, Remote Sens. Environ., 223, 82-94, https://doi.org/10.1016/j.rse.2019.01.015, 2019

Zhao, Y., Peth, S., Wang, X. Y., Lin, H., and Horn, R.: Controls of surface soil moisture spatial patterns and their temporal stability in a semi-arid steppe, Hydrol. Process., 24, 2507-2519, 10.1002/hyp.7665, 2010

An improved dataset of satellite observation-based global surface soil moisture covering 2003~2018 (RSSSM)

Yongzhe Chen[1,2], Xiaoming Feng[1,*], Bojie Fu[1,2]

[1] State Key Laboratory of Urban and Regional Ecology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, PR China.

[2] University of Chinese Academy of Sciences, Beijing 100049, PR China.

Correspondence to: Xiaoming Feng(fengxm@rcees.ac.cn)

Abstract. Soil moisture is an important variable linking the atmosphere and terrestrial ecosystems. However, long-term satellite monitoring of surface soil moisture is still lacking at the global scale. In this study, we conducted data calibration and data fusion of 11 well-acknowledged microwave remote sensing soil moisture products since 2003 through a neural network approach, with SMAP soil moisture data applied as the primary training target. The training efficiency was high (R^2 =0.95) due to the selection of 9 quality impact factors of microwave soil moisture products and the elaborate organization structure of multiple neural networks (5 rounds of simulations; 8 substeps; 67 independent neural networks; and more than one million zonal subnetworks). We achieved global satellite monitoring of surface soil moisture during 2003~2018 at 0.1° resolution. The temporal resolution is approximately 10 days, or to be specific, there are 3 data records within a month, for days 1~10, 11~20 and from 21 to the last day of that month. This new dataset, named RSSSM, is proved comparable to the in-situ surface soil moisture measurements at the International Soil Moisture Network sites (overall R^2 and RMSE values of 0.42 and 0.087 m^3/m^3), while the overall R^2 and RMSE values for the existing products (ASCAT-SWI, GLDAS Noah, ERA5-Land, CCI/ECV and GLEAM) are within the range of 0.31~0.41 and 0.095~0.142 m^3/m^3, respectively. The advantage of RSSSM is especially obvious in arid or relatively cold areas, and during growing seasons. Moreover, the persistent high data quality as well as complete spatial coverage ensure the applicability of RSSSM to studies on both the spatial and temporal patterns. Our new data suggest an increase in the global mean surface soil moisture. These data also reveal that without considering the deserts and rainforests, the surface moisture decline on consecutive rainless days is highest in summers over the low latitudes (30°S~30°N) but highest in winters over most midlatitude areas (30°N~60°N; 30°S~60°S). Notably, the error propagation is well controlled with the extension of the simulation period to the past, indicating that the data fusion algorithm proposed in this study will be even more meaningful in the future when more advanced microwave sensors are become operational. The dataset can be accessed at https://doi.pangaea.de/10.1594/PANGAEA.912597 (Chen, 2020).

删除了: new

删除了: -based …soil moisture products since 2003 through a neural network approach, with SMAP soil moisture data applied as the fundamental

删除了: proves to be…as high (R^2 =0.95) due to the selection of 9 quality impact factors of microwave soil moisture products and the elaborated

删除了: various

删除了: 74 …7 independent neural networks; and >10^6…ore than one million regional

删除了: , respectively

删除了: .

设置了格式: 字体: 非倾斜

设置了格式: 字体: 非倾斜

删除了: sites of

设置了格式

删除了: studies

设置了格式: 字体: 非倾斜

删除了: This new dataset, once validated against the International Soil Moisture Network (ISMN) records, is supposed to be superior to the existing products (ASCAT-

删除了: It…ur new data

删除了: s

删除了: , and

删除了: It

删除了: s…that without considering the deserts and

删除了: areas

删除了: is well controlled… indicating that the data fusion

删除了: 分页符

1 Introduction

Soil moisture plays an important role in modulating the exchange of water, carbon and energy between the land surface and atmosphere, as well as linking the global water, carbon and energy cycles (Dorigo et al., 2012; Karthikeyan et al., 2017a). Soil moisture has been endorsed by the Global Climate Observing System (GCOS) as an essential climate variable (Bojinski et al., 2014), as it is probably the best indicator of ecological droughts (Martínez-Fernández et al., 2016; Samaniego et al., 2018). However, due to the large uncertainty in global-scale soil moisture data, the applicability of these data in global ecosystem models is currently limited (Hashimoto et al., 2015; Stocker et al., 2019).

Reanalysis-based land surface model products are the most frequently used, mainly including the Global Land Data Assimilation System (GLDAS, with 0.25° resolution) (Rodell et al., 2004), European Reanalysis (ERA)-interim (0.75°) (Balsamo et al., 2015) and its successors ERA5 (0.25°) and ERA5-Land (0.1°) (Hoffmann et al., 2019)). Although these products can often predict temporal variations well due to the incorporation of high-quality precipitation data, the bias and root mean square error (RMSE) may be large (Bi et al., 2016; Gu et al., 2019). Moreover, the significant impacts of human activities such as irrigation and land cover changes on soil moisture are rarely considered (Kumar et al., 2015; Qiu et al., 2016). Apart from surface soil moisture that can be observed by satellites, the modeling method also provides information on the moisture in deeper soil layers.

With the advance of remote sensing technology, soil moisture products derived from microwave remote sensing, which have been proven to be superior to those derived from other electromagnetic wave bands (Karthikeyan et al., 2017a), have become an alternative to surface soil moisture monitoring (current satellite microwave sensors can detect only soil moisture within the top 5 cm of soil) (Feng et al., 2017; Jiao et al., 2016; Piles et al., 2018). Currently, global-scale soil moisture can be acquired from either passive (e.g., SMMR, SSM/I, TMI, WindSAT, AMSRE, AMSR2, SMOS, SMAP) or active sensors (e.g., ERS and ASCAT) but the valid temporal spans of all these sensors are limited, and the data quality and spatial coverage were considered to be unsatisfactory untill the launch of AMSRE in June 2002 (Karthikeyan et al., 2017b; Kawanishi et al., 2003). Currently, the ASCAT product is the longest continuous record of global surface soil moisture that is derived only from microwave remote sensing (Bartalis et al., 2007), and the temporal span of this product is from 2007 until present. Apart from the relatively

删除了: s
删除了: It
删除了: , which
删除了: . It is also … as it is probablyconsidered as
删除了: its…he applicability of these data in global ecosystem models are
删除了: and is thus a direct driving factor of evapotranspiration, carbon uptake and many other ecological processes, yet surrogates such as precipitation and vapor pressure deficit (VPD) are often applied in modeling due to…the large uncertainty in global-scale soil moisture data
删除了: The r
删除了: (e.g., …he Global Land Data Assimilation System (GLDAS, with spatial resolution of
删除了: ECMWF …RA)-interim (0.25
删除了: newly-published …uccessors: … ERA5 (0.25°) and
删除了: are the most frequently used
删除了: T
删除了: can hardly be
删除了: way…ethod also provides also the
删除了: prove
删除了: of
删除了: only
删除了: the penetrability of microwaves is usually <5 cm of
删除了: retrieved
删除了: isonly
删除了: is
删除了: the ASCAT product
删除了: whose …nd the temporal span of this product

short time series, satellite-based soil moisture products usually have lower accuracies than modeled products (Albergel et al., 2012; Chen et al., 2013) due to various disturbances, such as high vegetation cover, high open water fractions and complex topography (Draper et al., 2012; Fan et al., 2020; Ye et al., 2015). Moreover, large discrepancies exist among the soil moisture retrieved from various sensors and that by using different algorithms (Kim et al., 2015a; Mladenova et al., 2014). Although new sensors such as SMOS (Stillman and Zeng, 2018) and SMAP (Entekhabi et al., 2010), can produce significantly improved estimates because L-band microwaves (1~2 GHz) can penetrate the vegetation canopy better than the other bands (Burgin et al., 2017; Chen et al., 2018; Karthikeyan et al., 2017b; Kerr et al., 2016; Kim et al., 2018; Leroux et al., 2014; Stillman and Zeng, 2018), the applicability of both products is still limited. SMOS data have too much noise and too many missing values in Eurasia due to high radio frequency interference (RFI) (Oliva et al., 2012). While SMAP has the highest quality (the unbiased RMSE of the passive product can be close to its target of 0.04 m^3/m^3) and has filtered RFI (Chen et al., 2018; Colliander et al., 2017), the data are available since only March 2015.

Because both simulated and satellite-observed soil moisture products have advantages and disadvantages, the interest in data fusion is increasing. The European Space Agency (ESA) published a long-term surface soil moisture dataset called the Climate Change Initiative (CCI) or Essential Climate Variable (ECV), and the v4.5 product covers 1978~2018. Two steps contribute to the combined CCI product. The first step involves rescaling all microwave sensors' retrievals against the reference data (GLDAS Noah product) by cumulative distribution function (CDF) matching, while the second step merges the rescaled products together by selecting the best product in each subperiod or averaging the products weighted by the estimated errors (Dorigo et al., 2017; Gruber et al., 2017; Gruber et al., 2019; Liu et al., 2012). Although the CCI covers more than 40 years, the data before June 2002 have many missing values and are of low quality. Upon rescaling through CDF matching, the spatial patterns of the satellite products are generally replaced by those of GLDAS (Gruber et al., 2019; Liu et al., 2012; Liu et al., 2011b). Although the satellite-observed temporal patterns are retained, the merging algorithm is probably too simple (Liu et al., 2012) to harmonize the discrepancy among the temporal variations in various products (Feng et al., 2017). Another popular data product, Global Land Evaporation Amsterdam Model (GLEAM) soil moisture, is produced by assimilating CCI data (Burgin et al., 2017; Martens et al., 2017; Miralles et al., 2011). Currently, CCI soil moisture anomalies (the deviations to the

删除了: the accuracy of …atellite-based soil moisture products is …sually lower …ave lower accuracies than that of …

删除了: from for example…such as,

删除了: ,

删除了: ,

删除了: have

删除了: better

删除了: penetrability

删除了: Asia

删除了: only …vailable

删除了: there are increasing…he interests…in datae

删除了: integration…usion is increasing. The European Space Agency (ESA) published a long-term surface soil moisture dataset called the Climate Change Initiative (CCI) or Essential Climate Variable (ECV), and the v4.4 … product covers 1978~2018. Two aspects

删除了: combined

删除了: . … while The

删除了: is to merge

删除了: with …f lower

删除了: almost replaced by

删除了: Because the temporal variation in soil moisture is often better captured by model simulations than remote sensing inversions, CCI may undesirably combine the disadvantages of both.

删除了: GLEAM (…lobal Land Evaporation Amsterdam Model (GLEAM))…soil moisture, is produced by the …ssimilatingon…of the …

删除了: anomalies of

seasonal climatology, which indicate whether the soil moisture at a time point is more humid or drier than the multiyear average) are assimilated instead of the original CCI time series (Martens et al., 2017). Therefore, the observed spatial information is ignored, while the temporal changes are mainly driven by model simulations, meaning that remote sensing data are not optimally utilized.

Hence, the probably best practice for global long-term soil moisture mapping is to first develop a long-term surface soil moisture dataset only using the satellite data and then assimilate this dataset into model simulations. The first step is the major task. Rescaling the soil moisture data retrieved from each sensor by using CDF matching followed by averaging the rescaled data during a common period, which is adopted in CCI, will result in problems when unifying the temporal variations in different soil moisture products. Three methods have been proposed that target the use of the information acquired by one sensor to produce soil moisture data compatible with those retrieved from another. Based on physical-based equations (Wigneron et al., 2004), the regression between SMOS soil moisture and the dual-polarized brightness temperature (Tb) data from AMSRE is used to calibrate the AMSRE soil moisture time series (R-square =0.36) (Al-Yaari et al., 2016). An example of the second method uses the Land Parameter Retrieval Model (LPRM) (Owe et al., 2008) to retrieve soil moisture from SMOS and then match these data with the AMSRE-LPRM product by calibrating the LPRM parameters and applying a linear regression (Van der Schalie et al., 2017). Because machine learning can better characterize the nonlinear relationship between surface soil moisture and Tb (Rodriguez-Fernandez et al., 2015), researchers built a neural network that links SMOS soil moisture to the Tb and polarized reflectivity of AMSRE to produce a calibrated soil moisture data covering 9 years (2003~2011) (Rodríguez-Fernández et al., 2016). Among these three approaches, machine learning has been proven to be the best choice according to the connection between precipitation and the changes in soil moisture, as evaluated through a data assimilation technique and triple collocation analysis result (Van der Schalie et al., 2018).

A global long-term observational-based soil moisture product was recently developed by building a neural network between the SMOS product and the Tb data from AMSRE (2003~September 2011) and AMSR2 (July 2012~2015) (Yao et al., 2017). Some environmental factors, including land surface temperature (LST) derived from Tb at 36.5 GHz (Holmes et al., 2009) and the microwave vegetation index (MVI, an indicator of vegetation cover), were also incorporated as ancillary inputs. The

删除了： that …ndicate whether the soil moisture at a time point is more humid or drier than the multi-

删除了： (Lievens et al., 2017)

删除了： as well

删除了: data

删除了: one… common period, which is adopted in CCI, will result in problems whencan hardly…unifying the temporal variations of…n different soil moisture products. Three methods have been proposed that target the use of Targeting at using …he information acquired by one sensor produce soil moisture data compatible with that …hose retrieved from another, three methods have been proposed

删除了： (brightness temperature… data of

删除了: , R^2

删除了: is to use…ses the LPRM (

删除了:)…to retrieve soil moisture from SMOS and then match these datait

删除了: linking …hat links SMOS soil moisture with

删除了: proves …as been proven to be probably …he best choice according to the connection between precipitation and the changes in soil moisture, as evaluated through a data assimilation technique,…

删除了: G…obal long-term observational-based soil moisture product was recently developedhas been developed recently…

删除了: simple

删除了: of

删除了: , with… Some environmental factors, including land surface temperature (LST) derived from TB…b (V)

删除了: amount

training R-square value (R^2) of this product was only 0.45 (or correlation coefficient, r, equals 0.67) and the validation against in-situ measurements showed a temporal r of 0.52 and temporal RMSE of 0.084. The gap between the temporal spans of AMSRE and AMSR2 and the lack of SMOS data in Asia resulted in large quantities of missing data. As SMAP observations have become increasingly available, these data has been chosen as the training target, improving the training R^2 to 0.55, while the overall r and RMSE against measurements were 0.44 and 0.113 (Yao et al., 2019). Another study rebuilt a soil moisture time series over the Tibetan Plateau by using SMAP data as the reference of a random forest (Qu et al., 2019). For the environmental factors, while vegetation cover is not considered, elevation (DEM), IGBP land use cover type, grid location and the day of a year (DOY) were chosen as ancillary inputs. The training R^2 in this region reached 0.9, with a temporal accuracy higher than that of other products (temporal r=0.7; RMSE=0.07 in the unfrozen season). However, these data are regional (for Tibetan Plateau only), and with a temporal gap between AMSR-E and AMSR2 data (October 2011~June 2012). To be concluded, while previous studies have focused on developing long-term satellite-based surface moisture products using machine learning, there remain some major concerns remain that need to be solved. 1) The microwave observations from only three sensors at most are utilized, leading to large temporal and spatial gaps, and the limited training efficiency; 2) it remains unclear which environmental factors should be incorporated as ancillary inputs, and why; and 3) the training designed for soil moisture estimation at the global scale should more complex than that for only a specific region to ensure a satisfactory training efficiency. In this study, 11 high-quality microwave soil moisture products since 2003 are incorporated into 5 rounds of neural networks to achieve a spatially and temporally continuous simulation for 2003~2018, using as many sources of microwave observational data as possible as predictors in each neural network. The quality impact factors of microwave soil moisture retrievals are also determined and then utilized as ancillary inputs to improve the training efficiency. Moreover, we designed localized subnetworks instead of only one global-scale neural network to account for the regional differences in training rules.

删除了: training R-square value (R^2) is only 0.45 (or
设置了格式: 上标
设置了格式: 字体: 倾斜
删除了: s
删除了: *R*
设置了格式: 字体: 倾斜
删除了: (…nd temporal RMSE of =….084)… In addition,
删除了: t
删除了: ,…and the lack of SMOS data in Asia resulted in
删除了: *R*
设置了格式: 字体: 倾斜
删除了: are
删除了:
删除了: , MODIS LST, precipitation
删除了: a
删除了: predicators
删除了: *R*
设置了格式: 字体: 倾斜
删除了: this …hese data is
删除了: …ith a large …emporal gap between AMSR-E and
删除了: recent …revious studies have focused on developing
删除了: substantial success has not been achieved yet.
删除了: the high-quality
删除了: t
删除了:
删除了: three sensors
删除了: not fully …tilized, leading to too …arge temporal
删除了: ought to be…hould more complex than that for only
删除了: with…sing as many sources of microwave
删除了: By taking these matters into account, this study

2 Data and Methods

2.1 Data for the production of global long-term surface soil moisture data

2.1.1 Satellite-based surface soil moisture data products

SMAP currently has the highest quality of all remote sensing-based soil moisture products (Al-Yaari et al., 2019) and is thus chosen as the primary training target. The SMAP Enhanced L3 Radiometer Global Daily 9 km EASE-Grid Soil Moisture V002 (SPL3SMP_E_002, hereinafter SMAP_E for short), which was developed by improving the spatial interpolation of the original 36 km resolution SMAP soil moisture data (Chan et al., 2018), was adopted in this study. The nominal depth of SMAP_E is ~5 cm.

Previous studies often used Tb observations at various bands as network inputs (Rodríguez-Fernández et al., 2016). However, in this study, the well-acknowledged surface soil moisture products retrieved through mature algorithms (see Figure 1) are directly applied instead of Tb. This is because 1) the primary goal of this study is to calibrate and then fuse the existing popular microwave soil moisture products; and 2) the Tb signals at multiple bands contain too much information that is not related to soil moisture, which may weaken the training efficiency and lead to overfitting. Although the drawback is that the final soil moisture products may inherit the uncertainties associated with each retrieval method, this problem can be generally solved by including quality impact factors (see section 2.1.2). The first is the ASCAT soil moisture index product (ASCAT-SWI), which was developed by the ESA-Copernicus Land Monitoring Service (Albergel et al., 2008; Wagner et al., 1999). The saturation degree in the top soil layer (SWI_001) was converted to volumetric soil moisture by multiplication with soil porosity data included in the SMAP L4 Global Surface and Root Zone Soil Moisture Land Model Constants V004 dataset (hereinafter 'SMAP Constant'). Second, AMSR2-JAXA is the AMSR2 soil moisture retrieved by the Japan Aerospace Exploration Agency (JAXA) using Tb at the X-band (10.65 GHz) (Fujii et al., 2009), and version 3 data on the Global Portal System (G-Portal) were used. Third, AMSR2-LPRM-X stands for the AMSR2 soil moisture produced by applying the LPRM algorithm at the X-band (Parinussa et al., 2014) (C-band data such as AMSR2-LPRM-C or AMSRE-LPRM-C were not applied due to high RFI (Njoku et al., 2005)), and is obtained from NASA's Earthdata Search web. The fourth predictor, SMOS-IC, is a new SMOS soil moisture product created by INRA and CESBIO with the main goal of being as independent as possible from the

- 删除了: and methods
- 删除了: The s
- 删除了: is
- 删除了: currently
- 删除了: best
- 删除了: (Al-Yaari et al., 2019; Albergel et al., 2012)
- 删除了: …nd is thus chosen as the primary training target. The SMAP Enhanced L3 Radiometer Global Daily 9 km EASE-Grid Soil Moisture V002 (SPL3SMP_E_002, which is referred to as
- 删除了: i… developed by an …mproving theed
- 删除了: :
- 删除了: , while…the Tb observations …ignals at multiple
- 删除了: -
- 删除了: , the well-acknowledged surface soil moisture
- 删除了: T
- 删除了: 2 for details.…
- 删除了: ESA
- 删除了:
- 删除了: of
- 删除了: the
- 删除了: ly
- 删除了: ly
- 删除了: ,…was
- 删除了: incorporated
- 删除了: 4th one
- 设置了格式: 上标
- 删除了: made

auxiliary data, including the simulated soil moisture (Fernandez-Moran et al., 2017a; Fernandez-Moran et al., 2017b; Wigneron et al., 2007). The accuracy of SMOS-IC has been proven to be higher than that of other SMOS products (Al-Yaari et al., 2019; Ma et al., 2019), and the data version 105 offered by Centre Aval de Traitement des Données SMOS (CATDS) is adopted. TMI-LPRM-X is the X-band LPRM product of TMI and was created by the NASA Goddard Space Flight Center (GSFC), which is used as the 5th predictor. Fengyun 3B is a Chinese meteorological satellite with a Microwave Radiation Imager (MWRI) onboard (Yang et al., 2011; Yang et al., 2012). The National Satellite Meteorological Center product is retrieved using Tb at 10.7 GHz, which is denoted by 'FY-3B-NSMC' (the 6th predictor product). WindSat is flown on the Coriolis satellite (Gaiser et al., 2004), and the soil moisture retrieved by LPRM at the X-band (Parinussa et al., 2012) is provided by NASA (the 7th predictor). Three AMSRE products are also used, including the NASA product (AE_Land3) created by the National Snow and Ice Data Center (AMSRE-NSIDC)(Njoku et al., 2003), the JAXA product (AMSRE-JAXA) (Fujii et al., 2009; Koike et al., 2004) obtained from G-Portal and the LPRM product (AMSRE-LPRM) available at the NASA Earthdata Search. All data are reprojected to the WGS-1984 reference coordinate system and resampled to 0.1°.

To reduce noise and fill the gaps between sensor observation tracks (it takes at least 3 days for a microwave sensor to cover the whole globe), for every soil moisture product, both the daytime and nighttime observations within each 10-day period are combined by data averaging (the relative superiority of daytime and nighttime retrievals is not considered). For example, for SMAP, 11% of the global land surface has data for only 5 days or less within a 10-day period.

2.1.2 The quality impact factors of soil moisture retrievals

Environmental factors, including DEM, LST and vegetation cover (indicated by NDVI, MVI, etc.), were used as ancillary neural network inputs to improve the soil moisture simulation (Lu et al., 2015; Qu et al., 2019; Yao et al., 2017). According to these studies, these factors alone may not predict surface soil moisture well without the incorporation of any microwave remote sensing data, which can also be justified by the contribution analysis results (Figure S1a). This is because although they are somewhat related to soil moisture (e.g. soil moisture is generally limited in areas with low vegetation cover but high in forests (McColl et al., 2017)), the relationships are rather uncertain (e.g., at smaller scales, leaf area index (LAI) may have

删除了: proves to be higher than othe

设置了格式: 上标

删除了:

设置了格式: 上标

设置了格式: 上标

删除了: as well

删除了: -

删除了: s

删除了: ing

设置了格式: 字体: 加粗

删除了: precipitation,

删除了: for

删除了: d

删除了: estimation and data fusion

删除了: in general

删除了: however

a negative influence on soil moisture due to the variation in evapotranspiration (Naithani et al., 2013), or without clear impacts (Zhao et al., 2010); also, soil moisture can be either high or low in summers when vegetation peaks (Baldocchi et al., 2006; Méndez-Barroso et al., 2009)). However, these factors are quite essential due to their direct impacts on soil moisture retrieval through the radiative transfer model using microwave remote sensing data (Fan et al., 2020), and are retrieval quality impact factors. The detailed explanations are as follows: 1) the bias of soil moisture estimates derived from a certain sensor or a specific algorithm can be correlated with the degree of disturbances from various environmental factors. For example, in vegetated areas, LST is overestimated by LPRM (Ma et al., 2019), whereas soil moisture is underestimated by JAXA (Kim et al., 2015a), and the magnitudes of the biases are determined by vegetation amount or VOD. It indicates that the environmental factors are essential for a better calibration of various products, especially when soil moisture, which contains errors associated with the retrieval method, is directly applied instead of Tb; and 2) the relative performances of different products are also controlled by environmental factors; for example, the ASCAT product is preferable to AMSRE-LPRM in vegetated areas (Dorigo et al., 2010), while LST influences the relative superiority of the LPRM and JAXA algorithms (Kim et al., 2015a). Therefore, for improved data fusion, the weights assigned to different predictor soil moisture (or Tb) predictor data available at the same time should be determined by referring to these quality impact factors (Kim et al., 2015b).

In this study, 9 quality impact factors: LAI, water fraction, LST, land use cover, tree cover fraction, non-tree vegetation fraction, topographic complexity, and sand and clay fractions are selected and incorporated (see Figure 1). The reasons are as follows. Based on the two criteria above, the first environmental factor to be included is the ‘vegetation factor’ (i.e., vegetation water content, VWC). Plants can absorb or scatter radiation from soil and emit radiation, reducing the sensitivities of both radiometer and radar to soil moisture (Du et al., 2000; Owe et al., 2001). However, L-band microwaves can penetrate the vegetation layer better due to their longer wavelengths (Konings et al., 2017; Piles et al., 2018). On the other hand, although vegetation effects can be somewhat corrected (Jackson and Schmugge, 1991), different methods have different efficiencies. First-order radiative transfer models such as LPRM have difficulty describing the radiation attenuation by dense canopy (Crow et al., 2010), but the TU-Wien change detection algorithm applied to ASCAT can reduce vegetation impacts due to the implicit account of high-order scattering effects (Bartalis et al., 2007). Microwave vegetation indexes may contain large uncertainty and have coarse

删除了: However, the mechanism why these factors are helpful remains controversial: are these factors used as direct spatial predictors of soil moisture or just because they are related to the errors of satellite soil moisture retrievals (i.e., the quality impact factors of soil moisture)? We insist on the latter,…

删除了: proposing two main reasons for the incorporation of environmental factors

删除了: are

删除了: we determinedwhich environmental factors should be included.

删除了: T

删除了: the

删除了: its

删除了: can hardly describe

删除了: er

删除了: are in

resolutions (Liu et al., 2011a; Shi et al., 2008). NDVI becomes saturated at high vegetation cover (Huete et al., 2002). Because the LAI stands for the total leaf area per unit land, which is closely related to VWC assuming a relatively stable leaf equivalent water thickness (Yilmaz et al., 2008), LAI is a suitable surrogate. The Copernicus global 1 km resolution LAI (GEOV2-LAI) data are adopted here due to high accuracy and full coverage (Baret et al., 2013; Camacho et al., 2013; Verger et al., 2014). Because the sensor conversion from SPOT-VGT to PROBA-V in 2014 led to LAI data discontinuity in specific areas (Cammalleri et al., 2019), which may reduce neural network training and simulation efficiency, the GLASS (Global LAnd Surface Satellite) LAI product (Xiao et al., 2014; Xiao et al., 2016) from 2007~2017 is also used (Figure 1). The LAIs are averaged on a monthly scale and aggregated to 0.1° resolution. The second is the 'water fraction factor' (i.e., the fraction of water area in each pixel). Waters in land pixels dramatically decrease the Tb, leading to overestimation of soil moisture there. Because different methods are used to detect and correct small areas of water, either open water, wetlands or partly inundated wetlands and croplands (Entekhabi et al., 2010; Kerr et al., 2001; Mladenova et al., 2014; Njoku et al., 2003), microwave soil moisture data calibration and weight assignment based on the water fraction within land pixels make sense (Ye et al., 2015). In addition, the water fraction is a direct indicator of surface soil moisture. In this study, daily water area fraction derived from the Surface WAter Microwave Product Series (SWAMPS) v3.2 dataset (Schroeder et al., 2015) is applied. The third factor is the 'heat factor' (i.e., LST). Soil moisture retrievals from passive microwave sensors are based on the correlation between the soil dielectric constant, which is influenced by soil moisture, and the emissivity estimated as the ratio of Tb to soil physical temperature (Ts) (Karthikeyan et al., 2017a). Ts is approximate to LST and can be derived from the Tb at 36.5 GHz (Holmes et al., 2009; Parinussa et al., 2011), or from reanalysis datasets including ECMWF, MERRA and NCEP, or set as a constant of 293 K (Koike, 2013). Active microwave products are independent of LST (Ulaby et al., 1978). Because different LST estimates are used in the retrievals of different soil moisture products, while the bias of each LST estimate compared to the actual LST is influenced by the actual LST, we assume that that the actual LST can determine the accuracy of every LST estimate and finally the relative performances of various soil moisture products (Kim et al., 2015a). In this study, we averaged the MODIS monthly LST acquired from the ascending and descending passes of both TERRA and AQUA. The 4th factor is the 'land cover factor', which is added because the parameters essential for soil moisture retrieval (vegetation effect correction)

删除了: leaf area index (

删除了:)

删除了: an

删除了: water body

删除了: bodie

删除了: lower

删除了: there are

删除了: used of for detection and correction of watersmall area of water,

删除了: bodies

删除了: body

删除了: wa

删除了: is

删除了: conductivity

删除了: that

删除了: influenced

删除了: Because Because in the retrievals of different soil moisture products, different LST estimates are used, while the bias of each LST estimate have different biases compared to the actual valuesLST is influenced by actual LST,, we suppose

删除了:

删除了: may

删除了: bias

删除了: each

删除了:

删除了: product

删除了: ,

设置了格式: 上标

are set based on land use types (Griend and Wigneron, 2004; Jackson and Schmugge, 1991; Jackson et al., 1982; Panciera et al., 2009). Additionally, landscape heterogeneity influences the retrieval accuracy (Lakhankar et al., 2009; Lei et al., 2018; Ma et al., 2019). Here, both the annual MODIS land use cover maps and the MEaSUREs vegetation continuous fields (i.e. the cover fractions of trees, non-tree vegetation and bare ground (Hansen and Song, 2018)) are adopted. Apart from the above dynamic factors, there are also two static factors: the 'topographic factor' (i.e., topographic complexity or surface roughness) and the 'soil factor' (i.e., soil texture indicated by sand and clay fractions) (Neill et al., 2011). Both factors can influence the relationship between soil moisture and emissivity or the dielectric constant (Dobson et al., 1985; Karthikeyan et al., 2017a; Njoku and Chan, 2006), but are characterized and corrected differently, leading to different relative performances of various soil moisture products (Das and O'neill, 2010; Gao et al., 2006; Kim et al., 2015a). For topographic complexity, the static layer of the Copernicus ASCAT-SWI product (hereinafter the ASCAT Constant) is adopted while for soil texture, the SMAP Constant is used (topographic complexity data are not available from SMAP Constant; soil texture is not provided by ASCAT Constant). The contribution analysis results (Figure S1) show that because various microwave soil moisture data have already been included, precipitation data are not an essential indicator of soil moisture, and are not utilized as a physically based 'quality impact factor' either (see Text S1 for detailed explanations).

2.2 Methods for the production of global long-term surface soil moisture data

The global long-term surface soil moisture data production includes three basic parts, which are as follows. 1) Preprocessing: the production of high-quality neural network inputs; 2) neural network operation: the network training and soil moisture simulation; and 3) postprocessing: the correction of potential errors or deficiencies in the soil moisture simulation outputs. Because the temporal coverage of SMAP does not overlap with that of TMI, FY-3B, WindSat or AMSRE, several rounds of simulations are performed to fully utilize the satellite-based soil moisture data. Hence, the simulated soil moisture may also be converted to the training target of the next round's neural network, meaning that some postprocessing steps are also preprocessing steps. The basic flow of this process is shown in Figure 2.

删除了:

删除了: is

删除了:

删除了: P

删除了: and DOY are

删除了: is

删除了: included as

删除了: ancillary input

删除了: s

删除了: is

删除了: -

删除了:

删除了: Therefore, there are 9 quality impact factors: LAI, water body fraction, LST, land use cover, tree cover fraction, non-tree vegetation fraction, topographic complexity, and sand and clay fractions, which are essential inputs of all neural networks in this study.

删除了: **2.3 The basic flow of soil moisture calculation**

删除了: It should be noted that b

删除了: to make full use of the satellite-based soil moisture data, …

删除了: will

删除了: part of

2.2.1 Neural network design (1): localized neural networks

In this study, instead of a universal network, we devised localized neural networks. The data within each individual zone are used to train a zonal neural network (hereinafter a subnetwork), which is used for soil moisture simulation at that zone. By comparison, localized neural networks help improve the training efficiency; however, a smaller zonal size does not indicate a better simulation accuracy. We noticed that the LPRM algorithm-based products (AMSR2/TMI/WindSat/AMSRE-LPRM-X) were patchy, with clear boundaries between adjacent square-shaped zones over arid regions, while the patch size was exactly 1°×1°, which was probably due to the spatial distribution of parameters. This finding suggests that subnetworks should be built at the 1°×1° scale. Therefore, we divided the global extent except the polar areas (80°N~60°S) into 140×360 zones. Here, for a 0.1° pixel during a specific 10-day period, if all the input data (soil moisture products and quality impact factors) have valid values, one valid data point is provided. Therefore, the maximal number of valid data points applied to train a subnetwork = 100 × the number of 10-day periods within the training period. The subnetworks with less than 100 valid data points (e.g., those in oceans) were dropped, leaving usually >15,000 zonal subnetworks included in an independent neural network. The training was performed in MATLAB 2016a, and the number of nodes in the hidden layer (between the input and output layers (Stinchcombe and White, 1989)) of each subnetwork was set to 7.

2.2.2 Preprocessing and postprocessing steps

After standardization of the original soil moisture data, to improve the neural network training efficiency, the potential salt and pepper noises are removed. For each map (a specific 10-day period), within each 1°×1° zone, the soil moisture values are filtered to the level of three standard deviations relative to the mean in that zone. This preprocessing step is thus called '3σ denoising' (note that the denoise is conducted spatially, rather than temporally, so that the extreme events will not be treated). After neural network operation, boundary fuzzification is first applied, that is a step in both preprocessing and postprocessing. Because the localized 1°×1° network is applied instead of the global network, the boundary between nearby zones may be too obvious over some areas. To blur the boundary, a simple algorithm is applied, as shown in Figure S1. The soil moisture data with fuzzified boundaries are transformed into both the final product and the next round's training target. To produce the final

删除了: yet,…a smaller zonal size does not indicate a better simulation accuracy. We noticed that the LPRM algorithm-based products (AMSR2/TMI/WindSat/AMSRE-LPRM-X) are …ere patchy, with clear boundaries between adjacent square-shaped zones over arid places…egions, while the patch size wai… exactly 1°×1°, which wai

删除了: indicates

删除了: it can provide …ne valid data point is provided. Se

删除了: For each…e subnetworks with, if the number of valid pixels (for a 0.1° pixel in a given 10-day period, if all the subnetwork inputs have valid values, it is seen as a valid pixel; the maximal number of valid pixels=100×the number of 10-day periods)

删除了: valid data points

删除了: exceeds

删除了: , the subnetwork is…ere considered valid…ropped. There are… leaving usually more than …15,000 zonal valid …ubnetworks within…included in an individual

删除了: i… performed in the

删除了: software

删除了: i

删除了: each

删除了: e

删除了: which …hat is a step in

删除了: represented as Figure S1 is applied… The soil moisture data with fuzzified boundaries are transformed into both The boundary fuzzified soil moisture data are both transformed into product, two postprocessing steps are essential: filling of missing values and data masking. Because '3σ denoising' deleted suspicious soil moisture retrievals, the simulation outputs also contain few missing values, which can be simply filled by sequentially searching and averaging nearby valid values (Chen et al., 2019). While the snow/ice mask of the ASCAT-SWI product can be transferred to the simulation output, the potential snow or ice cover before 2007 should be identified. For a pixel in a specific ten-day period, if ice cover is reported by ASCAT-SWI in most years, it is also supposed to be covered by snow/ice, unless the thaw state is observed in the MEaSUREs Global Record of Daily Landscape Freeze/Thaw Status V4 dataset. The simulated soil moisture in the rainforests identified in the 'ASCAT Constant' is retained but not recommended due to the high uncertainty. On the other hand, to avoid error propagation with training times by ensuring a high-quality training target for the next round's simulation, we remove all suspicious values for every simulated result. This preprocessing step is performed by first obtaining the maximal and minimum values of SMAP_E soil moisture in each pixel. If the simulated value is out of the range of the SMAP data during 2015~2018, the value is considered suspicious and is not used as a training target. Subsequently, '3σ denoising' is performed again before the simulated soil moisture becomes secondary training target, which are referred to as SIM-1T, SIM-2T, and so on ('SIM' stands for the simulated soil moisture, the number after the hyphen indicates the round of simulation, and 'T' means it is applied as training target; the temporal spans of SIM-XT and SIM-X are the same, as shown in Figure 1).

2.2.3 Neural network design (2)- five rounds of simulations

The 11 available microwave soil moisture data products with different temporal spans are incorporated, and they are also utilized as fully as possible through up to 5 rounds of neural network-based simulations, with at least four different soil moisture products retrieved from three sensors applied as predictors in each round (see details below). While increasing the sources of soil moisture data inputs can be beneficial to the training efficiency, the spatial coverage of the simulation output is sacrificed because the overlapping area decreases with the increase in the number of soil moisture products. After all, most products have missing data in specific regions (e.g., mountains, wetlands and urban settlements), and some sensors are even unable to produce data at the global scale (TMI is limited to [N40°, S40°]; SMOS lacks data in Asia). To solve that dilemma,

删除了: e

删除了: on

删除了: t… the simulation output, the potential snow or ice cover before 2007 should be identified. For a grid …ixel in a specific ten-day period, if ice cover is reported by ASCAT-SWI in most years, it is also supposed to be potentially

删除了: if

删除了: . However, if

删除了:

删除了: the …haw state is observed according to…n the MEaSUREs Global Record of Daily Landscape Freeze/Thaw Status V4 dataset., the grid is not masked.

删除了: are …s retained but not recommended due to large

删除了: for every simulated result

删除了: performed a preprocessing step called suspicious value removal

删除了: d…all the

删除了: We …btaininged

删除了: thus …s not included in

删除了: the … training target. Subsequently, '3σ denoise' …enoising' is performed again before the simulated soil moisture becomes secondary training targets

删除了: A key characteristic of this study is that n

删除了: Not only the 11 available microwave soil moisture data with different temporal spans are all incorporated, but they are also utilized as fully as possible through up to 5 rounds of neural network-based simulations

删除了: output

删除了: of more soil moisture products is smaller

删除了: decreases with the increase in soil moisture data we classified all 0.1° pixels according to the available predictor soil moisture products over a 10-day period (for example, if there are at most four soil moisture data inputs in one round, there should be 4+3+2+1=10 combinations). However, to avoid soil moisture simulation under snow or ice cover (Section 2.2.2), not all combinations are considered. Then, corresponding to each selected combination, an independent neural network is trained. For data simulation in a 0.1° pixel, the most preferable independent neural network is expected to be trained using all the available soil moisture data sources in that pixel. However, in the 1° zone where it is located, the subnetwork belonging to that preferable independent neural network may not exist due to limited valid data points (see section 2.2.1). Then, an alternative subnetwork driven by the combination of fewer soil moisture data inputs should be applied instead. Hence, we should determine the neural network collocation that is the best choice for every pixel. Apart from applicability, the relative priority order of different neural networks was obtained by comprehensively considering the number and quality of input soil moisture products, the variety of sensors, the quantity of training samples indicated by the number of 10-day periods, and the relative accuracy of the training targets (the training target quality declines monotonically: SMAP>SIM-1T>SIM-2T>SIM-3T>SIM-4T). Sometimes, the two most likely priority orders are given, with the simulation results of the corresponding two substeps integrated later. Specifically, when the LAI data source changes, the division of a single round into several substeps is also essential. Based on these principles, five rounds of neural networks are designed as follows, with 8 substeps containing a total of 67 independent neural networks. The training period for each neural network and the simulation period for each substep are shown in Figure 1 (below the timeline), and the details are as follows:

For the first round's neural network (labeled as NN1), the potential training period is 2015D10~2018 ('D' is the ordinal of the 10-day period, so '2015D10' represents the period from April 1st to April 10th in 2015) because SMAP soil moisture data during that period are applied as the training target, while ASCAT-SWI10 (abbreviated as ASCAT), SMOS-IC (SMOS), AMSR2-JAXA and AMSR2-LPRM-X (AMSR2-LPRM) are the four soil moisture products used as predictors (details are in Tables S1~S2). Because all the four predictors have data since 2012D19, the potential soil moisture simulation period is 2012D19~2018, which is further divided into two parts: one is 2014~2018 (substep1), for which the PROBA-V LAI data that begins in 2014 are applied, whereas the other is 2012D19~2013 (substep2), for which GLASS LAI data are used (note: because

删除了: available …oil moisture products with valid data

删除了: in it

删除了: during

删除了: predictor

删除了: considered …onsidered(see Section 2.5) while some combinations are optional… Then, corresponding to selected combinations… different …n independent neural networks are…is trained. Moreover,

删除了: f…r data a …imulation in a 0.1° pixel, …the most preferable independent neural network is expected to be trained using all the available soil moisture data sources in that pixel. However,if

删除了: …the neural network driven by multiple predictor soil moisture has no subnetwork in the zone where it is located due to limited…imited valid pixels…ata points (see

删除了: which

删除了: a …very specific

删除了: …of different neural networks should be

删除了: ere

删除了: determined

删除了: monotonously…onotonically: SMAP>SIM-

删除了: he priority order of independent networks cannot b

删除了: operations… are designed as follows, with 8

删除了: of…each neural network and the simulation period

删除了: l…ed as NN1), the potential training period is

设置了格式

删除了: April 1st to 10th of 2015

设置了格式

删除了: is…re applied as the training target, while ASCAT

删除了: was…s further divided into two parts: one is

GLASS LAI covers the period from the beginning of our study period until 2017, the training period for substep 2 is 2015D10~2017). Please refer to Tables S1~S2 for details. The simulation results of the two substeps (SIM-1-1 and SIM-1-2) are combined as SIM-1 and then transformed into a secondary training target, denoted as SIM-1T. In the second round of simulation, the training target can be either SMAP or SIM-1T, while the soil moisture input data are ASCAT, SMOS, TMI-LPRM-X (TMI) and FY-3B-NSMC (FY). The simulation output, SIM-2, covers the period of 2011D20~2012D18, which is constrained by the common period of the four predictors (Table S3~S4). In the third round of neural network operation, the simulation period is 2010D16~2011D19. SMAP, SIM-1T and SIM-2T are combined and used as the training targets (the training periods are within the range of 2011D20~2017D36), while the predictor soil moisture data are ASCAT, SMOS, TMI and WindSat-LPRM-X (WINDSAT). There are two substeps in round 3 that are distinguished by whether the priority order of the neural networks is determined mainly based on the training sample quantity and the training target quality (SIM-3-1), or by first considering the number of predictor soil moisture products (SIM-3-2, Table S5~S8). Because these two methods emphasize different aspects of neural network quality, in some pixels, SIM-3-1 will be advantageous, but in others, SIM-3-2 could be better. Hence, an algorithm is devised to combine the advantages of both simulations (SIM-3), which is described in Table S9. Next, the 4th round is for simulations during 2007D01~2010D15. SIM-2T and SIM-3T are combined to be the training target and ASCAT, WINDSAT, TMI, AMSRE-JAXA, AMSRE-LPRM-X (AMSRE-LPRM) and AMSRE-NSIDC are all applied as predictors (LAI data now come from SPOT-VGT). Two substeps are also needed. In the first substep, neural networks are sorted by paying the greatest attention to the number of soil moisture inputs and the sensors they are derived from, while the training sample size and training target quality are prioritized to create an alternative estimate (Tables S10~S13). Afterwards, SIM-4 is obtained by reasonably integrating these two results. In the final round, the soil moisture simulation is extended to as early as 2003. SIM-2T, SIM-3T and SIM-4T together are the training targets, while the predictor soil moisture data entering the neural networks consist of WINDSAT, TMI, AMSRE-JAXA, AMSRE-LPRM and AMSRE-NSIDC (Table S14~S15).

删除了: l

删除了: The first round of simulation has two substeps. In substep 1, GEOV2 PROBA-V LAI data are used for soil moisture simulation during 2014~2018 (note: the other 8 quality impact factors are applied as well) whereas in substep 2, the GLASS LAI are applied, and the simulation period is 2012D19~2013D36 (i.e., September 2012 to 2013; D is the ordinal of the ten-day period). SMAP soil moisture is the training target, while ASCAT-SWI10 (abbreviated as ASCAT), SMOS-IC (SMOS), AMSR2-JAXA and AMSR2-LPRM-X (AMSR2-LPRM) are the four predictor soil moisture products

删除了: (details are in… Tables S1~S2 for details)

删除了: ,

删除了: the… secondary training target (

删除了: by

删除了:)

删除了: be …MAP or SIM-1T, while the input

删除了: spans from …overs the period of 2011D20 to …2?

删除了: that

删除了: (…he simulation period is

删除了: to be…nd used as the training targets (the training

删除了: …eural networks is determined mainly by …ased

删除了: With

删除了: being …ombined to be the training targets

删除了: s

删除了: included

删除了: more …he greatest attention to the number of soil

删除了: estimation …Tables S10~S13). Afterwards, SIM-4

设置了格式: 字体: 10 磅

2.3 Methods for the validation of surface soil moisture products

For the evaluation of global soil moisture data, the International Soil Moisture Network (ISMN) dataset (Dorigo et al., 2011; Dorigo et al., 2013) is the most frequently used (Al-Yaari et al., 2019; Albergel et al., 2012; Dorigo et al., 2015; Fernandez-Moran et al., 2017b; Gao et al., 2020; Karthikeyan et al., 2017b; Kerr et al., 2016; Kim et al., 2015b; Kolassa et al., 2018; Lievens et al., 2017; Zhang et al., 2019). Because SMAP, the training target, is the soil moisture within 0~5 cm, the simulated soil moisture is used for that soil layer as well. Accordingly, the measurements used for validation are limited to ≤ 5 cm in depth (e.g., the Russian networks were not applicable for this reason). The quality flags of ISMN (Dorigo et al., 2013) are also checked to retain only the ‘good quality’ data. After data screening and processing (for example, in the case of high spatial variability in soil moisture, we excluded the pixels with average annual maximal water area fractions greater than 5%, see Text S2), more than 100,000 10-day averaged soil moisture records acquired from 728 stations of 29 networks are selected for validation of the soil moisture product, which as a temporal resolution of 10 days. More than 90% of the stations are located in relatively flat areas with topographic complexity less than 10%. The detailed information of these stations and the periods of the data used are listed in Table S16, while the spatial distribution of these stations is shown in Figure 3. The major climate types of the sites are determined from the Köppen-Geiger climate classification map (see Table 1 for the description (Kottek et al., 2006)).

Most ISMN networks are dense networks, as the stations are very close to each other, often within the same 0.1° pixel, whereas others are sparse networks (see Text S2 and Figure 3). In addition, various sensors are simultaneously operated at some stations. Hence, to make full use of all the high-quality records, and to reduce the problem caused by the scale difference between simulation and measurement, the site-scale 10-day averaged soil moisture data are further aggregated to a 0.1° pixel-scale by averaging all the data (different stations or different sensors) within the pixel (Gruber et al., 2020). Specifically, if soil moisture is not simulated due to snow or ice cover, the corresponding measurement is not useful. This process resulted in a final collection of ~40,000 pixel-scale 10-day period soil moisture records within the validation dataset.

The soil moisture datasets to be evaluated include the RSSSM product in this study (Remote Sensing Surface Soil Moisture, covering 2003~2018), SMAP_E (the primary training target, covering April 2015~2018), the longest record of satellite-based

删除了: 2.7

删除了: V

删除了: (Al-Yaari et al., 2019; Karthikeyan et al., 2017b)

删除了: supposed

删除了: of

删除了: whose… which as a …temporal resolution ofis

删除了: ower than

删除了: S…2… The major climate types of the sites are determined from the updated map of the …öppen-Geiger climate classification map (see Table 2 … for the descriptions…of each type …

删除了: probably …ften within the same 0.1° grid

删除了: some

删除了: S…2…. In addition, various sensors aa

删除了: available

删除了: good

删除了: grid…ixel-scale by averaging the same period’s values within the same grid

删除了: ….Specifically, if soil moisture is not simulated due to snow or ice cover, the corresponding measurement is excluded…ot useful. This process resulted in a final collection of ~40,000 grid

删除了: simulated (satellite-based) soil moisture…SSSM product in this study (Remote Sensing Surface Soil Moisturehereinafter ‘SIM’… covering 2003~2018), (the primary training target, covering April 2015~2018), the longest record of existing …atellite-derived soil moisture: ASCAT-SWI (converted to volumetric fraction; data period is 2007~2018), the reanalysis-based soil moisture: GLDAS Noah V2.1 and ERA5-Land (data were resampled, 10-day averaged and then evaluated during 2003~2018) as well as the soil moisture datasets developed by combining both satellite observations and model simulations: CCI v4.5 and GLEAM v3.3 (for v3.3a, the radiation and air temperature forcing data come from ERA5, whereas for v3.3b, all meteorological data are satellite-based, yet the data after September 2018 are not available). The overall performance of any soil moisture product is first evaluated using all of the validation datasets, with Pearson R-square (R^2) and RMSE values (unit: m^3 m^{-3}) adopted as the main indicators. The next step is temporal pattern validation. For pixels with enough (>20) 10-day averaged in situ records, we compare the estimated soil moisture during all periods against the corresponding measurements, with the calculated Pearson correlation coefficient (*r*) and RMSE. Several supplementary indexes are also added, including bias, unbiased RMSE (ubRMSE) and the correlation coefficient between the anomalies (anomalies *r*, abbreviated here as 'A.R'; A.R can better indicate the simulation accuracy of interannual variations; soil moisture anomalies are calculated by Eq. 1). Next, we compare the means and medians of the above evaluation indexes for different soil moisture products and test whether the differences are significant. Moreover, the relative performances of various products in different climatic zones are analyzed. Finally, we perform spatial pattern validation. In detail, for every 10-day period, we compare all the soil moisture measurements that are upscaled to 0.1° during that period with the corresponding estimated values. The spatial pattern evaluation indexes include the correlation coefficient (*r*), RMSE, bias and ubRMSE values (Eq. 2). The relative superiority of all products during different 10-day periods in a year, and the changes in data coverage as well as data quality with time are also investigated.

$$\overline{SSM(k)} = \frac{\sum_{y=1}^{ny} SSM(y,k)}{ny} \quad (ny \geq 3)\,;\ SSM\ is\ either\ estimated\ or\ measured$$

SSM: *surface soil moisture*; k: *the ordinal of* 10 *day period in a year*; y: *a year with measured SSM in* k^{th} 10 *day period*; ny: *number of those years*

$$SSM_{anom}(y,k) = SSM(y,k) - \overline{SSM(t)}$$

$SSM_{anom}(y,t)$: *the anomalies of surface soil moisture during the* t^{th} 10 *day period in year* y (Eq. 1)

$$\overline{SSM_{est}} = \frac{\sum_{i=1}^{ng} SSM_{est,i}}{ng}\ ;\quad \overline{SSM_{act}} = \frac{\sum_{i=1}^{ng} SSM_{act,i}}{ng} \quad (ng \geq 20)$$

i: *a grid with upscaled surface soil moisture measurements during a specific* 10 *day period*; ng: *the number of those grids in the globe*

删除了: record… ASCAT-SWI (converted to volumetric
删除了: s
删除了: time series
删除了: is
删除了: is…only updated to…ot available September
删除了: . The next step is the temporal pattern validation.
删除了: *R*
设置了格式: 字体: 倾斜
删除了: calculated
删除了: *R*
设置了格式: 字体: 倾斜
删除了: , … A.R can better indicate the
删除了: were
删除了: Then
删除了: d
删除了: /
删除了: ed
删除了: climate …limatic zones are
删除了: were
删除了: Last
删除了: ed
删除了: F
删除了: d…all the soil moisture measurements that were
删除了: soil moisture…alues data in all grids
删除了: with stations are evaluated by t…he spatial pattern
设置了格式: 字体: 倾斜
删除了: ,…RMSE, bias and ubRMSE values (Eq. 2). The
删除了: of…data coverage as well as data quality with time

$$ubRMSE_{spatial} = \sqrt{\sum_{i=1}^{ng}[(SSM_{est,i} - \overline{SSM}_{est}) - (SSM_{est,i} - \overline{SSM}_{act})]^2/ng} \quad \text{(Eq. 2)}$$

2.4 Methods for the intra-annual variation analysis of surface soil moisture

Because the original resolution of SMAP soil moisture is ~0.4° while that of most predictor soil moisture products is 0.25°, the intra-annual variation analysis of RSSSM is performed at 0.5° resolution. We also exclude high-latitude areas (60°N~90°N) where the available data are limited due to frequent ice cover. Fourier functions can characterize intra-annual variation well (Brooks et al., 2012; Hermance et al., 2007). Therefore, for the remaining areas (60°S~60°N), based on a total of 36×16 (years) =576 data points, we fit the intra-annual cycle of soil moisture using the Fourier function with the period fixed to 1 year (36 10-day periods). The number of terms is set to 1 unless the intra-annual cycle is obviously asymmetrical and can be much better characterized by a two-term Fourier function. Subsequently, the highest peak and lowest trough values of surface soil moisture as well as the corresponding locations in time (the ordinal of 10 days) are exported.

The direct driving factor of the variation in surface soil moisture is precipitation, for which we adopted the GPM IMERG Precipitation V06- Final Run data (Huffman et al., 2019). Apart from a direct correlation analysis, we also explore the relationship between the intra-annual cycles of precipitation and surface moisture using Fourier fitting (the derived fitting function is dropped if the adjusted R^2 is lower than 0.1), with the peak time difference in each 0.5° grid calculated (if both cycles have two peaks, the average locations of the two peaks are calculated). Because RSSSM indicates the average soil moisture condition during every 10-day period, we evaluate the surface soil moisture decline after 20 consecutive days (i.e., two adjacent 10-day periods) without effective precipitation to explore the impact of dry periods on surface soil moisture. Effective precipitation is calculated by precipitation minus canopy interception, which is estimated by the modified Merriam canopy interception model (Kozak et al., 2010; Merriam, 1960). If the total effective precipitation within two consecutive 10-day periods (20 days) is less than a given threshold (initially set to 10 mm), we consider that the soil moisture change in the latter period compared to the previous period is mostly due to surface evaporation and percolation (capillary rise is negligible (McColl et al., 2017)), and thus should be negative. Hence, for a 0.5° grid, if the number of negative values does not meet two times the number of positive values, the precipitation threshold is reduced by 1 mm until that condition is satisfied. This loop

删除了: 2.8

删除了: I

删除了: It should be noted that although the nominal spatial resolution of the surface soil moisture product (SIM) is 0.1°,

删除了: the intra-annual variation analysis would be more robust if we aggregate SIM toof RSSSM was performed at 0.5° resolution. because the original resolution of SMAP soil moisture is ~0.4° while the resolution of most soil moisture products to be calibrated and fused is 0.25°. Then, Wwe also excluded the high latitude areas…(60°N~90°N)

删除了: the

删除了: F

删除了: ted

删除了: by …sing the Fourier function (probably the best choice for periodic function fitting(Brooks et al., 2012; Hermance et al., 2007)

删除了:)

删除了: f…nal Rr

删除了: d

删除了: used

删除了: the

删除了: the …ourier function to characterize the intra-annual

删除了: the simulated surface soil moisture…SSSM only

删除了: studied

删除了: ten …0 consecutive rainless

删除了: that

删除了: combining

删除了: (the vegetation canopy cover and LAI is acquired

删除了: if each time when the number of negative values is terminated when there are less than 36 available data points in dry periods (the maximal number of data points is 576), and then the grid is excluded from the analysis. In desert areas, the random noise of the surface soil moisture product can hide the signal of moisture changes, while in wet areas (e.g., rainforests), 20 days without effective precipitation seldom occurs, leading to no results over most areas. In the remaining areas, the intra-annual variation in the surface moisture loss during dry days can be fitted by the Fourier function as well, which is then analyzed using the above methods.

3 Results

3.1 The neural network training efficiency: a comparison between RSSSM and SMAP

To examine the training and simulation efficiency of the neural network, we compare the simulated surface soil moisture (RSSSM) with the training target SMAP during April 2015~2018. The R^2 reaches up to 0.95, while the RMSE is 0.031 m^3/m^3 (Figure 4a). If only the pixels with measured data are considered, the consistency between RSSSM and SMAP becomes even stronger, with an R^2 of 0.97 and an RMSE of 0.016 (Figure 4b). When validated against site measurements, the R^2 and RMSE values are 0.46 and 0.083, respectively, for both RSSSM and SMAP (Figure 4c and 4d). All these findings justify the high training and prediction efficiency of the neural network set designed in this study.

For temporal accuracy, according to Table 2, RSSSM is just slightly lower than SMAP (the differences in the five indicators, r, RMSE, bias, ubRMSE and A.R, are all nonsignificant). Figure 5 indicates generally the same level of temporal accuracy for RSSSM and SMAP under all climates. RSSSM cannot adequately characterize the temporal variation in soil moisture in the 'Dfc' (snow climate, fully humid, see Table 1) region because the training target, SMAP, does not have a high temporal accuracy in this area, probably due to frequent freezing and melting processes.

Next, we compare the spatial accuracy of RSSSM and SMAP. The spatial correlation of RSSSM is somewhat reduced compared to the training target, while the RMSE is slightly increased (Table 3), indicating a subtle loss of detailed spatial information through neural network operation. Because ISMN stations are mostly located in the middle to high latitudes of the Northern Hemisphere, Figure 6 shows that: 1) the accuracy of RSSSM is highest in summers (growing seasons) and lowest in winters, which is inherited from its origin, SMAP, probably due to the impact of freezing on soil moisture retrieval; and 2)

删除了: are…less than 36 …the maximal number of data points is 576), and then with

删除了: rainy …et areas (e.g., rainforests), 20 days without effective precipitation seldom occurs, leading to no results over most regions

删除了: places

删除了: **SIM**

删除了: d

删除了: SIM

删除了: -

删除了: 2a…a). If only the grids …ixels with measured data are considered, the consistency between RSSSMSIM

删除了: is

删除了: 2b…b). When validated against site measurements, the R^2 and RMSE values are 0.46 and 0.083, respectively, for

删除了: ng

删除了: As f

删除了: 3…, RSSSMSIM…is just slightly lower than SMAP

设置了格式: 字体: 倾斜

删除了: 3

删除了: the

删除了: SIM…and SMAP under all climates. RSSSMSIM

删除了: well

删除了: 2

删除了: there

删除了: d

删除了: SIM…and SMAP. The spatial correlation of

删除了: -

删除了: from microwave sensors

RSSSM has a similar spatial accuracy as SMAP in most periods, except for May to June and November to December.

3.2 The accuracy comparison between RSSSM and popular global long-term soil moisture products

3.2.1 The data quality comparison between RSSSM and the satellite-derived product

The satellite-derived global surface soil moisture product, ASCAT-SWI, now covers 12 years, 2007~2018. During that period, the overall R^2 and RMSE for RSSSM are 0.44 and 0.086, respectively (Figure 7), which appear to be much better than those for ASCAT-SWI (R^2=0.33, RMSE=0.100). If the data period of SMAP (2015D10~2018) is excluded, the overall R^2 and RMSE for RSSSM are 0.43 and 0.087, respectively, which are still better than those for ASCAT-SWI (R^2=0.33, RMSE=0.1). However, RSSSM overestimates soil moisture when it is low, which is a problem inherited from the SMAP product (Figure 4), and is a bit nonlinearly correlated with the measured values (Figure 7a).

According to the temporal validation results (Table 4), the evaluation indexes including *r*, RMSE, bias and ubRMSE for RSSSM are all significantly ($p<0.05$) better than those for ASCAT-SWI (anomalies *r* for RSSSM is also higher, but not significant). The temporal accuracy of RSSSM appears to be obviously higher in all climatic zones except for polar areas (Dsb, Dwc and ET). Specifically, in arid areas (BWh and BWk), the temporal correlation coefficients for ASCAT-SWI are low and even negative, but are high for RSSSM (Figure 8).

The spatial accuracy of RSSSM is found significantly higher than ASCAT-SWI when any evaluation index is considered (Table 5). Moreover, the results show that RSSSM is generally superior to ASCAT-SWI throughout the year, especially during the growing seasons (Figure 9).

3.2.2 The data quality comparison between RSSSM and land surface model products

First, the overall accuracies of RSSSM and GLDAS Noah V2.1 surface soil moisture data during 2003~2018 are compared. While RSSSM is nonlinearly correlated with measured soil moisture, the relationship between GLDAS soil moisture and the measurements appears to be slightly more nonlinear, resulting in a smaller R^2 of 0.39 and higher RMSE of 0.097 for GLDAS product compared to RSSSM (R^2: 0.42; RMSE: 0.087, see Figure 10). When excluding the SMAP (training target) data period, the R^2 and RMSE for RSSSM are 0.41 and 0.089, respectively, which are also superior to those for GLDAS (R^2: 0.37; RMSE:

删除了: SIM…shares …as a similar spatial accuracy as SMAP in most periods, but the inferiority of SIM occurs during…

删除了: **SIM**

删除了: **SIM**

删除了: SIM …SSSM are 0.44 and 0.086, respectively (Figure 5…), respectively

删除了: s

删除了: , still better than

删除了: 5…), the evaluation indexes including *R*

设置了格式: 字体: 倾斜

删除了: SIM…are all significantly ($p<0.05$) better than those for ASCAT-SWI (anomalies *R*

设置了格式: 字体: 倾斜

删除了: SIM…is also higher, but not significant). The temporal accuracy of RSSSMSIM…appears to be obviously higher in all climatic zones except for polar areas (Dsb, Dwc and ET). Specifically, in arid areas (BWh and BWk), the temporal correlation coefficients for ASCAT-SWI are low

删除了: SIM…is considered to be

删除了: as well,

删除了: one …valuation index is considered (Table 6…).

删除了: all year round

删除了: 7

删除了: **SIM**

删除了: SIM

删除了: T…e relationship between GLDAS soil moisture

删除了: a bit

删除了: is obviously …onlinear, resulting in a smaller R^2 of

删除了: ,

0.099).

The higher temporal accuracy of RSSSM than GLDAS can be justified by comparing the indicators including *r*, RMSE and ubRMSE (Table 6). The advantage of RSSSM over GLDAS could be identified in almost all climatic regions, especially the cold areas such as BWk, Dfa, Dfc, Dwc and ET (Figure 11), perhaps because the soil thawing and freezing processes are not simulated well. The spatial accuracy of RSSSM indicated by *r*, RMSE, bias and ubRMSE is found to be significantly higher than GLDAS as well (Table 7). The spatial correlation of RSSSM is somewhat higher than that of GLDAS during March to May and September to November, and the spatial RMSE is lower all year round except for January and February (Figure 12). ERA5-Land is a newly published reanalysis-based model product with 0.1° resolution. The overall quality validation (Figure S2) reveals a frequent overestimation of soil moisture by ERA5-Land as well as a nonlinear relationship between the predicted and measured values. Accordingly, although the R^2 for ERA5-Land is 0.41, only slightly lower than that of RSSSM (0.42), the RMSE for ERA5-Land, 0.123, which is much higher than that for RSSSM (0.087) during their common period. Without considering the SMAP period, the conditions are the same (the R^2 for RSSSM and ERA5-Land are 0.41 and 0.38; the RMSE values for these two products are 0.089 and 0.125, respectively). The temporal correlation indicated by *r* and A.R is somewhat higher for ERA5-Land in general (Table S17), but in most cold areas (Dfa, Dwc and ET), the opposite condition occurs (Figure S3a, S3d). The temporal ubRMSE values for RSSSM and ERA5-Land do not differ significantly, but RSSSM usually performs better in relatively arid places (Figure S3c). While the relative temporal accuracies of RSSSM and ERA5-Land are unclear, the spatial pattern of RSSSM is more accurate than that of ERA5-Land considering the significantly better spatial correlation, RMSE, bias and ubRMSE (Table S18). The considerable advantage of RSSSM over ERA5-Land exists throughout the year, especially during the growing seasons from March to November (Figure S4).

3.2.3 The data quality comparison between RSSSM and the soil moisture products derived from both satellite data and model simulations

CCI is a typical surface soil moisture dataset developed by combining satellite observations and model simulations. However, validation against measurements indicates that the CCI product is not of very good quality; the overall R^2 is only 0.31 with an RMSE value of up to 0.095 (Figure S5, when the SMAP data period is excluded, the R^2 and RMSE for CCI are 0.28 and 0.098,

删除了: supposed …igher temporal accuracy of RSSSMSIM…than GLDAS can be justified by comparing the indicators including *R*

设置了格式: 字体: 倾斜

删除了: 7…). The advantage of RSSSMSIM…over GLDAS could be identified in almost all climatic regions, especially the cold areas such as BWk, Dfa, Dfc, Dwc and ET (Figure 11), perhaps because the soil thawing and freezing processes are not simulated well in those regions…
The spatial accuracy of RSSSM,SIM, which is…indicated by *R*…

设置了格式: 字体: 倾斜

删除了: ,…is expected to be…ound to be significantly

删除了: 3

删除了: just

删除了: SIM… the RMSE for ERA5-Land, 0.123, which is

删除了: of

删除了: 0.123

删除了: condition is

删除了: indicators (…ndicated by *R*

设置了格式: 字体: 倾斜

删除了:) for ERA5-Land are …is somewhat higher than

删除了: accuracy

删除了: SIM

删除了: is

删除了: SIM…ought to be…s more accurate than that of

删除了: all year round,

删除了: S5

删除了: **SIM**

删除了: s…and model simulations. However, validation compared to 0.41 and 0.089 for RSSSM). The temporal pattern of RSSSM, indicated by *r* and RMSE, is found to be significantly better than CCI (Table S19), and under all climate conditions (Figure S6). Our results indicate that RSSSM also shows a consistently higher spatial accuracy than CCI, especially during the growing seasons (Table S20 and Figure S7).

Next, we focus on the interannual change in data quality. According to Figure 13a~c, while the correlation coefficient for RSSSM does not vary significantly among different years, the RMSE and ubRMSE values in earlier periods are somewhat raised compared to those after 2012. Though the data quality of RSSSM can hardly be maintained as well, the degradation degree is much slighter than that of CCI. The comparison of the spatial coverages of the 10-day scale RSSSM and CCI data (rainforests are excluded) shows that RSSSM covers all land surfaces except for permafrost, while the interannual variation in coverage is also negligible throughout the entire period (the intra-annual cycles of data coverages result from the changes in frozen areas), which are preferable to CCI, whose data coverage before 2007 is limited (Figure 13d).

GLEAM products also contain satellite information due to the assimilation of CCI data, but model simulation plays a much more important role. By validation, the overall R^2 and RMSE values for the GLEAM v3.3a product (2003~2018) are 0.38 and 0.142, whereas those for the v3.3b product are 0.36 and 0.13, respectively. Both estimates are nonlinearly correlated with and are generally higher than the measured values (Figure S8). Therefore, with an R^2 of 0.42 and an RMSE of 0.087, RSSSM is found to be superior to GLEAM v3.3a/b in general (if the SMAP data period is excluded, RSSSM's R^2 and RMSE values are 0.41 and 0.089, respectively, which are still better than both GLEAM v3.3a (R^2: 0.35; RMSE: 0.141) and GLEAM v3.3a (R^2: 0.34; RMSE: 0.128)). The temporal and spatial accuracies of GLEAM products and RSSSM are compared in Tables S21~S24. The advantage of GLEAM is its ability to characterize the temporal variations in soil moisture, with higher temporal correlation achieved in most climatic regions (Figure S9a and S9d). However, the main potential disadvantage is the obvious overestimation, which leads to significantly higher RMSE values than RSSSM in all regions and all periods (Figure S9b and Figure S10b). Moreover, the spatial pattern of GLEAM products is less convincing than that of RSSSM, considering the lower spatial correlation coefficients, especially in spring (March to May) and autumn (September to November) (Figure S10a). Therefore, the potential advantages of RSSSM can exceed those of GLEAM.

In conclusion, surface soil moisture developed mainly based on land surface models (GLEAM and ERA5-Land) has high

删除了: SIM…SSSM, indicated by *R*

删除了: supposed to be…ound to be significantly better than CCI (Table S19), and under all climate conditions (Figure S7…

设置了格式: 字体: 倾斜

删除了: s

删除了: SIM

删除了: occupies

删除了: S8

删除了: 11a…3a~c, while the correlation coefficient for SIM …SSSM does not vary significantly among different years, the RMSE and ubRMSE values in earlier periods are somewhat raised compared to those after 2012. The main reasons are as follows: 1) five rounds of simulations were performed, with the output converted into the training target of the next round's neural networks. Hence, as the simulation…

删除了: By comparing

删除了: SIM

删除了: , it is shown

删除了: SIM …SSSM covers all land surfaces except for

删除了: coverages' interannual variation in coverage is

删除了: …Figure 11d…3d) and can thus ensure a supposed…

删除了: S9…8). Therefore, with an R^2 of 0.42 and an RMSE…

删除了: ,

删除了: SIM …SSSM are compared in Tables S21~S24. The…

删除了: of

删除了: R …orrelation achieved in most climatic regions

删除了: S…ring (March to May) and aA

删除了: S11a…10a). Therefore, the potential advantages of…

删除了: have

temporal accuracy, but their absolute values and spatial patterns are relatively unreliable, whereas RSSSM shows good performances in all aspects. Generally, this study indicates that the expected order of data applicability among various global long-term surface soil moisture products is RSSSM (applicable to all studies)> GLEAM (suitable for temporal variation studies)> ERA5-Land (applicable to temporal pattern studies)> GLDAS Noah V2.1 (somewhat applicable to all studies)> ASCAT-SWI> CCI. The training R^2 of the previous neural networks designed for global surface soil moisture mapping is 0.45~0.55, while the temporal *r* and RMSE values against measurements are 0.52 and 0.084 (Yao et al., 2017), and the overall R^2 and RMSE are 0.2 and 0.113 (Yao et al., 2019). In this study, by elaborating the neural network, the training R^2 is elevated to 0.95, with improvements also to the temporal *r* and RMSE (0.69 and 0.08) as well as overall R^2 and RMSE (0.42 and 0.087) values. In addition, our 10-day period average product is both spatially and temporally continuous over 16 years, with a high spatial resolution, and covers all land except for frozen ground. Hence, our product could be more useful than previous machine learning products.

3.3 The spatial and temporal patterns of the calculated surface soil moisture

For the calculated global surface soil moisture, the spatial pattern averaged during 2003~2018 is shown in Figure 14a (the maps for separate months are shown in Figure S11a). The above validation results show that except for RSSSM, GLDAS has the highest spatial accuracy, so the spatial map of GLDAS surface moisture is attached below (Figure S11b). By comparison, the spatial patterns of RSSSM and GLDAS are similar, but some differences also exist (see the regions circled in red). Obviously, RSSSM has a higher spatial heterogeneity and probably more reflections on wetlands and irrigated fields (e.g., the Hetao Irrigation Area in China), whereas GLDAS appears patchy in arid areas. The latitudinal pattern comparison in Figure S12a also implies that RSSSM contains more detailed spatial information.

For the interannual variation, because the GLEAM v3.3a product is proven to have the best accuracy in characterizing the temporal anomalies of soil moisture, and also covers the whole world, this product is selected as the reference to justify our calculation. According to Figure S12b, both GLEAM and RSSSM support a significant rising trend in global mean surface soil moisture during 2003~2018, while the average rates are both approximately 0.03 m^3 m^{-3} yr^{-1} (Figure S12b). The spatial

删除了: SIM …SSSM shows good performances in all aspects. Generally, this study indicates that the expected order of data applicability among various global long-term surface soil moisture products is SIM

删除了: studies of

删除了: *R*

设置了格式: 字体: 倾斜

设置了格式: 上标

删除了: 44

删除了: with the

设置了格式: 字体: 倾斜

删除了: *R* ... and RMSE (0.69 and 0.08) as well as overall *R*

设置了格式: 字体: 倾斜

删除了: 65

删除了: also improved… In addition, our 10-day period average product is both spatially and temporally continuous

删除了: it …overs all land except for frozen ground (data are

删除了: -…earninged

删除了: 12a …4a (the maps for separate months are shown

删除了: A

删除了: V…lidation results shows…that apart from…except

删除了: there also exist some differences

删除了: SIM…has a higher spatial heterogeneity and,

删除了: places

删除了: S13a …12a also implies that RSSSMSIM

删除了: As f…or the interannual variation, because the

删除了: which…and also covers the globe…hole world,

删除了: around

删除了: S13b

patterns of the interannual trends in RSSSM and GLEAM are shown in Figure 14c~d, which are generally consistent. Soil moisture gains are found over the border between the USA and Canada, as well as over Paraguay, Kazakhstan, Northeastern and Southern China (the regions circled in blue), while soil moisture declines took place in North Asia and eastern Brazil (the regions with red circles). The main discrepancy between the soil moisture trends predicted by the two products lies in Central Africa, the Arabian Peninsula and northwestern Canada.

Because the validation against measurements proves that the intra-annual soil moisture variation in the 'Dfc' climate region cannot be captured by SMAP or RSSSM, the acquired intra-annual analysis results in this region are not considered. Over low-latitude areas (30°S~30°N), surface soil moisture peaks in summers (seasons are opposite in the Northern and Southern Hemispheres); however, in most of the midlatitude areas (30°S~60°S; 30°N~60°N) except for eastern Asia (i.e., east of the Yenisei River), the soil moisture is high in winters (nongrowing seasons) and low in summers (Figure 15a and Figure S13a). The intra-annual range of surface moisture is largest in the tropical monsoon climate regions, including the African savannas, the Orinoco Plain, the Ganges plain and the plains in the Indochina Peninsula, as well as some seasonal frozen areas, whereas it is lowest in arid places (Figure 15b; Figure S13b~c). Precipitation is a direct driver of surface soil moisture changes (Figure S14a~b), and the intra-annual cycle of soil moisture often strictly follows that of precipitation as long as it exists (Figure 15c and Figure S14c). Therefore, considering that precipitation is highest in summer at low latitudes, where plants often grow in all seasons, whereas in the westerlies, rainfall is temporally even (eastern Asia is an exception perhaps due to monsoon and topographic conditions) yet with much higher evapotranspiration in summer, the global intra-annual patterns of soil moisture can be explained. The peak time difference between surface moisture and precipitation is approximately one 10-day period, or six days on average at global scale (Figure 15d), which is expected to be related to the 'time lag' effect. On dry days, the fastest surface moisture decline is expected in summers when evapotranspiration is high. However, this study reveals that at midlatitudes, the opposite condition occurs: the surface water loss without rain is lowest in summer (Figure 15e and Figure S15a). Further analysis proves a positive correlation between surface moisture and its rate of decline, with r>0.8 over 85% of the area (Figure S15b~c), indicating that because soil moisture in the westerlies is often high in winters, the available surface water for evaporation and percolation loss is limited in summer, and plants tend to utilize water in deeper soil layers. When

删除了: SIM
删除了: 12c
删除了: both
删除了: area
删除了: SIM
删除了: -
删除了: -
删除了: 13a
删除了: S13a
删除了: 13b
删除了: S14b
删除了: S15a
删除了: 13c
删除了: S15c
删除了: in the
删除了: 13d
删除了: one may expect
删除了: in
删除了: 13e
删除了: S16a
删除了: *R*
设置了格式: 字体: 倾斜
删除了: places
删除了: S16b
删除了: er

droughts occurred during a random period, the mean surface moisture decline is highest in the tropical monsoon climate regions (Figure 15f). Therefore, if sufficient water during rainy seasons is lacking there, significant water loss (Figure S15d) may destroy the local ecosystem.

4 Discussion and conclusions

In this study, an improved global long-term satellite-based surface soil moisture dataset, RSSSM, was developed mainly based on 11 microwave soil moisture products. Our product is more comparable to the in situ measurements at ISMN stations than most existing popular surface soil moisture datasets. Our product is temporally continuous during 2003~2018, and covers the whole globe except for frozen grounds (CCI has limited spatial coverage before 2007, when ASCAT data are unavailable), ensuring its applicability to global long-term studies or ecosystem modeling. However, the achieved accuracy (R^2=0.42; RMSE=0.087) for surface soil moisture is still lower than that for many other terrestrial essential climate variables. The target RMSE for surface soil moisture set by GCOS is 0.04 m^3 m^{-3}, which is much lower than the value met in this study, indicating the need to further improve the global soil moisture data quality.

Fortunately, this study provides a novel approach that has the potential to create increasingly better soil moisture products in future. The RMSE and ubRMSE values in earlier periods are somewhat higher than those after 2012, which is because: 1) five rounds of simulations were performed, with the output converted into the training target of the next round's neural networks, leading to a little error propagation as the simulation period extended to the past; and 2) the quality of microwave soil moisture data is generally lower in earlier periods due to the relatively unadvanced microwave sensors with low signal-to-noise ratio (SNR). However, due to the elaborate design of the neural network set (localized networks, full use of 11 microwave soil moisture products, the determination of quality impact factors and the organization of 67 independent neural networks), high training efficiency is achieved, resulting in limited amplification of noise and high maintenance of valid information during 16 years of simulation. This method turns out to be better than the simple CDF matching algorithm, which may not efficiently calibrate the low-quality soil moisture data retrieved from earlier sensors. The overall data accuracy of RSSSM is only slightly lower than that of SMAP, the primary training target. Therefore, if microwave sensors with higher SNR or better penetration

删除了: 13f

删除了: if it lacks enough water during rainy seasons …here, the …

删除了: S16d

删除了: ruin

删除了: new …lobal long-term satellite-based surface soil moisture dataset, RSSSM, was developed mainly by fusing

删除了: -

删除了: Through validation against …easurements at ISMN stations, our product proved to be superior to all…than the

删除了: It

删除了: -…ontinuous during 2003~2018, and covers the whole globe except for frozen grounds (CCI has limited spatial coverage before 2007, when ASCAT data areis…unavailable), ensuring its applicability toin…global

删除了: …owever, the achieved accuracy (R^2=0.42;

删除了: what we have met

设置了格式: 字体: 10 磅

删除了: provided

删除了: ing…d to the past; and 2) the quality of microwave

删除了: D

删除了: d

删除了: multiple rounds of simulation, various sources of…1

删除了: ttle

删除了: little

删除了: s

删除了: the …igh maintenance of valid

删除了: -

删除了: …raining target)…

of vegetation canopy than SMAP are launched in the future (for example, the upcoming P-band microwave sensors (Etminan et al., 2020; Ye et al., 2020)), we can develop a temporally continuous soil moisture dataset beginning in 2003 by using the soil moisture or Tb retrieved from the new sensors as the reference. This upcoming product is expected to have even higher accuracy than the SMAP product (we will update the complete RSSSM product then). In that sense, the data fusion algorithm proposed here will be very meaningful in the future.

Another way to improve global surface soil moisture data accuracy as well as the temporal resolution is to combine satellite-based products with land surface models such as GLEAM. Remote sensing inversion can delineate more detailed spatial information on soil moisture, whereas reanalysis-based models have advantages in characterizing temporal variations, and even on a daily scale, except for irrigated croplands. Furthermore, because root-zone soil moisture is the direct factor influencing vegetation growth, it often plays a more important role than surface moisture in ecosystems; however, this factor cannot be obtained from microwave remote sensing. Hence, combining the advantages of observation and model simulation helps to improve the data accuracies of both surface and root-zone soil moisture. Unfortunately, while the CCI algorithm integrates the disadvantages of both methods, GLEAM incorporated only very limited observational information. We propose that one possible approach is to use the pixel-specific confidence range and the spatial pattern of satellite-based soil moisture (e.g., our product: RSSSM) to constrain the model parameters or add supplementary modules if necessary. In detail, RSSSM can be used as the initial surface soil moisture map. Then, after each time of soil moisture simulation in multiple layers (both root-zone and surface), the model efficiency is examined through a spatial correlation test between the simulated surface moisture and RSSSM. In addition, whether the simulated value falls within the confidence range (e.g., ±20%) of that reported by RSSSM should also be tested. By recurrent adjustments, the model parameters in each pixel can be optimized. For irrigated croplands, if irrigation is not considered in the models, the simulated surface soil moisture will soon fall below the confidence range, and the spatial correlation will also decline regardless of the parameters that are provided. Therefore, a well-designed irrigation module (Chen et al., 2019) should be introduced. Finally, for regions with massive human-induced land cover changes (e.g., afforestation), optical remote sensing should be applied for better estimation of evapotranspiration.

删除了: by using the soil moisture or Tb retrieved from the new sensors as the reference, we can develop a temporally continuous soil moisture dataset since 2003,

删除了: which

删除了: the

删除了: influential

删除了: of

删除了: ,

删除了: which

删除了: only

删除了: observed

删除了: grid

删除了: SIM

删除了: SIM

删除了: SIM

删除了: SIM

删除了: grid

删除了: no matter which parameters are given

5 Data availability

The global surface soil moisture dataset, RSSSM, is available at: https://doi.pangaea.de/10.1594/PANGAEA.912597 (Chen, 2020).

Author contributions

Yongzhe Chen conducted the research, completed the original draft and revised it. The corresponding author, Xiaoming Feng, supervised the research and revised the draft. Bojie Fu administered the project and funded the research. All co-authors reviewed the manuscript and contributed to the writing process.

删除了: administrated

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the National Key Research and Development Program of China (NO. 2017YFA0604700), the National Science Foundation of China (41722104) and the Chinese Academy of Sciences (QYZDY-SSW-DQC025). We are also grateful to all the data contributors who made it possible to complete this work.

References

Al-Yaari, A., Wigneron, J. P., Dorigo, W., Colliander, A., Pellarin, T., Hahn, S., Mialon, A., Richaume, P., Fernandez-Moran, R., Fan, L., Kerr, Y. H., and De Lannoy, G.: Assessment and inter-comparison of recently developed/reprocessed microwave satellite soil moisture products using ISMN ground-based measurements, Remote Sens. Environ., 224, 289-303, https://doi.org/10.1016/j.rse.2019.02.008, 2019

Al-Yaari, A., Wigneron, J. P., Kerr, Y., de Jeu, R., Rodriguez-Fernandez, N., van der Schalie, R., Al Bitar, A., Mialon, A., Richaume, P., Dolman, A., and Ducharne, A.: Testing regression equations to derive long-term global soil moisture datasets from passive microwave observations, Remote Sens. Environ., 180, 453-464, https://doi.org/10.1016/j.rse.2015.11.022, 2016

Albergel, C., de Rosnay, P., Gruhier, C., Muñoz-Sabater, J., Hasenauer, S., Isaksen, L., Kerr, Y., and Wagner, W.: Evaluation of remotely sensed and modelled soil moisture products using global ground-based in situ observations, Remote Sens. Environ., 118, 215-226, https://doi.org/10.1016/j.rse.2011.11.017, 2012

Albergel, C., Rüdiger, C., Pellarin, T., Calvet, J. C., Fritz, N., Froissard, F., Suquia, D., Petitpa, A., Piguet, B., and Martin, E.: From near-surface to root-zone soil moisture using an exponential filter: an assessment of the method based on in-situ observations and model simulations, Hydrol. Earth Syst. Sci., 12, 1323-1337, http://doi.org/10.5194/hess-12-1323-2008, 2008

Baldocchi, D., Tang, J., and Xu, L.: How switches and lags in biophysical regulators affect spatial-temporal variation of soil respiration in an oak-grass savanna, Journal of Geophysical Research: Biogeosciences, 111, https://doi.org/10.1029/2005JG000063, 2006

Balsamo, G., Albergel, C., Beljaars, A., Boussetta, S., Brun, E., Cloke, H., Dee, D., Dutra, E., Muñoz-Sabater, J., Pappenberger, F., de Rosnay, P., Stockdale, T., and Vitart, F.: ERA-Interim/Land: a global land surface reanalysis data set, Hydrol. Earth Syst. Sci., 19, 389-407, http://doi.org/10.5194/hess-19-389-2015, 2015

Baret, F., Weiss, M., Lacaze, R., Camacho, F., Makhmara, H., Pacholcyzk, P., and Smets, B.: GEOV1: LAI and FAPAR essential climate variables and FCOVER global time series capitalizing over existing products. Part1: Principles of development and production, Remote Sens. Environ., 137, 299-309, https://doi.org/10.1016/j.rse.2012.12.027, 2013

Bartalis, Z., Wagner, W., Naeimi, V., Hasenauer, S., Scipal, K., Bonekamp, H., Figa, J., and Anderson, C.: Initial soil moisture retrievals from the METOP-A Advanced Scatterometer (ASCAT), Geophys. Res. Lett., 34, http://doi.org/10.1029/2007GL031088, 2007

Bi, H., Ma, J., Zheng, W., and Zeng, J.: Comparison of soil moisture in GLDAS model simulations and in situ observations over the Tibetan Plateau, J. Geophys. Res.-Atmos., 121, 2658-2678, http://doi.org/10.1002/2015JD024131, 2016

Bojinski, S., Verstraete, M., Peterson, T. C., Richter, C., Simmons, A., and Zemp, M.: The Concept of Essential Climate Variables in Support of Climate Research, Applications, and Policy, Bull. Amer. Meteorol. Soc., 95, 1431-1443, http://doi.org/10.1175/BAMS-D-13-00047.1, 2014

Brooks, E. B., Thomas, V. A., Wynne, R. H., and Coulston, J. W.: Fitting the Multitemporal Curve: A Fourier Series Approach to the Missing Data Problem in Remote Sensing Analysis, IEEE Trans. Geosci. Remote Sensing, 50, 3340-3353, http://doi.org/10.1109/TGRS.2012.2183137, 2012

Burgin, M., Colliander, A., Njoku, E., Chan, S., Cabot, F., Kerr, H. Y., Bindlish, R., Jackson, T., Entekhabi, D., and Yueh, S.:

删除了:

设置了格式: 字体: 10 磅

A Comparative Study of the SMAP Passive Soil Moisture Product With Existing Satellite-Based Soil Moisture Products, IEEE Trans. Geosci. Remote Sensing, 55, 2959-2971, http://doi.org/10.1109/TGRS.2017.2656859, 2017

Camacho, F., Cernicharo, J., Lacaze, R., Baret, F., and Weiss, M.: GEOV1: LAI, FAPAR essential climate variables and FCOVER global time series capitalizing over existing products. Part 2: Validation and intercomparison with reference products, Remote Sens. Environ., 137, 310-329, https://doi.org/10.1016/j.rse.2013.02.030, 2013

Cammalleri, C., Verger, A., Lacaze, R., and Vogt, J. V.: Harmonization of GEOV2 fAPAR time series through MODIS data for global drought monitoring, Int. J. Appl. Earth Obs. Geoinf., 80, 1-12, https://doi.org/10.1016/j.jag.2019.03.017, 2019

Chan, S. K., Bindlish, R., O'Neill, P., Jackson, T., Njoku, E., Dunbar, S., Chaubell, J., Piepmeier, J., Yueh, S., Entekhabi, D., Colliander, A., Chen, F., Cosh, M. H., Caldwell, T., Walker, J., Berg, A., McNairn, H., Thibeault, M., Martínez-Fernández, J., Uldall, F., Seyfried, M., Bosch, D., Starks, P., Holifield Collins, C., Prueger, J., van der Velde, R., Asanuma, J., Palecki, M., Small, E. E., Zreda, M., Calvet, J., Crow, W. T., and Kerr, Y.: Development and assessment of the SMAP enhanced passive soil moisture product, Remote Sens. Environ., 204, 931-941, https://doi.org/10.1016/j.rse.2017.08.025, 2018

Chen, F., Crow, W. T., Bindlish, R., Colliander, A., Burgin, M. S., Asanuma, J., and Aida, K.: Global-scale evaluation of SMAP, SMOS and ASCAT soil moisture products using triple collocation, Remote Sens. Environ., 214, 1-13, https://doi.org/10.1016/j.rse.2018.05.008, 2018

Chen, Y.: A new dataset of satellite observation-based global surface soil moisture covering 2003-2018, https://doi.org/10.1594/PANGAEA.912597, 2020

Chen, Y., Feng, X., Fu, B., Shi, W., Yin, L., and Lv, Y.: Recent Global Cropland Water Consumption Constrained by Observations, Water Resour. Res., 55, 3708-3738, http://doi.org/10.1029/2018WR023573, 2019

Chen, Y., Yang, K., Qin, J., Zhao, L., Tang, W., and Han, M.: Evaluation of AMSR-E retrievals and GLDAS simulations against observations of a soil moisture network on the central Tibetan Plateau, J. Geophys. Res.-Atmos., 118, 4466-4475, http://doi.org/10.1002/jgrd.50301, 2013

Colliander, A., Jackson, T. J., Bindlish, R., Chan, S., Das, N., Kim, S. B., Cosh, M. H., Dunbar, R. S., Dang, L., Pashaian, L., Asanuma, J., Aida, K., Berg, A., Rowlandson, T., Bosch, D., Caldwell, T., Caylor, K., Goodrich, D., al Jassar, H., Lopez-Baeza, E., Martínez-Fernández, J., González-Zamora, A., Livingston, S., McNairn, H., Pacheco, A., Moghaddam, M., Montzka, C., Notarnicola, C., Niedrist, G., Pellarin, T., Prueger, J., Pulliainen, J., Rautiainen, K., Ramos, J., Seyfried, M., Starks, P., Su, Z., Zeng, Y., van der Velde, R., Thibeault, M., Dorigo, W., Vreugdenhil, M., Walker, J. P., Wu, X., Monerris, A., O'Neill, P. E., Entekhabi, D., Njoku, E. G., and Yueh, S.: Validation of SMAP surface soil moisture products with core validation sites, Remote Sens. Environ., 191, 215-231, https://doi.org/10.1016/j.rse.2017.01.021, 2017

Crow, W. T., Wagner, W., and Naeimi, V.: The Impact of Radar Incidence Angle on Soil-Moisture-Retrieval Skill, IEEE Geosci. Remote Sens. Lett., 7, 501-505, https://doi.org/10.1109/LGRS.2010.2040134, 2010

Das, N. and O'neill, P.: Selection of Soil Attributes Datasets for the SMAP Mission, 2010.

Dobson, M. C., Ulaby, F. T., Hallikainen, M. T., and El-rayes, M. A.: Microwave Dielectric Behavior of Wet Soil-Part II: Dielectric Mixing Models, IEEE Trans. Geosci. Remote Sensing, GE-23, 35-46, https://doi.org/10.1109/TGRS.1985.289498, 1985

Dorigo, W., de Jeu, R., Chung, D., Parinussa, R., Liu, Y., Wagner, W., and Fernández-Prieto, D.: Evaluating global trends

(1988–2010) in harmonized multi-satellite surface soil moisture, Geophys. Res. Lett., 39, https://doi.org/10.1029/2012GL052988, 2012

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

Dorigo, W. A., Gruber, A., De Jeu, R. A. M., Wagner, W., Stacke, T., Loew, A., Albergel, C., Brocca, L., Chung, D., Parinussa, R. M., and Kidd, R.: Evaluation of the ESA CCI soil moisture product using ground-based observations, Remote Sens. Environ., 162, 380-395, https://doi.org/10.1016/j.rse.2014.07.023, 2015

Dorigo, W. A., Scipal, K., Parinussa, R. M., Liu, Y. Y., Wagner, W., de Jeu, R. A. M., and Naeimi, V.: Error characterisation of global active and passive microwave soil moisture datasets, Hydrol. Earth Syst. Sci., 14, 2605-2616, https://doi.org/10.5194/hess-14-2605-2010, 2010

Dorigo, W. A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A., Drusch, M., Mecklenburg, S., van Oevelen, P., Robock, A., and Jackson, T.: The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements, Hydrol. Earth Syst. Sci., 15, 1675-1698, https://doi.org/10.5194/hess-15-1675-2011, 2011

Dorigo, W. A., Xaver, A., Vreugdenhil, M., Gruber, A., Hegyiová, A., Sanchis-Dufau, A. D., Zamojski, D., Cordes, C., Wagner, W., and Drusch, M.: Global Automated Quality Control of In Situ Soil Moisture Data from the International Soil Moisture Network, Vadose Zone J., 12, https://doi.org/10.2136/vzj2012.0097, 2013

Draper, C. S., Reichle, R. H., De Lannoy, G. J. M., and Liu, Q.: Assimilation of passive and active microwave soil moisture retrievals, Geophys. Res. Lett., 39, http://doi.org/10.1029/2011GL050655, 2012

Du, Y., Ulaby, F. T., and Dobson, M. C.: Sensitivity to soil moisture by active and passive microwave sensors, IEEE Trans. Geosci. Remote Sensing, 38, 105-114, https://doi.org/10.1109/36.823905, 2000

Entekhabi, D., Njoku, E., O'neill, P., Kellogg, K. H., Crow, W., Edelstein, W. N., Entin, J. K., Goodman, S. D., Jackson, T., Johnson, F. M., Kimball, J., Piepmeier, J., Koster, R. D., Martin, E., McDonald, C. K., Moghaddam, M., Moran, M. S., Reichle, R., Shi, J. C., Spencer, D., Thurman, S. W., Tsang, L., and Zyl, J. V.: The Soil Moisture Active Passive (SMAP) Mission, Proc. IEEE, 98, 704-716, https://doi.org/10.1109/JPROC.2010.2043918, 2010

Etminan, A., Tabatabaeenejad, A., and Moghaddam, M.: Retrieving Root-Zone Soil Moisture Profile From P-Band Radar via Hybrid Global and Local Optimization, IEEE Trans. Geosci. Remote Sensing, doi: https://doi.org/10.1109/TGRS.2020.2965569, 2020. 1-9, https://doi.org/10.1109/TGRS.2020.2965569, 2020

Fan, X., Liu, Y., Gan, G., and Wu, G.: SMAP underestimates soil moisture in vegetation-disturbed areas primarily as a result of biased surface temperature data, Remote Sens. Environ., 247, 111914, https://doi.org/10.1016/j.rse.2020.111914, 2020

Feng, X., Li, J., Cheng, W., Fu, B., Wang, Y., Lü, Y., and Shao, M. A.: Evaluation of AMSR-E retrieval by detecting soil moisture decrease following massive dryland re-vegetation in the Loess Plateau, China, Remote Sens. Environ., 196, 253-264, https://doi.org/10.1016/j.rse.2017.05.012, 2017

Fernandez-Moran, R., Al-Yaari, A., Mialon, A., Mahmoodi, A., Al Bitar, A., De Lannoy, G., Rodriguez-Fernandez, N., Lopez-

Baeza, E., Kerr, Y., and Wigneron, J.-P.: SMOS-IC: An Alternative SMOS Soil Moisture and Vegetation Optical Depth Product, Remote Sens., 9, https://doi.org/10.3390/rs9050457, 2017a

Fernandez-Moran, R., Wigneron, J. P., De Lannoy, G., Lopez-Baeza, E., Parrens, M., Mialon, A., Mahmoodi, A., Al-Yaari, A., Bircher, S., Al Bitar, A., Richaume, P., and Kerr, Y.: A new calibration of the effective scattering albedo and soil roughness parameters in the SMOS SM retrieval algorithm, Int. J. Appl. Earth Obs. Geoinf., 62, 27-38, https://doi.org/10.1016/j.jag.2017.05.013, 2017b

Fujii, H., Koike, T., and Imaoka, K.: Improvement of the AMSR-E Algorithm for Soil Moisture Estimation by Introducing a Fractional Vegetation Coverage Dataset Derived from MODIS Data, J. Meteorol. Soc. Japan, 29, 282-292, https://doi.org/10.11440/rssj.29.282, 2009

Gaiser, P. W., Germain, K. M. S., Twarog, E. M., Poe, G. A., Purdy, W., Richardson, D., Grossman, W., Jones, W. L., Spencer, D., Golba, G., Cleveland, J., Choy, L., Bevilacqua, R. M., and Chang, P. S.: The WindSat spaceborne polarimetric microwave radiometer: sensor description and early orbit performance, IEEE Trans. Geosci. Remote Sensing, 42, 2347-2361, https://doi.org/10.1109/TGRS.2004.836867, 2004

Gao, H., Wood, E. F., Jackson, T. J., Drusch, M., and Bindlish, R.: Using TRMM/TMI to Retrieve Surface Soil Moisture over the Southern United States from 1998 to 2002, J. Hydrometeorol., 7, 23-38, https://doi.org/10.1175/JHM473.1, 2006

Gao, L., Sadeghi, M., and Ebtehaj, A.: Microwave retrievals of soil moisture and vegetation optical depth with improved resolution using a combined constrained inversion algorithm: Application for SMAP satellite, Remote Sens. Environ., 239, 111662, https://doi.org/10.1016/j.rse.2020.111662, 2020

Griend, A. A. V. d. and Wigneron, J.: On the measurement of microwave vegetation properties: some guidelines for a protocol, IEEE Trans. Geosci. Remote Sensing, 42, 2277-2289, https://doi.org/10.1109/TGRS.2004.832243, 2004

Gruber, A., De Lannoy, G., Albergel, C., Al-Yaari, A., Brocca, L., Calvet, J. C., Colliander, A., Cosh, M., Crow, W., Dorigo, W., Draper, C., Hirschi, M., Kerr, Y., Konings, A., Lahoz, W., McColl, K., Montzka, C., Muñoz-Sabater, J., Peng, J., Reichle, R., Richaume, P., Rüdiger, C., Scanlon, T., van der Schalie, R., Wigneron, J. P., and Wagner, W.: Validation practices for satellite soil moisture retrievals: What are (the) errors?, Remote Sens. Environ., 244, 111806, https://doi.org/10.1016/j.rse.2020.111806, 2020

Gruber, A., Dorigo, W. A., Crow, W., and Wagner, W.: Triple Collocation-Based Merging of Satellite Soil Moisture Retrievals, IEEE Trans. Geosci. Remote Sensing, 55, 6780-6792, https://doi.org/10.1109/TGRS.2017.2734070, 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

Gu, X., Li, J., Chen, Y. D., Kong, D., and Liu, J.: Consistency and Discrepancy of Global Surface Soil Moisture Changes From Multiple Model-Based Data Sets Against Satellite Observations, J. Geophys. Res.-Atmos., 124, 1474-1495, https://doi.org/10.1029/2018JD029304, 2019

Hansen, M. and Song, X. P.: Vegetation Continuous Fields (VCF) Yearly Global 0.05 Deg, NASA EOSDIS Land Processes DAAC, https://doi.org/10.5067/MEaSUREs/VCF/VCF5KYR.001, 2018

Hashimoto, S., Carvalhais, N., Ito, A., Migliavacca, M., Nishina, K., and Reichstein, M.: Global spatiotemporal distribution of soil respiration modeled using a global database, Biogeosciences, 12, 4121-4132, https://doi.org/10.5194/bg-12-4121-2015, 2015

Hermance, J. F., Jacob, R. W., Bradley, B. A., and Mustard, J. F.: Extracting Phenological Signals From Multiyear AVHRR NDVI Time Series: Framework for Applying High-Order Annual Splines With Roughness Damping, IEEE Trans. Geosci. Remote Sensing, 45, 3264-3276, http://doi.org/10.1109/TGRS.2007.903044, 2007

Hoffmann, L., Günther, G., Li, D., Stein, O., Wu, X., Griessbach, S., Heng, Y., Konopka, P., Müller, R., Vogel, B., and Wright, J. S.: From ERA-Interim to ERA5: the considerable impact of ECMWF's next-generation reanalysis on Lagrangian transport simulations, Atmos. Chem. Phys., 19, 3097-3124, https://doi.org/10.5194/acp-19-3097-2019, 2019

Holmes, T. R. H., De Jeu, R. A. M., Owe, M., and Dolman, A. J.: Land surface temperature from Ka band (37 GHz) passive microwave observations, J. Geophys. Res.-Atmos., 114, https://doi.org/10.1029/2008JD010257, 2009

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., and Ferreira, L. G.: Overview of the radiometric and biophysical performance of the MODIS vegetation indices, Remote Sens. Environ., 83, 195-213, https://doi.org/10.1016/S0034-4257(02)00096-2, 2002

Huffman, G., Bolvin, D., Braithwaite, D., Hsu, K., Joyce, R., and Xie, P.: Integrated Multi-satellitE Retrievals for GPM (IMERG), version 6, NASA's Precipitation Processing Center, 2019

Jackson, T. J. and Schmugge, T. J.: Vegetation effects on the microwave emission of soils, Remote Sens. Environ., 36, 203-212, https://doi.org/10.1016/0034-4257(91)90057-D, 1991

Jackson, T. J., Schmugge, T. J., and Wang, J. R.: Passive microwave sensing of soil moisture under vegetation canopies, Water Resour. Res., 18, 1137-1142, https://doi.org/10.1029/WR018i004p01137, 1982

Jiao, Q., Li, R., Wang, F., Mu, X., Li, P., and An, C.: Impacts of Re-Vegetation on Surface Soil Moisture over the Chinese Loess Plateau Based on Remote Sensing Datasets, Remote Sens., 8, https://doi.org/10.3390/rs8020156, 2016

Karthikeyan, L., Pan, M., Wanders, N., Kumar, D. N., and Wood, E. F.: Four decades of microwave satellite soil moisture observations: Part 1. A review of retrieval algorithms, Adv. Water Resour., 109, 106-120, https://doi.org/10.1016/j.advwatres.2017.09.006, 2017a

Karthikeyan, L., Pan, M., Wanders, N., Kumar, D. N., and Wood, E. F.: Four decades of microwave satellite soil moisture observations: Part 2. Product validation and inter-satellite comparisons, Adv. Water Resour., 109, 236-252, https://doi.org/10.1016/j.advwatres.2017.09.010, 2017b

Kawanishi, T., Sezai, T., Ito, Y., Imaoka, K., Takeshima, T., Ishido, Y., Shibata, A., Miura, M., Inahata, H., and Spencer, R. W.: The Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E), NASDA's contribution to the EOS for global energy and water cycle studies, IEEE Trans. Geosci. Remote Sensing, 41, 184-194, https://doi.org/10.1109/TGRS.2002.808331, 2003

Kerr, Y. H., Al-Yaari, A., Rodriguez-Fernandez, N., Parrens, M., Molero, B., Leroux, D., Bircher, S., Mahmoodi, A., Mialon, A., Richaume, P., Delwart, S., Al Bitar, A., Pellarin, T., Bindlish, R., Jackson, T. J., Rüdiger, C., Waldteufel, P., Mecklenburg, S., and Wigneron, J. P.: Overview of SMOS performance in terms of global soil moisture monitoring after six years in operation, Remote Sens. Environ., 180, 40-63, https://doi.org/10.1016/j.rse.2016.02.042, 2016

Kerr, Y. H., Waldteufel, P., Wigneron, J., Martinuzzi, J., Font, J., and Berger, M.: Soil moisture retrieval from space: the Soil

Moisture and Ocean Salinity (SMOS) mission, IEEE Trans. Geosci. Remote Sensing, 39, 1729-1735, https://doi.org/10.1109/36.942551, 2001

Kim, H., Parinussa, R., Konings, A. G., Wagner, W., Cosh, M. H., Lakshmi, V., Zohaib, M., and Choi, M.: Global-scale assessment and combination of SMAP with ASCAT (active) and AMSR2 (passive) soil moisture products, Remote Sens. Environ., 204, 260-275, https://doi.org/10.1016/j.rse.2017.10.026, 2018

Kim, S., Liu, Y. Y., Johnson, F. M., Parinussa, R. M., and Sharma, A.: A global comparison of alternate AMSR2 soil moisture products: Why do they differ?, Remote Sens. Environ., 161, 43-62, https://doi.org/10.1016/j.rse.2015.02.002, 2015a

Kim, S., Parinussa, R. M., Liu, Y. Y., Johnson, F. M., and Sharma, A.: A framework for combining multiple soil moisture retrievals based on maximizing temporal correlation, Geophys. Res. Lett., 42, 6662-6670, http://doi.org/10.1002/2015GL064981, 2015b

Koike, T.: Soil moisture algorithm descriptions of GCOM-W1 AMSR2 (Rev. A), Earth Observation Research Center, Japan Aerospace Exploration Agency, 2013.

Koike, T., Nakamura, Y., Kaihotsu, I., Davaa, G., Matsuura, N., Tamagawa, K., and Fujii, H.: Development of an advanced microwave scanning radiometer (AMSR-E) algorithm for soil moisture and vegetation water content, Proceedings of Hydraulic Engineering, 48, 217-222, https://doi.org/10.2208/prohe.48.217, 2004

Kolassa, J., Reichle, R. H., Liu, Q., Alemohammad, S. H., Gentine, P., Aida, K., Asanuma, J., Bircher, S., Caldwell, T., Colliander, A., Cosh, M., Holifield Collins, C., Jackson, T. J., Martínez-Fernández, J., McNairn, H., Pacheco, A., Thibeault, M., and Walker, J. P.: Estimating surface soil moisture from SMAP observations using a Neural Network technique, Remote Sens. Environ., 204, 43-59, https://doi.org/10.1016/j.rse.2017.10.045, 2018

Konings, A. G., Piles, M., Das, N., and Entekhabi, D.: L-band vegetation optical depth and effective scattering albedo estimation from SMAP, Remote Sens. Environ., 198, 460-470, https://doi.org/10.1016/j.rse.2017.06.037, 2017

Kottek, M., Grieser, J., Beck, C., Rudolf, B., and Rubel, F.: World Map of the Köppen-Geiger climate classification updated, Meteorol. Z., 15, 259-263, https://doi.org/10.1127/0941-2948/2006/0130, 2006

Kozak, J. A., Ahuja, L. R., Green, T. R., and Ma, L.: Modelling crop canopy and residue rainfall interception effects on soil hydrological components for semi-arid agriculture, Hydrol. Process., 21, 229-241, https://doi.org/10.1002/hyp.6235, 2010

Kumar, S. V., Peters-Lidard, C. D., Santanello, J. A., Reichle, R. H., Draper, C. S., Koster, R. D., Nearing, G., and Jasinski, M. F.: Evaluating the utility of satellite soil moisture retrievals over irrigated areas and the ability of land data assimilation methods to correct for unmodeled processes, Hydrol. Earth Syst. Sci., 19, 4463-4478, https://doi.org/10.5194/hess-19-4463-2015, 2015

Lakhankar, T., Ghedira, H., Temimi, M., Azar, E. A., and Khanbilvardi, R.: Effect of Land Cover Heterogeneity on Soil Moisture Retrieval Using Active Microwave Remote Sensing Data, Remote Sens., 1, https://doi.org/10.3390/rs1020080, 2009

Lei, F., Crow, W. T., Shen, H., Su, C.-H., Holmes, T. R. H., Parinussa, R. M., and Wang, G.: Assessment of the impact of spatial heterogeneity on microwave satellite soil moisture periodic error, Remote Sens. Environ., 205, 85-99, https://doi.org/10.1016/j.rse.2017.11.002, 2018

Leroux, D. J., Kerr, Y. H., Bitar, A. A., Bindlish, R., Jackson, T. J., Berthelot, B., and Portet, G.: Comparison Between SMOS, VUA, ASCAT, and ECMWF Soil Moisture Products Over Four Watersheds in U.S, IEEE Trans. Geosci. Remote Sensing, 52, 1562-1571, https://doi.org/10.1109/TGRS.2013.2252468, 2014

Lievens, H., Martens, B., Verhoest, N. E. C., Hahn, S., Reichle, R. H., and Miralles, D. G.: Assimilation of global radar backscatter and radiometer brightness temperature observations to improve soil moisture and land evaporation estimates, Remote Sens. Environ., 189, 194-210, https://doi.org/10.1016/j.rse.2016.11.022, 2017

Liu, Y. Y., de Jeu, R. A. M., McCabe, M. F., Evans, J. P., and van Dijk, A. I. J. M.: Global long-term passive microwave satellite-based retrievals of vegetation optical depth, Geophys. Res. Lett., 38, https://doi.org/10.1029/2011GL048684, 2011a

Liu, Y. Y., Dorigo, W. A., Parinussa, R. M., de Jeu, R. A. M., Wagner, W., McCabe, M. F., Evans, J. P., and van Dijk, A. I. J. M.: Trend-preserving blending of passive and active microwave soil moisture retrievals, Remote Sens. Environ., 123, 280-297, https://doi.org/10.1016/j.rse.2012.03.014, 2012

Liu, Y. Y., Parinussa, R. M., Dorigo, W. A., De Jeu, R. A. M., Wagner, W., van Dijk, A. I. J. M., McCabe, M. F., and Evans, J. P.: Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals, Hydrol. Earth Syst. Sci., 15, 425-436, http://doi.org/10.5194/hess-15-425-2011, 2011b

Lu, Z., Chai, L., Ye, Q., and Zhang, T.: Reconstruction of time-series soil moisture from AMSR2 and SMOS data by using recurrent nonlinear autoregressive neural networks, 26-31 July 2015 2015, 980-983, https://doi.org/10.1109/IGARSS.2015.7325932.

Ma, H., Zeng, J., Chen, N., Zhang, X., Cosh, M. H., and Wang, W.: Satellite surface soil moisture from SMAP, SMOS, AMSR2 and ESA CCI: A comprehensive assessment using global ground-based observations, Remote Sens. Environ., 231, 111215, https://doi.org/10.1016/j.rse.2019.111215, 2019

Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A., and Verhoest, N. E. C.: GLEAM v3: satellite-based land evaporation and root-zone soil moisture, Geosci. Model Dev., 10, 1903-1925, https://doi.org/10.5194/gmd-10-1903-2017, 2017

Martínez-Fernández, J., González-Zamora, A., Sánchez, N., Gumuzzio, A., and Herrero-Jiménez, C. M.: Satellite soil moisture for agricultural drought monitoring: Assessment of the SMOS derived Soil Water Deficit Index, Remote Sens. Environ., 177, 277-286, https://doi.org/10.1016/j.rse.2016.02.064, 2016

McColl, K. A., Alemohammad, S. H., Akbar, R., Konings, A. G., Yueh, S., and Entekhabi, D.: The global distribution and dynamics of surface soil moisture, Nat. Geosci., 10, 100, https://doi.org/10.1038/ngeo2868, 2017

Méndez-Barroso, L. A., Vivoni, E. R., Watts, C. J., and Rodríguez, J. C.: Seasonal and interannual relations between precipitation, surface soil moisture and vegetation dynamics in the North American monsoon region, J. Hydrol., 377, 59-70, https://doi.org/10.1016/j.jhydrol.2009.08.009, 2009

Merriam, R. A.: A note on the interception loss equation, Int. J. Digit. Earth, 65, 3850-3851, https://doi.org/10.1029/jz065i011p03850, 1960

Miralles, D. G., Holmes, T. R. H., De Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A., and Dolman, A. J.: Global land-surface evaporation estimated from satellite-based observations, Hydrol. Earth Syst. Sci., 15, 453-469, https://doi.org/10.5194/hess-15-453-2011, 2011

Mladenova, I. E., Jackson, T. J., Njoku, E., Bindlish, R., Chan, S., Cosh, M. H., Holmes, T. R. H., de Jeu, R. A. M., Jones, L., Kimball, J., Paloscia, S., and Santi, E.: Remote monitoring of soil moisture using passive microwave-based techniques —

Theoretical basis and overview of selected algorithms for AMSR-E, Remote Sens. Environ., 144, 197-213, https://doi.org/10.1016/j.rse.2014.01.013, 2014

Naithani, K. J., Baldwin, D. C., Gaines, K. P., Lin, H., and Eissenstat, D. M.: Spatial Distribution of Tree Species Governs the Spatio-Temporal Interaction of Leaf Area Index and Soil Moisture across a Forested Landscape, PLOS ONE, 8, e58704, https://doi.org/10.1371/journal.pone.0058704, 2013

Neill, P. E. O., Podest, E., and Njoku, E. G.: Utilization of ancillary data sets for SMAP algorithm development and product generation, 24-29 July 2011 2011, 2436-2439, https://doi.org/10.1109/IGARSS.2011.6049703.

Njoku, E. G., Ashcroft, P., Chan, T. K., and Li, L.: Global survey and statistics of radio-frequency interference in AMSR-E land observations, IEEE Trans. Geosci. Remote Sensing, 43, 938-947, https://doi.org/10.1109/TGRS.2004.837507, 2005

Njoku, E. G. and Chan, T. K.: Vegetation and surface roughness effects on AMSR-E land observations, Remote Sens. Environ., 100, 190-199, https://doi.org/10.1016/j.rse.2005.10.017, 2006

Njoku, E. G., Jackson, T. J., Lakshmi, V., Chan, T. K., and Nghiem, S. V.: Soil moisture retrieval from AMSR-E, IEEE Trans. Geosci. Remote Sensing, 41, 215-229, https://doi.org/10.1109/TGRS.2002.808243, 2003

Oliva, R., Daganzo, E., Kerr, Y. H., Mecklenburg, S., Nieto, S., Richaume, P., and Gruhier, C.: SMOS Radio Frequency Interference Scenario: Status and Actions Taken to Improve the RFI Environment in the 1400–1427-MHz Passive Band, IEEE Trans. Geosci. Remote Sensing, 50, 1427-1439, https://doi.org/10.1109/TGRS.2012.2182775, 2012

Owe, M., de Jeu, R., and Holmes, T.: Multisensor historical climatology of satellite-derived global land surface moisture, J. Geophys. Res.-Earth Surf., 113, https://doi.org/10.1029/2007JF000769, 2008

Owe, M., Jeu, R. d., and Walker, J.: A methodology for surface soil moisture and vegetation optical depth retrieval using the microwave polarization difference index, IEEE Trans. Geosci. Remote Sensing, 39, 1643-1654, https://doi.org/10.1109/36.942542, 2001

Panciera, R., Walker, J. P., Kalma, J. D., Kim, E. J., Saleh, K., and Wigneron, J.-P.: Evaluation of the SMOS L-MEB passive microwave soil moisture retrieval algorithm, Remote Sens. Environ., 113, 435-444, https://doi.org/10.1016/j.rse.2008.10.010, 2009

Parinussa, R. M., Holmes, T. R. H., and Jeu, R. A. M. d.: Soil Moisture Retrievals From the WindSat Spaceborne Polarimetric Microwave Radiometer, IEEE Trans. Geosci. Remote Sensing, 50, 2683-2694, https://doi.org/10.1109/TGRS.2011.2174643, 2012

Parinussa, R. M., Holmes, T. R. H., Wanders, N., Dorigo, W. A., and de Jeu, R. A. M.: A Preliminary Study toward Consistent Soil Moisture from AMSR2, J. Hydrometeorol., 16, 932-947, https://doi.org/10.1175/JHM-D-13-0200.1, 2014

Parinussa, R. M., Holmes, T. R. H., Yilmaz, M. T., and Crow, W. T.: The impact of land surface temperature on soil moisture anomaly detection from passive microwave observations, Hydrol. Earth Syst. Sci., 15, 3135-3151, https://doi.org/10.5194/hess-15-3135-2011, 2011

Piles, M., Schalie, R. v. d., Gruber, A., Muñoz-Marí, J., Camps-Valls, G., Mateo-Sanchis, A., Dorigo, W., and Jeu, R. d.: Global Estimation of Soil Moisture Persistence with L and C-Band Microwave Sensors, 22-27 July 2018 2018, 8259-8262, https://doi.org/10.1109/IGARSS.2018.8518161.

Qiu, J., Gao, Q., Wang, S., and Su, Z.: Comparison of temporal trends from multiple soil moisture data sets and precipitation:

The implication of irrigation on regional soil moisture trend, Int. J. Appl. Earth Obs. Geoinf., 48, 17-27, https://doi.org/10.1016/j.jag.2015.11.012, 2016

Qu, Y., Zhu, Z., Chai, L., Liu, S., Montzka, C., Liu, J., Yang, X., Lu, Z., Jin, R., Li, X., Guo, Z., and Zheng, J.: Rebuilding a Microwave Soil Moisture Product Using Random Forest Adopting AMSR-E/AMSR2 Brightness Temperature and SMAP over the Qinghai–Tibet Plateau, China, Remote Sens., 11, https://doi.org/10.3390/rs11060683, 2019

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C. J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D., and Toll, D.: The Global Land Data Assimilation System, Bull. Amer. Meteorol. Soc., 85, 381-394, https://doi.org/10.1175/BAMS-85-3-381, 2004

Rodríguez-Fernández, J. N., Kerr, H. Y., Van der Schalie, R., Al-Yaari, A., Wigneron, J.-P., De Jeu, R., Richaume, P., Dutra, E., Mialon, A., and Drusch, M.: Long Term Global Surface Soil Moisture Fields Using an SMOS-Trained Neural Network Applied to AMSR-E Data, Remote Sens., 8, https://doi.org/10.3390/rs8110959, 2016

Rodriguez-Fernandez, N., Aires, F., Richaume, P., Kerr, Y. H., Prigent, C., Kolassa, J., Cabot, F., Jiménez, C., Mahmoodi, A., and Drusch, M.: Soil Moisture Retrieval Using Neural Networks: Application to SMOS, IEEE Trans. Geosci. Remote Sensing, 53, 5991-6007, https://doi.org/10.1109/TGRS.2015.2430845, 2015

Samaniego, L., Thober, S., Kumar, R., Wanders, N., Rakovec, O., Pan, M., Zink, M., Sheffield, J., Wood, E. F., and Marx, A.: Anthropogenic warming exacerbates European soil moisture droughts, Nat. Clim. Chang., 8, 421-426, https://doi.org/10.1038/s41558-018-0138-5, 2018

Schroeder, R., McDonald, C. K., Chapman, D. B., Jensen, K., Podest, E., Tessler, D. Z., Bohn, J. T., and Zimmermann, R.: Development and Evaluation of a Multi-Year Fractional Surface Water Data Set Derived from Active/Passive Microwave Remote Sensing Data, Remote Sens., 7, https://doi.org/10.3390/rs71215843, 2015

Shi, J., Jackson, T., Tao, J., Du, J., Bindlish, R., Lu, L., and Chen, K. S.: Microwave vegetation indices for short vegetation covers from satellite passive microwave sensor AMSR-E, Remote Sens. Environ., 112, 4285-4300, https://doi.org/10.1016/j.rse.2008.07.015, 2008

Stillman, S. and Zeng, X.: Evaluation of SMAP Soil Moisture Relative to Five Other Satellite Products Using the Climate Reference Network Measurements Over USA, IEEE Trans. Geosci. Remote Sensing, 56, 6296-6305, https://doi.org/10.1109/TGRS.2018.2835316, 2018

Stinchcombe and White: Universal approximation using feedforward networks with non-sigmoid hidden layer activation functions, 1989 1989, 613-617 vol.611, http://doi.org/10.1109/IJCNN.1989.118640.

Stocker, B. D., Zscheischler, J., Keenan, T. F., Prentice, I. C., Seneviratne, S. I., and Peñuelas, J.: Drought impacts on terrestrial primary production underestimated by satellite monitoring, Nat. Geosci., 12, 264-270, https://doi.org/10.1038/s41561-019-0318-6, 2019

Ulaby, F. T., Batlivala, P. P., and Dobson, M. C.: Microwave Backscatter Dependence on Surface Roughness, Soil Moisture, and Soil Texture: Part I-Bare Soil, IEEE Trans. Geosci. Electronics, 16, 286-295, https://doi.org/10.1109/TGE.1978.294586, 1978

Van der Schalie, R., De Jeu, R., Parinussa, R., Rodríguez-Fernández, N., Kerr, Y., Al-Yaari, A., Wigneron, J.-P., and Drusch, M.: The Effect of Three Different Data Fusion Approaches on the Quality of Soil Moisture Retrievals from Multiple Passive

Microwave Sensors, Remote Sens., 10, https://doi.org/10.3390/rs10010107, 2018

Van der Schalie, R., de Jeu, R. A. M., Kerr, Y. H., Wigneron, J. P., Rodríguez-Fernández, N. J., Al-Yaari, A., Parinussa, R. M., Mecklenburg, S., and Drusch, M.: The merging of radiative transfer based surface soil moisture data from SMOS and AMSR-E, Remote Sens. Environ., 189, 180-193, https://doi.org/10.1016/j.rse.2016.11.026, 2017

Verger, A., Baret, F., and Weiss, M.: Near Real-Time Vegetation Monitoring at Global Scale, IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens., 7, 3473-3481, https://doi.org/10.1109/JSTARS.2014.2328632, 2014

Wagner, W., Lemoine, G., and Rott, H.: A Method for Estimating Soil Moisture from ERS Scatterometer and Soil Data, Remote Sens. Environ., 70, 191-207, https://doi.org/10.1016/S0034-4257(99)00036-X, 1999

Wigneron, J., Calvet, J., Rosnay, P. d., Kerr, Y., Waldteufel, P., Saleh, K., Escorihuela, M. J., and Kruszewski, A.: Soil moisture retrievals from biangular L-band passive microwave observations, IEEE Geosci. Remote Sens. Lett., 1, 277-281, https://doi.org/10.1109/LGRS.2004.834594, 2004

Wigneron, J. P., Kerr, Y., Waldteufel, P., Saleh, K., Escorihuela, M. J., Richaume, P., Ferrazzoli, P., de Rosnay, P., Gurney, R., Calvet, J. C., Grant, J. P., Guglielmetti, M., Hornbuckle, B., Mätzler, C., Pellarin, T., and Schwank, M.: L-band Microwave Emission of the Biosphere (L-MEB) Model: Description and calibration against experimental data sets over crop fields, Remote Sens. Environ., 107, 639-655, https://doi.org/10.1016/j.rse.2006.10.014, 2007

Xiao, Z., Liang, S., Wang, J., Chen, P., Yin, X., Zhang, L., and Song, J.: Use of General Regression Neural Networks for Generating the GLASS Leaf Area Index Product From Time-Series MODIS Surface Reflectance, IEEE Trans. Geosci. Remote Sensing, 52, 209-223, https://doi.org/10.1109/TGRS.2013.2237780, 2014

Xiao, Z., Liang, S., Wang, J., Xiang, Y., Zhao, X., and Song, J.: Long-Time-Series Global Land Surface Satellite Leaf Area Index Product Derived From MODIS and AVHRR Surface Reflectance, IEEE Trans. Geosci. Remote Sensing, 54, 5301-5318, https://doi.org/10.1109/TGRS.2016.2560522, 2016

Yang, H., Weng, F., Lv, L., Lu, N., Liu, G., Bai, M., Qian, Q., He, J., and Xu, H.: The FengYun-3 Microwave Radiation Imager On-Orbit Verification, IEEE Trans. Geosci. Remote Sensing, 49, 4552-4560, https://doi.org/10.1109/TGRS.2011.2148200, 2011

Yang, J., Zhang, P., Lu, N., Yang, Z., Shi, J., and Dong, C.: Improvements on global meteorological observations from the current Fengyun 3 satellites and beyond, Int. J. Digit. Earth, 5, 251-265, https://doi.org/10.1080/17538947.2012.658666, 2012

Yao, P., Lu, H., Yue, S., Yang, F., Lyu, H., Yang, K., McColl, K. A., Gianotti, D., and ENTekhabi, D.: Estimating Surface Soil Moisture from AMSR2 Tb with Artificial Neural Network Method and SMAP Products, IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 6998-7001, https://doi.org/10.1109/IGARSS.2019.8898152, 2019.

Yao, P., Shi, J., Zhao, T., Lu, H., and Al-Yaari, A.: Rebuilding Long Time Series Global Soil Moisture Products Using the Neural Network Adopting the Microwave Vegetation Index, Remote Sens., 9, https://doi.org/10.3390/rs9010035, 2017

Ye, N., Walker, J. P., Guerschman, J., Ryu, D., and Gurney, R. J.: Standing water effect on soil moisture retrieval from L-band passive microwave observations, Remote Sens. Environ., 169, 232-242, https://doi.org/10.1016/j.rse.2015.08.013, 2015

Ye, N., Walker, J. P., Yeo, I., Jackson, T. J., Kerr, Y., Kim, E., McGrath, A., PopStefanija, I., Goodberlet, M., and Hills, J.: Toward P-Band Passive Microwave Sensing of Soil Moisture, IEEE Geosci. Remote Sens. Lett., 169, 232-242,

删除了: doi: https://doi.org/10.1109/LGRS.2020.2976204, 2020. 1-5…

设置了格式: 字体: (默认) Times New Roman

设置了格式: 字体: (默认) Times New Roman

设置了格式: 字体: (默认) Times New Roman

设置了格式: 字体: (默认) Times New Roman

https://doi.org/10.1109/LGRS.2020.2976204, 2020

Yilmaz, M. T., Hunt, E. R., and Jackson, T. J.: Remote sensing of vegetation water content from equivalent water thickness using satellite imagery, Remote Sens. Environ., 112, 2514-2522, https://doi.org/10.1016/j.rse.2007.11.014, 2008

Zhang, R., Kim, S., and Sharma, A.: A comprehensive validation of the SMAP Enhanced Level-3 Soil Moisture product using ground measurements over varied climates and landscapes, Remote Sens. Environ., 223, 82-94, https://doi.org/10.1016/j.rse.2019.01.015, 2019

Zhao, Y., Peth, S., Wang, X. Y., Lin, H., and Horn, R.: Controls of surface soil moisture spatial patterns and their temporal stability in a semi-arid steppe, Hydrol. Process., 24, 2507-2519, https://doi.org/10.1002/hyp.7665, 2010

Tables

Table 1: Description of the Köppen-Geiger climate classification types at all the selected ISMN stations.

Climate_Köppen	General description
Aw	Equatorial savannah with dry winter
BSk	Steppe climate, cold and arid
BWh	Desert climate, hot and arid
BWk	Desert climate, cold and arid
Cfa	Warm temperate climate, fully humid, hot summer
Cfb	Warm temperate climate, fully humid, warm summer
Csa	Warm temperate climate with dry, hot summer
Csb	Warm temperate climate with dry, warm summer
Dfa	Snow climate, fully humid, hot summer
Dfb	Snow climate, fully humid, warm summer
Dfc	Snow climate, fully humid, cool summer and cold winter
Dsb	Snow climate with dry, warm summer
Dwc	Snow climate with cool summer and cold, dry winter
ET	Tundra climate

删除了: Table 1: The global surface soil moisture products used in this study.
Product name ...

带格式的: 段落间距段后: 12 磅

设置了格式: 非上标/ 下标

删除了: 2

设置了格式: 非上标/ 下标

设置了格式: 字体: (中文) +中文标题 (等线 Light)

删除了: The d

带格式的: 行距: 单倍行距

带格式的: 行距: 单倍行距

带格式的: 行距: 单倍行距

带格式的: 行距: 单倍行距

带格式的: 行距: 单倍行距

带格式的: 行距: 单倍行距

带格式的: 行距: 单倍行距

带格式的: 行距: 单倍行距

带格式的: 行距: 单倍行距

带格式的: 行距: 单倍行距

带格式的: 行距: 单倍行距

带格式的: 行距: 单倍行距

带格式的: 行距: 单倍行距

带格式的: 行距: 单倍行距

带格式的: 行距: 单倍行距

Table 2: The mean and median values of the five evaluation indexes (correlation coefficient: *r*, RMSE, bias, unbiased RMSE (ubRMSE), and the anomalies r (A.R)) on the temporal accuracy of the surface soil moisture simulated in this study (RSSSM) and the other surface soil moisture products, when validated using the ISMN in-situ measurements. Note: 1) for the comparison of RSSSM against the SMAP_E (SMAP) product, the validation period is from April 2015 to 2018; 2) for the comparison between RSSSM and ASCAT-SWI (ASCAT), the period is 2007~2018; 3) the comparison period for RSSSM and GLDAS Noah v2.1 (GLDAS), ERA5-Land (ERA5-L), CCI or GLEAM v3.3a (GLE-a) surface soil moisture product are 2003~2018; 4) the common period for RSSSM and GLEAM v3.3b (GLE-b) is from 2003 to September 2018.

Index	*r*		RMSE		bias		ubRMSE		A.R	
Product	RSSSM	SMAP	RSSSM	SMAP	RSSSM	SMAP	RSSSM	SMAP	RSSSM	SMAP
Mean	0.756	0.762	0.075	0.074	0.015	0.016	0.043	0.043	0.700	0.707
Median	0.795	0.798	0.067	0.066	0.009	0.013	0.043	0.043	0.720	0.744
Product	RSSSM	ASCAT	RSSSM	ASCAT	RSSSM	ASCAT	RSSSM	ASCAT	RSSSM	ASCAT
Mean	0.687	0.561	0.079	0.095	0.002	-0.007	0.047	0.062	0.627	0.554
Median	0.735	0.627	0.074	0.088	-0.001	-0.010	0.048	0.062	0.654	0.595
Product	RSSSM	GLDAS	RSSSM	GLDAS	RSSSM	GLDAS	RSSSM	GLDAS	RSSSM	GLDAS
Mean	0.689	0.613	0.080	0.091	0.001	0.028	0.047	0.051	0.620	0.519
Median	0.737	0.661	0.075	0.082	-0.002	0.029	0.048	0.049	0.661	0.567
Product	RSSSM	ERA5-L	RSSSM	ERA5-L	RSSSM	ERA5-L	RSSSM	ERA5-L	RSSSM	ERA5-L
Mean	0.689	0.734	0.080	0.112	0.001	0.082	0.047	0.050	0.620	0.648
Median	0.737	0.758	0.075	0.094	-0.002	0.073	0.048	0.049	0.661	0.672
Product	RSSSM	CCI	RSSSM	CCI	RSSSM	CCI	RSSSM	CCI	RSSSM	CCI
Mean	0.690	0.642	0.080	0.091	0.002	-0.002	0.047	0.049	0.620	0.530
Median	0.735	0.666	0.074	0.080	-0.002	0.006	0.049	0.047	0.658	0.552
Product	RSSSM	GLE-a	RSSSM	GLE-a	RSSSM	GLE-a	RSSSM	GLE-a	RSSSM	GLE-a
Mean	0.689	0.735	0.080	0.126	0.001	0.093	0.047	0.047	0.620	0.681
Median	0.737	0.771	0.075	0.119	-0.002	0.104	0.048	0.046	0.661	0.715
Product	RSSSM	GLE-b	RSSSM	GLE-b	RSSSM	GLE-b	RSSSM	GLE-b	RSSSM	GLE-b
Mean	0.688	0.729	0.080	0.117	0.001	0.077	0.047	0.046	0.618	0.670
Median	0.730	0.762	0.075	0.112	-0.002	0.091	0.048	0.045	0.659	0.705

删除了: 3
设置了格式: 字体: (中文) +中文标题 (等线 Light)
删除了: values
删除了: R
删除了: with the unit of m3 m-3
删除了: R
删除了:
删除了: simulated
设置了格式: 非上标/ 下标
设置了格式: 非上标/ 下标
删除了: SIM
删除了: SMAP during April 2015~2018,
删除了: compared with
删除了: site
删除了: or
删除了: or
删除了: of
带格式表格
带格式表格
删除了:
Index
带格式表格
设置了格式: 字体: 倾斜

Table 3: The mean and median values of the four evaluation indexes (*r*, RMSE, bias and ubRMSE) on the spatial pattern accuracy of RSSSM and the other global long-term surface soil moisture products (SMAP_E, ASCAT-SWI, GLDAS Noah v2.1, ERA5-Land, CCI, GLEAM v3.3a and GLEAM v3.3b) in every 10-day period. For each pair of comparisons, the evaluation indexes are for the common period of the two products, which is the same as Table 2. The abbreviations for the products are also the same as those in Table 2.

Index	*r*		RMSE		bias		ubRMSE	
Product	RSSSM	SMAP	RSSSM	SMAP	RSSSM	SMAP	RSSSM	SMAP
Mean	0.652	0.659	0.084	0.084	0.016	0.016	0.082	0.081
Median	0.655	0.664	0.082	0.081	0.019	0.019	0.080	0.078
Product	RSSSM	ASCAT	RSSSM	ASCAT	RSSSM	ASCAT	RSSSM	ASCAT
Mean	0.636	0.561	0.087	0.102	0.005	-0.010	0.085	0.097
Median	0.650	0.572	0.086	0.100	0.007	-0.009	0.085	0.095
Product	RSSSM	GLDAS	RSSSM	GLDAS	RSSSM	GLDAS	RSSSM	GLDAS
Mean	0.617	0.593	0.090	0.097	-0.005	0.035	0.086	0.087
Median	0.643	0.630	0.089	0.096	0.001	0.041	0.086	0.086
Product	RSSSM	ERA5-L	RSSSM	ERA5-L	RSSSM	ERA5-L	RSSSM	ERA5-L
Mean	0.616	0.575	0.090	0.125	-0.005	0.077	0.086	0.095
Median	0.641	0.633	0.089	0.125	0.001	0.082	0.086	0.092
Product	RSSSM	CCI	RSSSM	CCI	RSSSM	CCI	RSSSM	CCI
Mean	0.618	0.497	0.090	0.099	-0.004	0.003	0.086	0.093
Median	0.647	0.554	0.089	0.098	0.002	0.006	0.086	0.093
Product	RSSSM	GLE-a	RSSSM	GLE-a	RSSSM	GLE-a	RSSSM	GLE-a
Mean	0.617	0.576	0.090	0.139	-0.005	0.105	0.086	0.089
Median	0.643	0.616	0.089	0.142	0.001	0.112	0.086	0.088
Product	RSSSM	GLE-b	RSSSM	GLE-b	RSSSM	GLE-b	RSSSM	GLE-b
Mean	0.616	0.560	0.090	0.128	-0.005	0.088	0.086	0.090
Median	0.643	0.613	0.089	0.130	0.001	0.094	0.086	0.089

删除了: 4

删除了: five

设置了格式: 字体: (中文) +中文标题 (等线 Light)

删除了: statistics

删除了: R

删除了: ,

删除了: and A.R

删除了: SIM

删除了: SMAP

删除了: in

删除了: each

删除了: during April 2015~2018

删除了: of

带格式表格

Figures

删除了: Index

设置了格式: 字体: 倾斜

设置了格式: 字体: 倾斜

带格式的: 段落间距段前: 0 磅, 段后: 0 磅

设置了格式: 字体: 倾斜

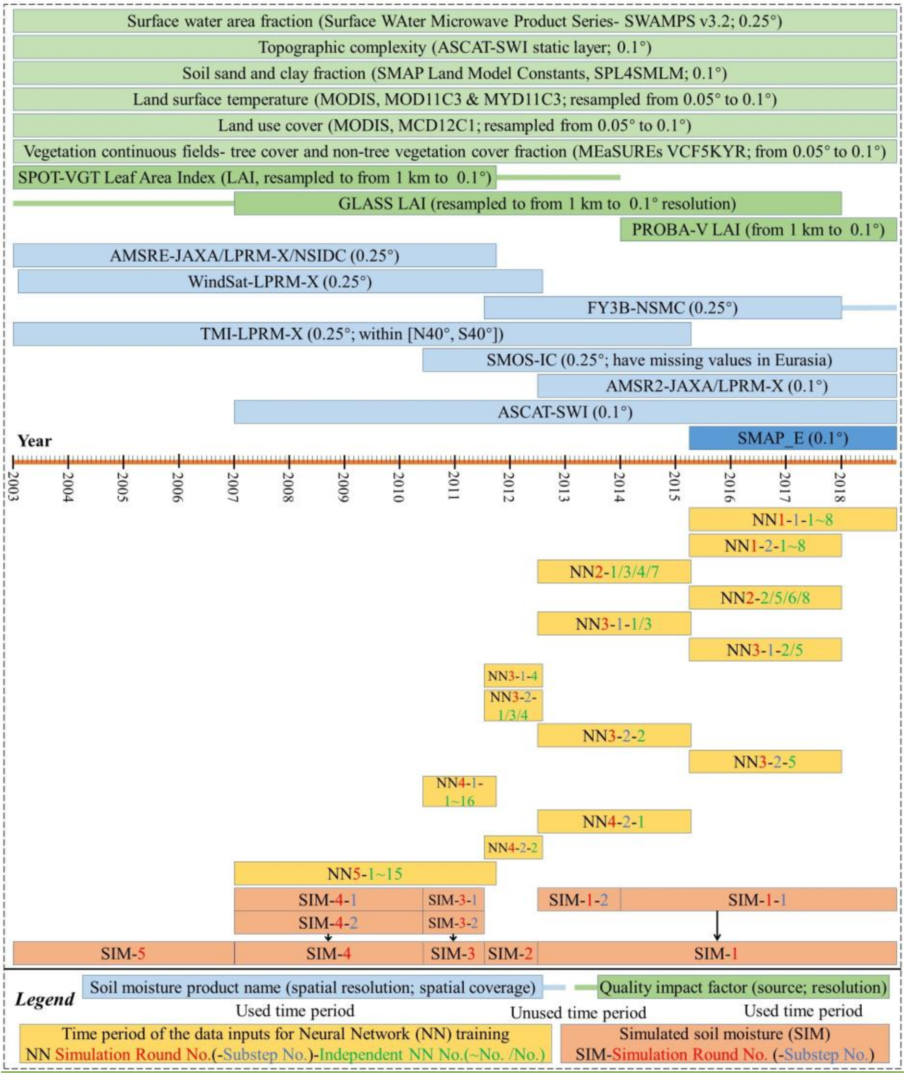
带格式表格

带格式表格

设置了格式: 字体: 倾斜

设置了格式: 字体: 倾斜

设置了格式: 字体: 倾斜



Surface water area fraction (Surface WAter Microwave Product Series- SWAMPS v3.2; 0.25°)
Topographic complexity (ASCAT-SWI static layer; 0.1°)
Soil sand and clay fraction (SMAP Land Model Constants, SPL4SMLM; 0.1°)
Land surface temperature (MODIS, MOD11C3 & MYD11C3; resampled from 0.05° to 0.1°)
Land use cover (MODIS, MCD12C1; resampled from 0.05° to 0.1°)
Vegetation continuous fields- tree cover and non-tree vegetation cover fraction (MEaSUREs VCF5KYR; from 0.05° to 0.1°)
SPOT-VGT Leaf Area Index (LAI, resampled to from 1 km to 0.1°)
GLASS LAI (resampled to from 1 km to 0.1° resolution)
PROBA-V LAI (from 1 km to 0.1°)
AMSRE-JAXA/LPRM-X/NSIDC (0.25°)
WindSat-LPRM-X (0.25°)
FY3B-NSMC (0.25°)
TMI-LPRM-X (0.25°; within [N40°, S40°])
SMOS-IC (0.25°; have missing values in Eurasia)
AMSR2-JAXA/LPRM-X (0.1°)
ASCAT-SWI (0.1°)
Year
SMAP_E (0.1°)
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
NN1-1-1~8
NN1-2-1~8
NN2-1/3/4/7
NN2-2/5/6/8
NN3-1-1/3
NN3-1-2/5
NN3-1-4
NN3-2-1/3/4
NN3-2-2
NN3-2-5
NN4-1-1~16
NN4-2-1
NN4-2-2
NN5-1~15
SIM-4-1
SIM-3-1
SIM-1-2
SIM-1-1
SIM-4-2
SIM-3-2
SIM-5
SIM-4
SIM-3
SIM-2
SIM-1
Legend
Soil moisture product name (spatial resolution; spatial coverage)
Quality impact factor (source; resolution)
Used time period
Unused time period
Used time period
Time period of the data inputs for Neural Network (NN) training
NN Simulation Round No.(-Substep No.)-Independent NN No.(~No. /No.)
Simulated soil moisture (SIM)
SIM-Simulation Round No. (-Substep No.)

Figure 1: The timeline figure showing the time periods of the soil moisture datasets and the 'quality impact factor' products (e.g., LAI dataset) used in this study (listed above the timeline), as well as the periods of data applied for the training of the 67 independent neural networks and the neural network simulation outputs (i.e., simulated soil moisture) of eight substeps (listed below the timeline).

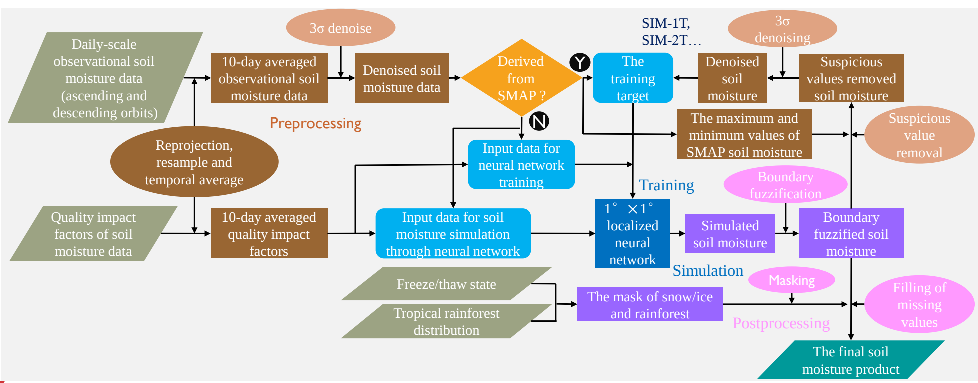
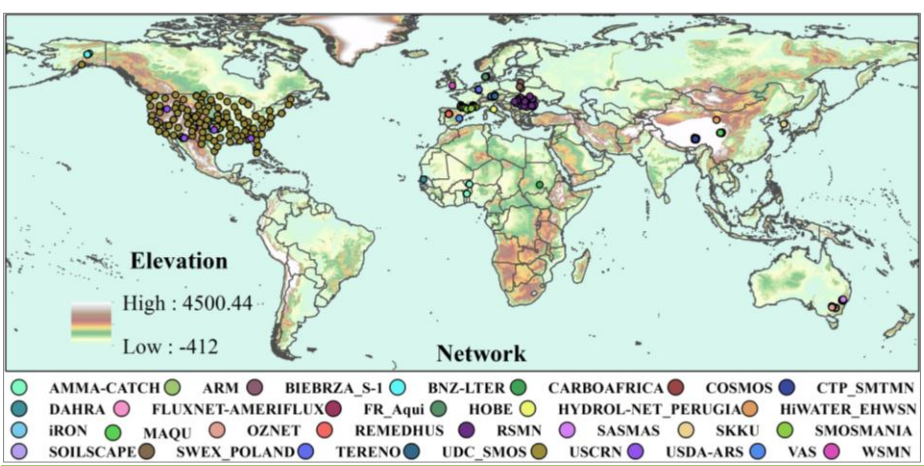
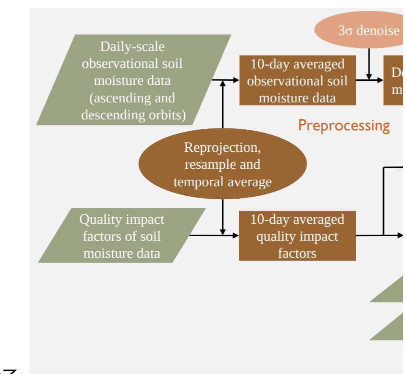


Figure 2: The flow chart for the production of global surface soil moisture data (RSSSM).



删除了:



带格式的: 居中

Figure 3. The global distribution of ISMN networks and stations.

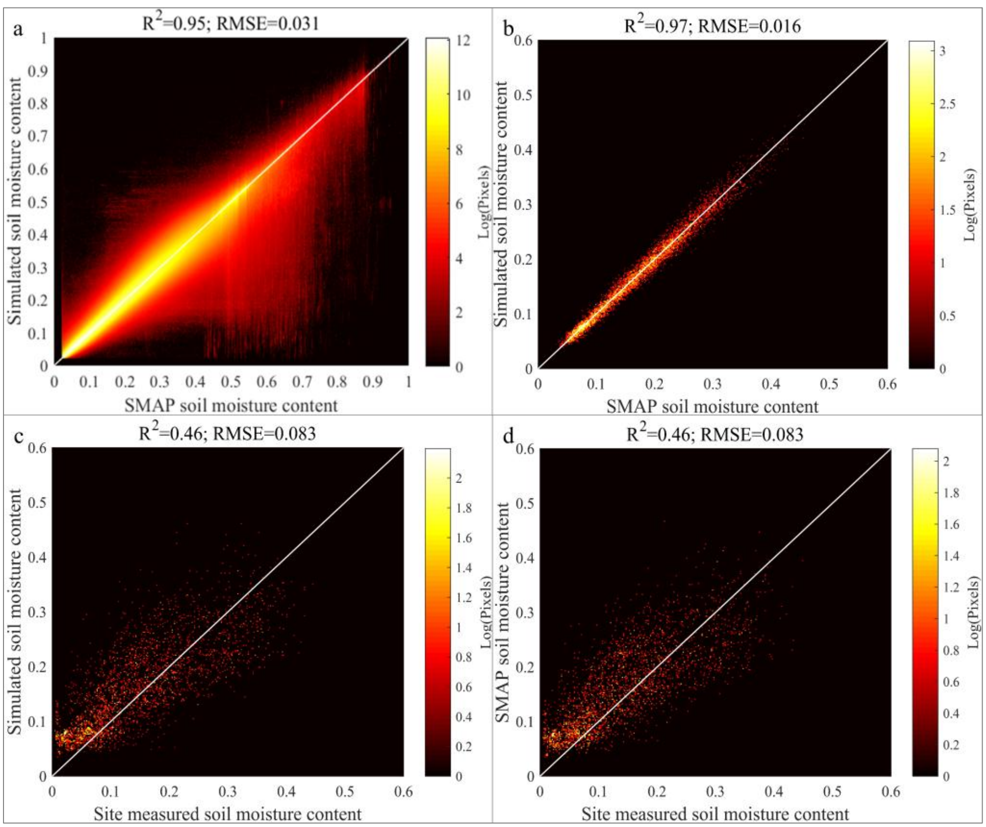


Figure 4: Comparison between the neural network simulated surface soil moisture (RSSSM) and SMAP data. The scatter plots are between: (a) RSSSM and SMAP values at all pixels; b) RSSSM and SMAP values at only the pixels with measurements; (c) RSSSM and the site measured soil moisture; and (d) SMAP and the site measurements during April 2015~2018. All plots are represented as the point density on a logarithmic scale, while the units for soil moisture content and RMSE values are m^3 m^{-3}.

删除了: 2

删除了: SIM

删除了: SIM

删除了: grids

删除了: SIM

删除了: grids

删除了: SIM

删除了: of

删除了: is

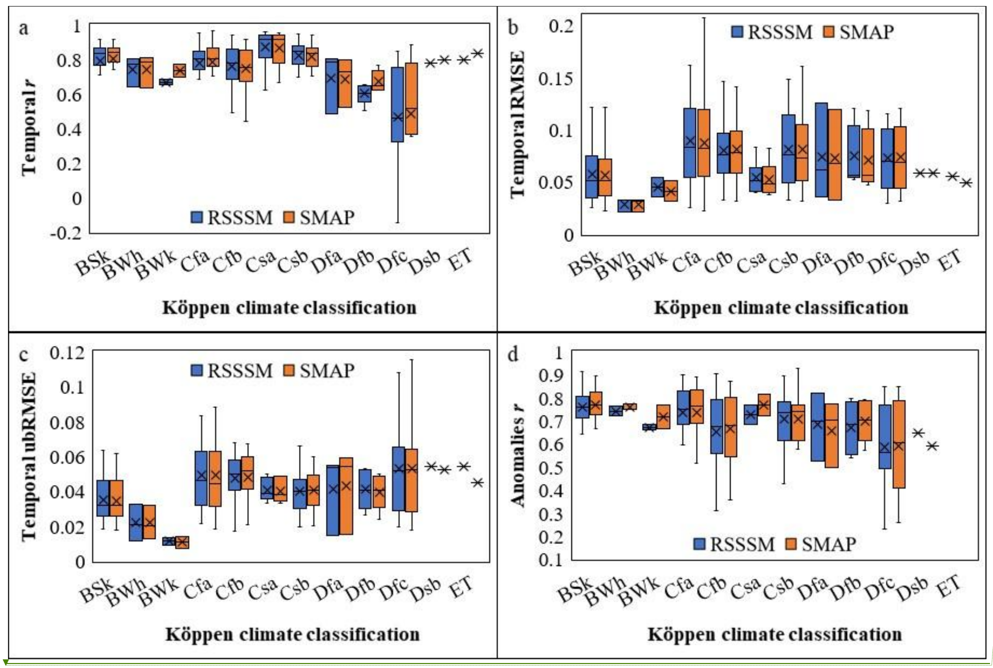
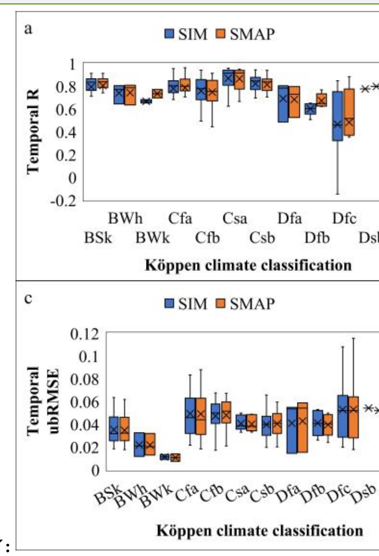


Figure 5: Comparison between the temporal accuracy of RSSSM and SMAP in regions with different Köppen-Geiger climate types. The four indexes are (a) *r*, (b) RMSE, (c) ubRMSE and (d) Anomalies *r* (A.R). The lengths of the error bars are 1.5 times that of the interquartile range, while the upper and lower boundaries and the central lines of the boxes indicate the 75th, 50th and 25th percentile values, with mean values marked by '×' (the forms of all the following boxplots are the same).



删除了:

删除了: 3

删除了: SIM

设置了格式: 字体: 倾斜

删除了: R

删除了: ***R***

设置了格式: 字体: 倾斜

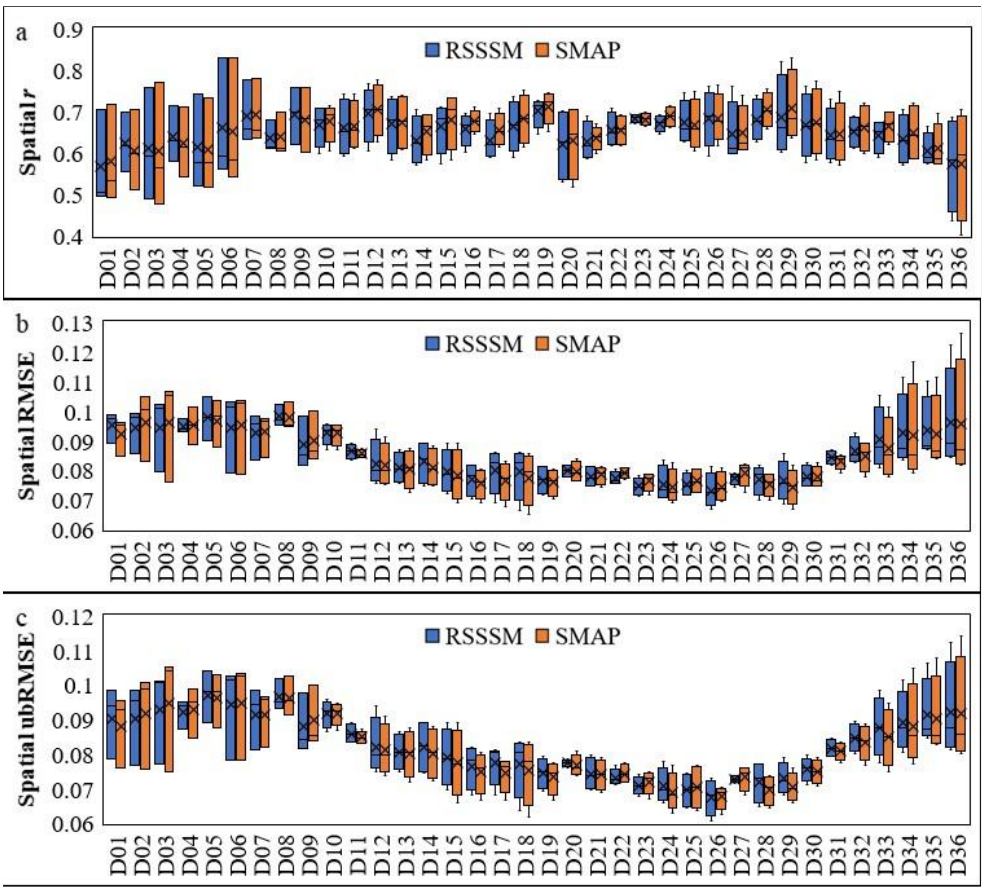


Figure 6: Comparison between the spatial pattern accuracy of RSSSM and SMAP in different 10-day periods during April 2015~2018. The three evaluation indexes are (a) *r*, (b) RMSE and (c) ubRMSE. The length of each box/error bar is determined from the evaluation index values in three (January to March) or four (April to December) years.

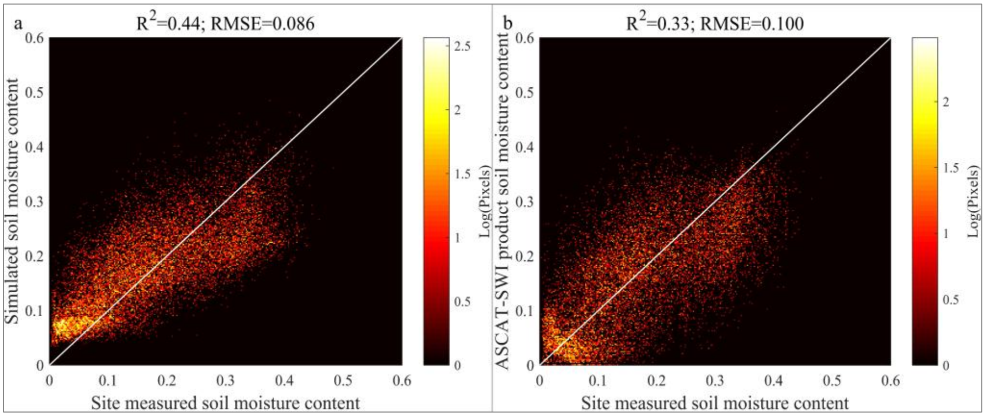


Figure 7: The overall data accuracy comparison between RSSSM and the ASCAT-SWI data product. The scatter plot is between (a) RSSSM or (b) ASCAT-SWI soil moisture and the site measured values during 2007~2018. The unit of all plots is the density of points on a logarithmic scale.

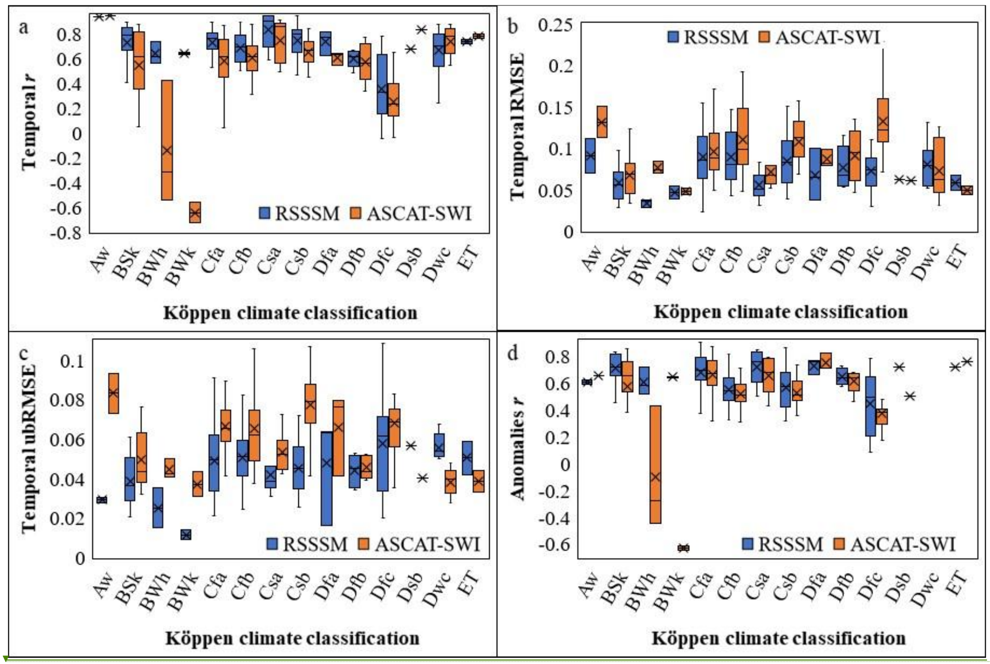
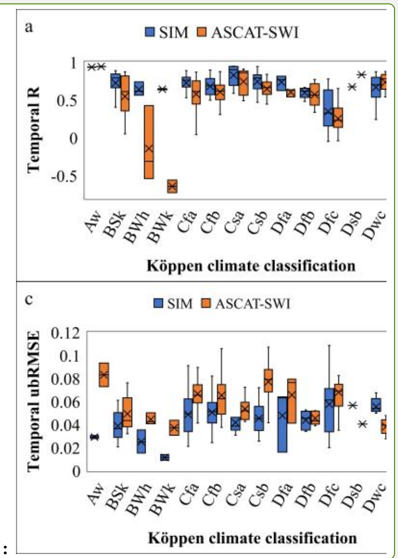


Figure 8: Comparison between the temporal accuracy of RSSSM and ASCAT-SWI in different Köppen-Geiger climatic regions. The four indexes are (a) *r*, (b) RMSE, (c) ubRMSE, and (d) Anomalies R (A.R).

删除了:



删除了: :

删除了: *R*

设置了格式: 字体: 倾斜

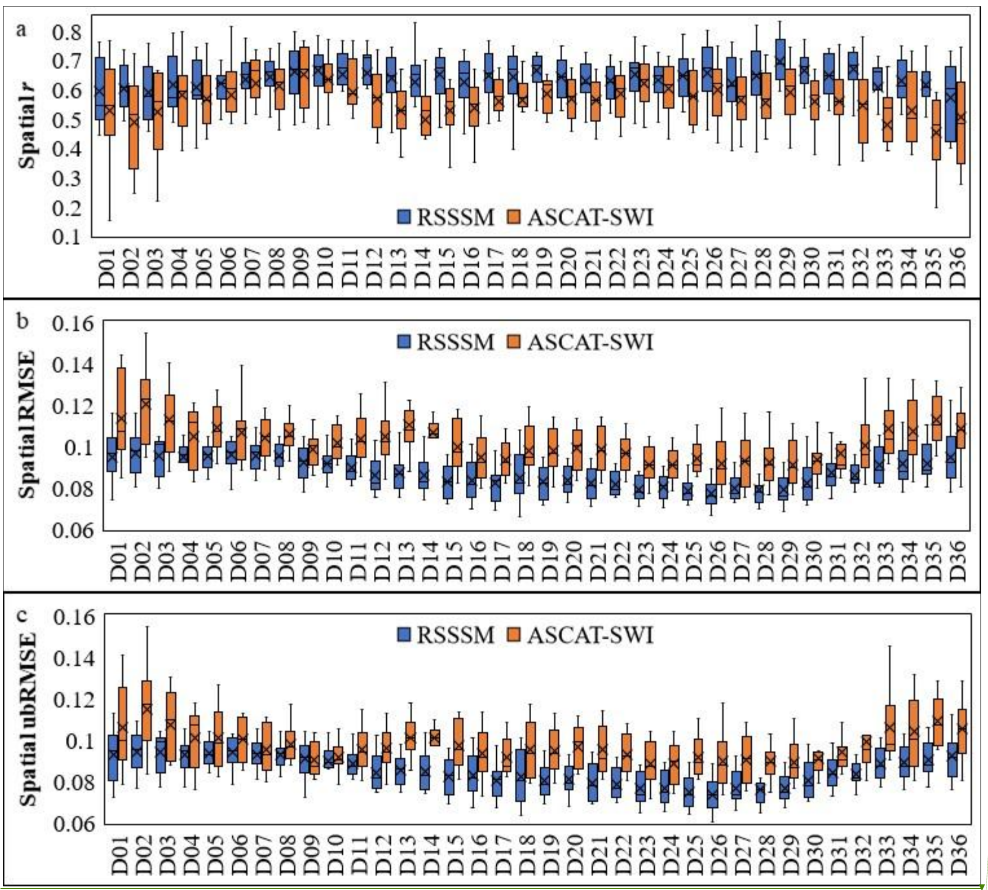
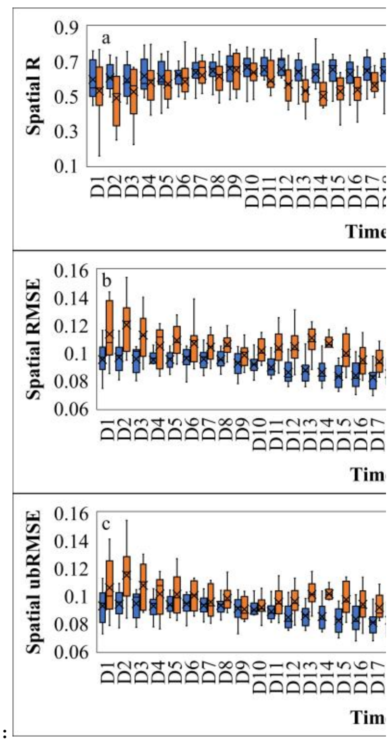


Figure 9: Comparison between the spatial accuracy of RSSSM and ASCAT-SWI during different 10-day periods. The evaluation indexes are (a) *r*, (b) RMSE, and (c) ubRMSE.

删除了:



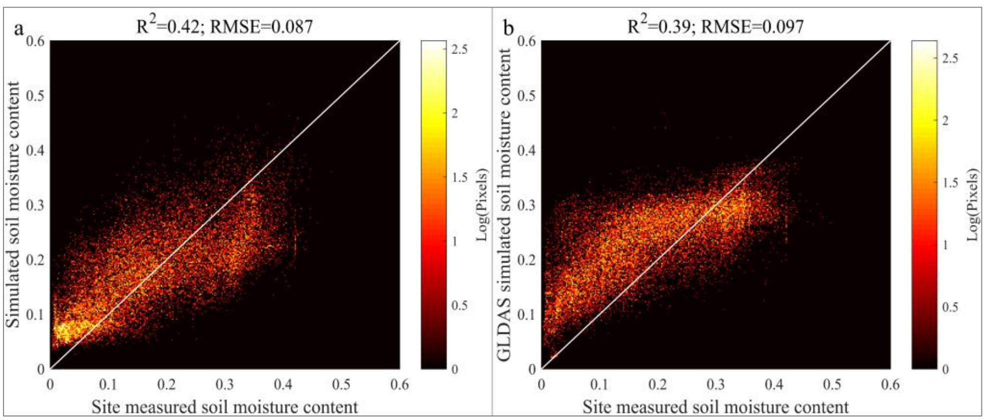


Figure 10: The overall data accuracy comparison between RSSSM and the surface soil moisture simulated by GLDAS Noah V2.1. The scatter plot is between the (a) RSSSM or (b) GLDAS soil moisture and the measured soil moisture during 2003~2018.

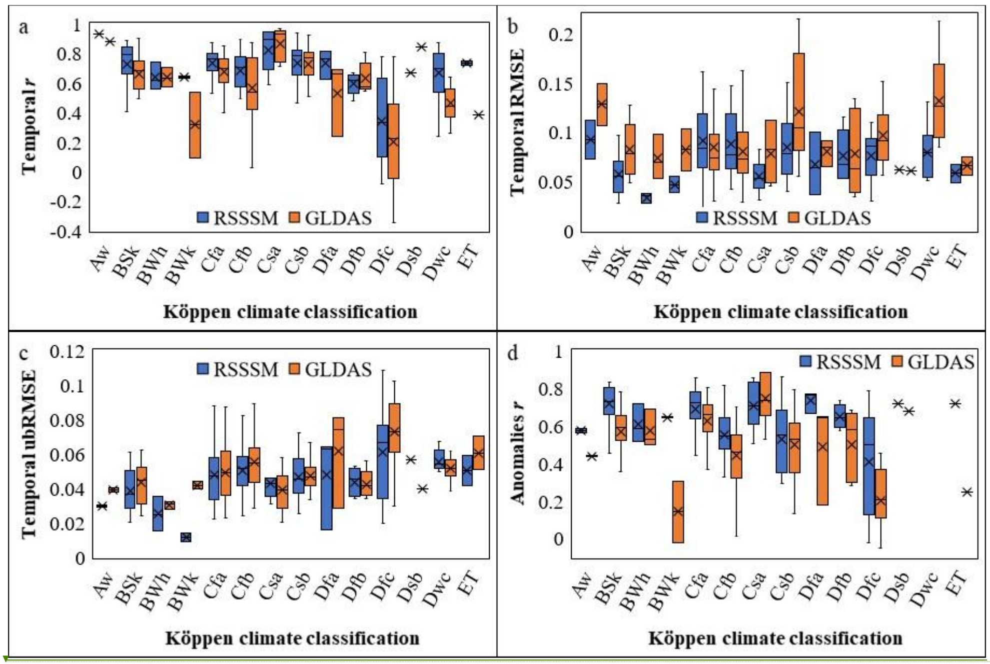
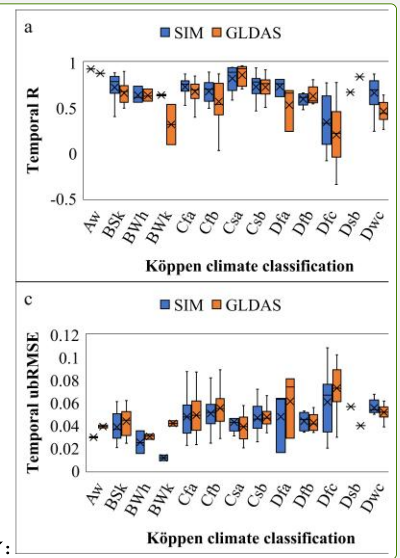


Figure 11: Comparison between the temporal accuracy of RSSSM and GLDAS surface soil moisture in regions with different Köppen-Geiger climate types. The four indexes are (a) *r*, (b) RMSE, (c) ubRMSE, and (d) Anomalies *r*.

删除了:



删除了: :

删除了: *R*

设置了格式: 字体: 倾斜

删除了: *R*

设置了格式: 字体: 倾斜

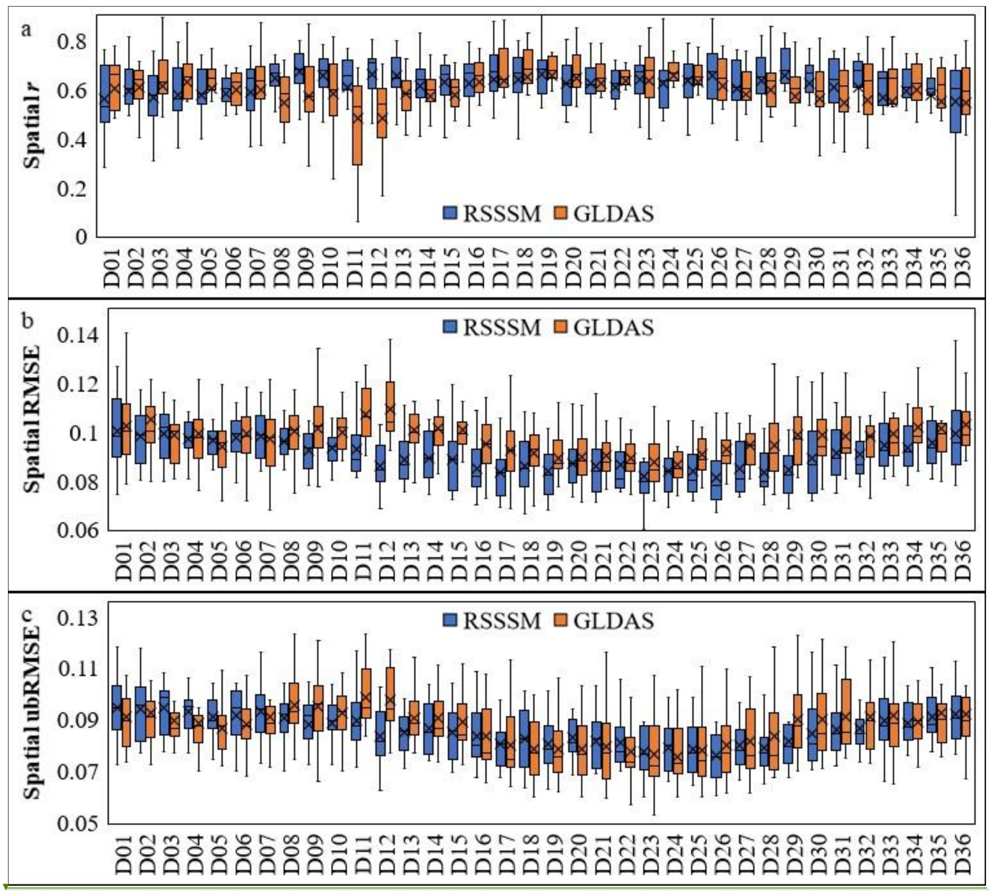
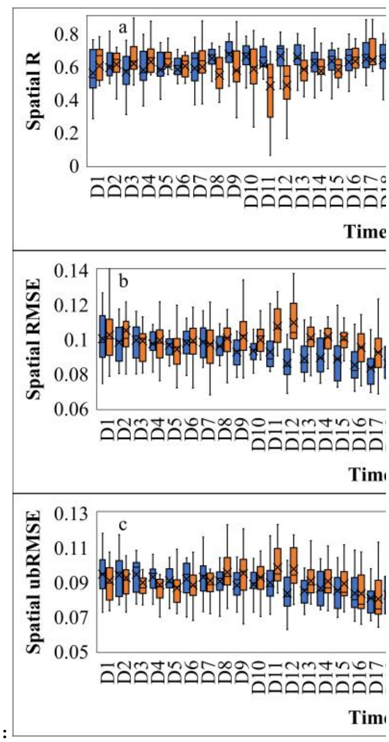


Figure 12: Comparison between the spatial accuracy of RSSSM and GLDAS during different 10-day periods. The evaluation indexes are the same as those in Figure 7.

删除了:



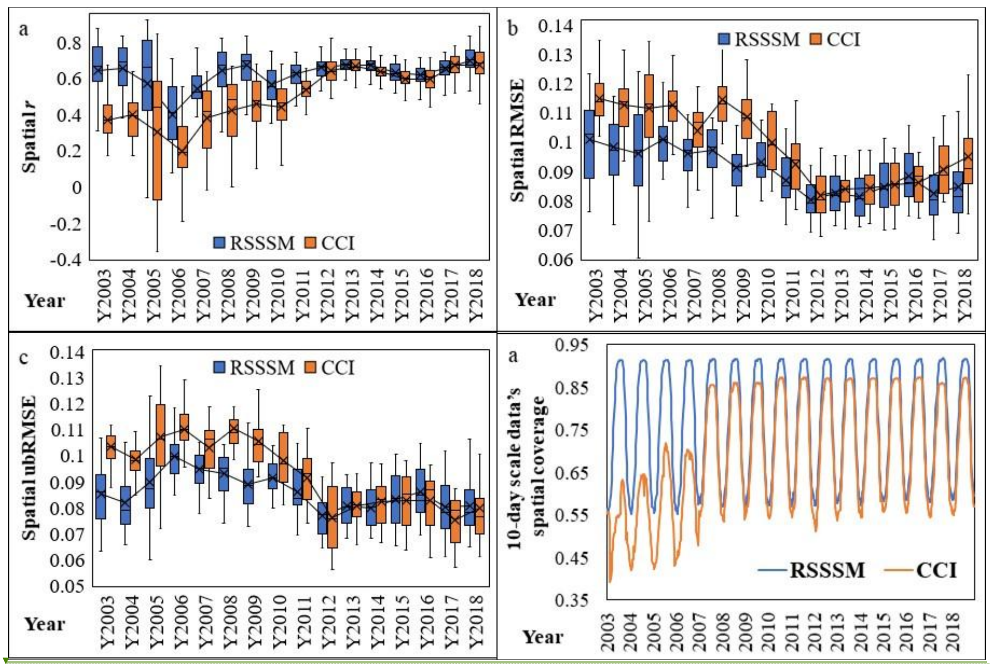
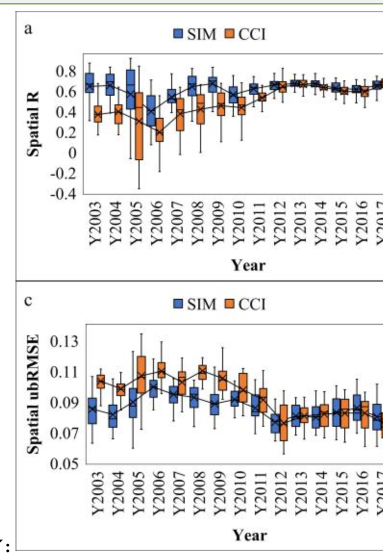


Figure 13: Changes in the data quality and data spatial coverages of RSSSM and CCI soil moisture with year. The interannual changes in (a) spatial correlation coefficients (*r*), (b) spatial RMSE, (c) spatial ubRMSE values, and (d) the spatial coverages of 10-day period data of RSSSM and CCI.

删除了:



删除了: 11

删除了: SIM

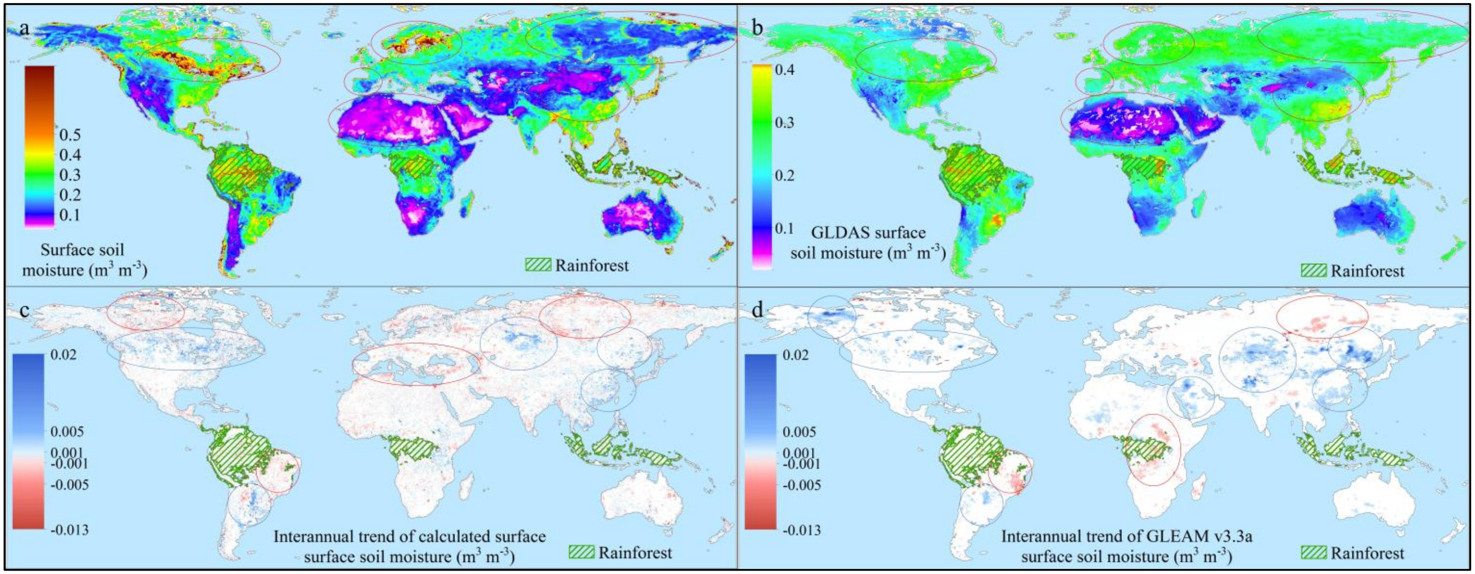


Figure 14: The spatial and temporal patterns of the neural network simulated surface soil moisture (RSSSM) and comparison against other products: (a~b) the global map of (a) calculated RSSSM and (b) GLDAS Noah V2.1 soil moisture (averaged during 2003~2018); (c~d) the interannual trend map of (c) calculated RSSSM and (d) GLEAM v3.3a soil moisture from 2003 to 2018. The circled regions in (a~b) are the places with obvious differences between RSSSM and the other products, while the circled regions in (c~d) are those with significant trends.

删除了: 12

删除了: calcula

删除了: calcula

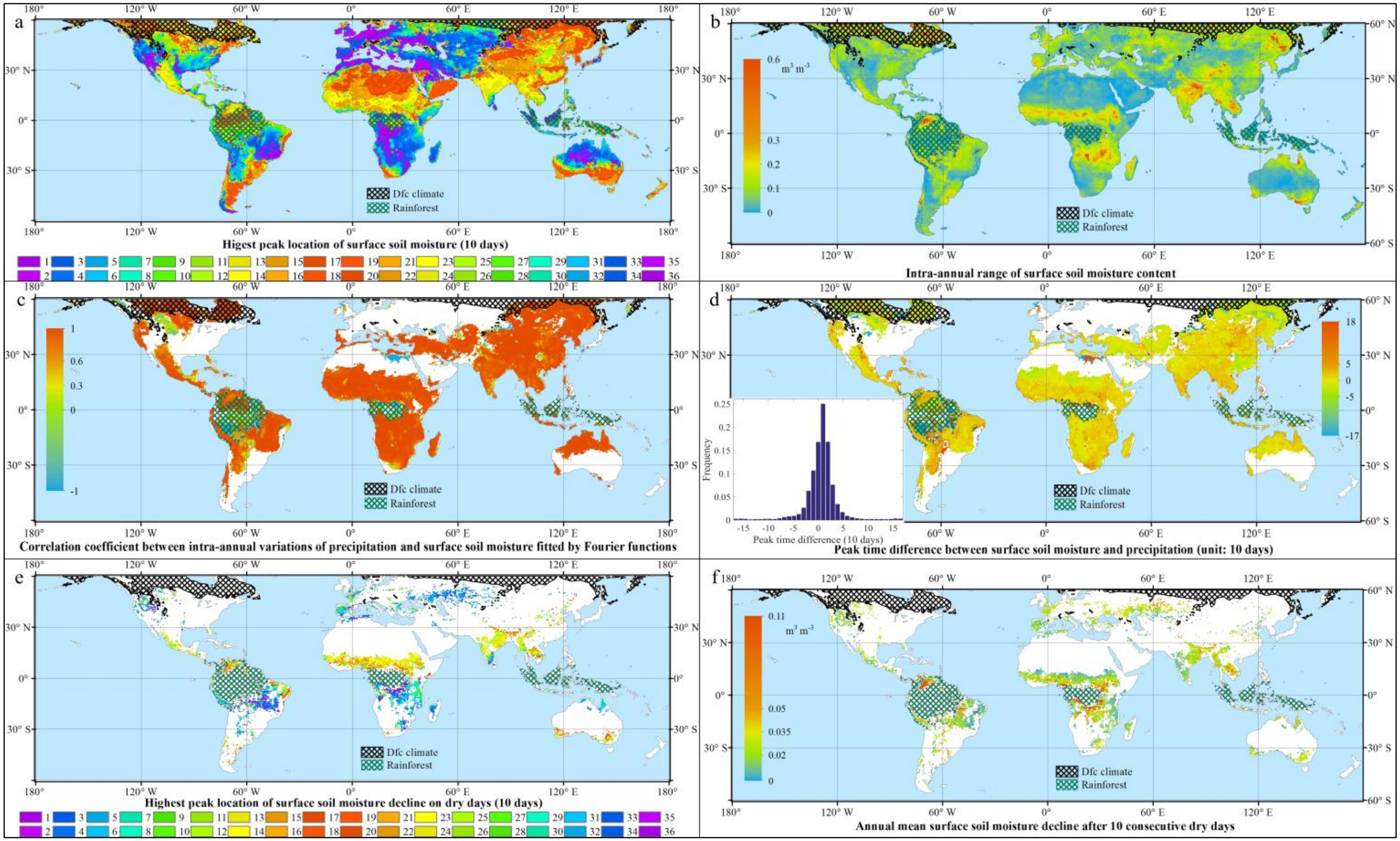


Figure 15: The intra-annual variation in global surface soil moisture and its relationship with precipitation. (a) The spatial pattern of the time when surface soil moisture reaches its maximum in a year (unit: 10 days, note that the seasons are opposite in the Northern and Southern Hemispheres); (b) the intra-annual variation range of surface soil moisture; (c) the map of the correlation coefficient between the intra-annual variations in precipitation and surface soil moisture (both are fitted by Fourier periodic functions); (d) the peak time difference between the surface soil moisture and precipitation (unit: 10 days), with the frequency histogram shown as the inset; (e) the 10-day period with the fastest surface soil moisture loss on rainless days in every 0.5° grid over the world; and (f) map of the annual mean surface soil moisture decline after 10 consecutive dry days (if assuming that the dry period occurs randomly throughout a year).

删除了:
删除了: at glo
删除了: globa
设置了格式: 字
删除了: 13
删除了: distrib
of the calculate
删除了:
删除了: the glo
删除了: intra-a
删除了: and
删除了: the ins
删除了: spatial
删除了: highest
删除了: decline
删除了: dry
删除了: (unit:
删除了: estima
删除了: declin
删除了: in
删除了: over
删除了: