**Interactive comment on '*An improved dataset of satellite observation-based global surface soil moisture covering 2003~2018 (RSSSM)*'**

By Yongzhe Chen, Xiaoming Feng, Bojie Fu

**To Reviewer #2:**

We thank referee #2 for the valuable comments that will help us improve the quality and readability 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.

***********

**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 methods, is not clear and we apologize for some confusing or incorrect terminologies. These problems will make the readers and reviewers misunderstand several rather complex algorithms, which are also the major innovation points of this study, including the selection of quality impact factors as neural network 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 doubts or queries, we explained our design of the method in greater detail and 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.

**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 the spatial correlation coefficient, spatial RMSE and unbiased RMSE, etc.). Please see section 2.3 in the Methods in the revised manuscript for details, while the accuracy comparison among all products is in section 3.2 of the results. The validation results indicate that our RSSSM product is comparable to the site measurements, in terms of both $R^2$ and the 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 and relatively cold areas (Figure 8, Figure 11, Figure S7 and Table 2). The temporal accuracy of RSSSM is similar to that of ERA5-Land and GLEAM v3.3 products (the 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 superior to all other products (please refer to Figure 9, Figure 12, Figure S5, Figure S8, Figure S11 and Table 3) throughout almost the entire year, especially during the growing seasons. Based on these findings, we propose that our product (RSSSM) has 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. Therefore, we corrected the sentences as follows: '*This new dataset, named RSSSM, is 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 spatial and temporal patterns.*' We have also corrected all the relevant unclear statements (e.g., supposed superior, expected to be better, …) throughout the manuscript.

**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 apologize for the unsuitable terminologies. We corrected the sentence

'the penetrability of microwaves is usually <5 cm of soil' to '*current satellite microwave sensors can detect only soil moisture within the top 5 cm of soil*' following this comment and your Specific comment 5, and corrected '*dielectric conductivity*' to '*dielectric constant*'.

**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.

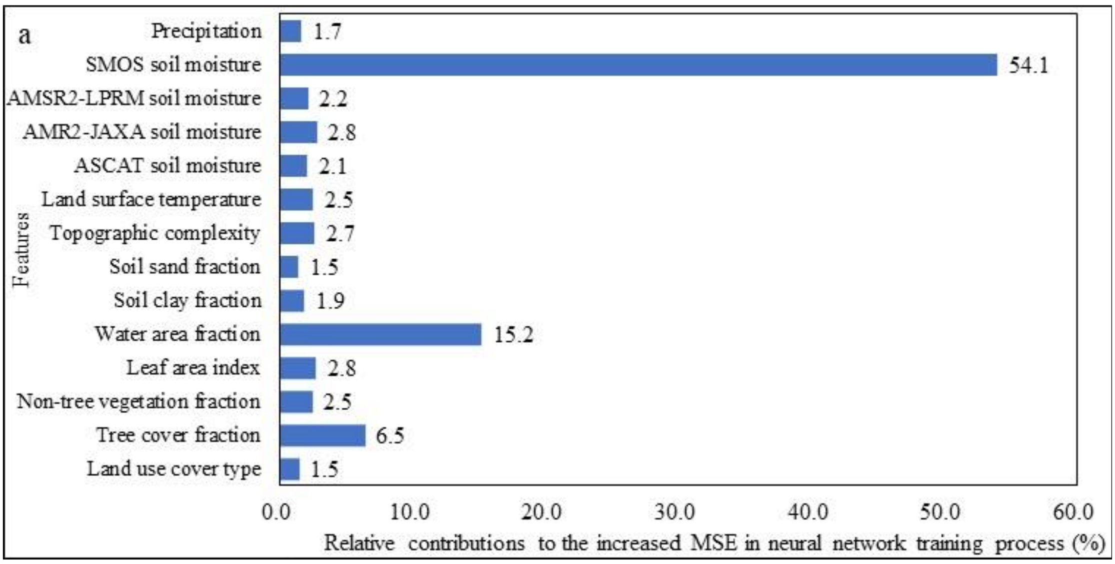
**Response:** We checked the work by *Qu* et al. (Qu et al., 2019) carefully and found that precipitation data were actually not applied as an input for the random forest in that study.

Furthermore, other recent studies focusing on long-term soil moisture mapping based on microwave remote sensing data also did not incorporate precipitation as ancillary inputs for the neural networks (Santi et al., 2016; Yao et al., 2019; Yao et al., 2017). Hence, we conducted 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 globe, for 11% of global land, there will be only 5 or fewer observations for random days within a 10-day period. By taking the average of these available data, this study focuses on only the mean soil moisture condition during that 10-day period. Then, to see how much the incorporation of precipitation data can improve the neural network training efficiency, we calculated the 10-day averaged GPM Final-Run precipitation, which can indicate the overall precipitation water availability (the antecedent precipitation index is not used because it must be calculated on a daily scale, and the attenuation coefficient is difficult to determine at a 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 soil moisture predictor 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 an increase in MSE during 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 contribute to only 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 result 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 are related to surface moisture conditions (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 already be indicated by the microwave soil moisture products. Therefore, 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 mechanisms (see section 2.1.2). Therefore, theoretically, these variables should be added to the neural network, even though the land use cover type and soil sand fraction data have been proven 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 neural network training. This result is true because precipitation directly contributes to increases in soil moisture. However, NARX is not suitable for global-scale long-term continuous soil moisture mapping because the base map (i.e., the soil moisture at the beginning of the simulation period) is difficult 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 are otherwise available. Therefore, if NARX is adopted, we can only estimate long-term soil moisture in the tropics and subtropics with air temperatures consistently higher than 0 °C. Finally, 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, simulations based on previous soil moisture and current precipitation products will lead to errors in regions 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 reduced in irrigated croplands and afforestation/deforestation areas as well.



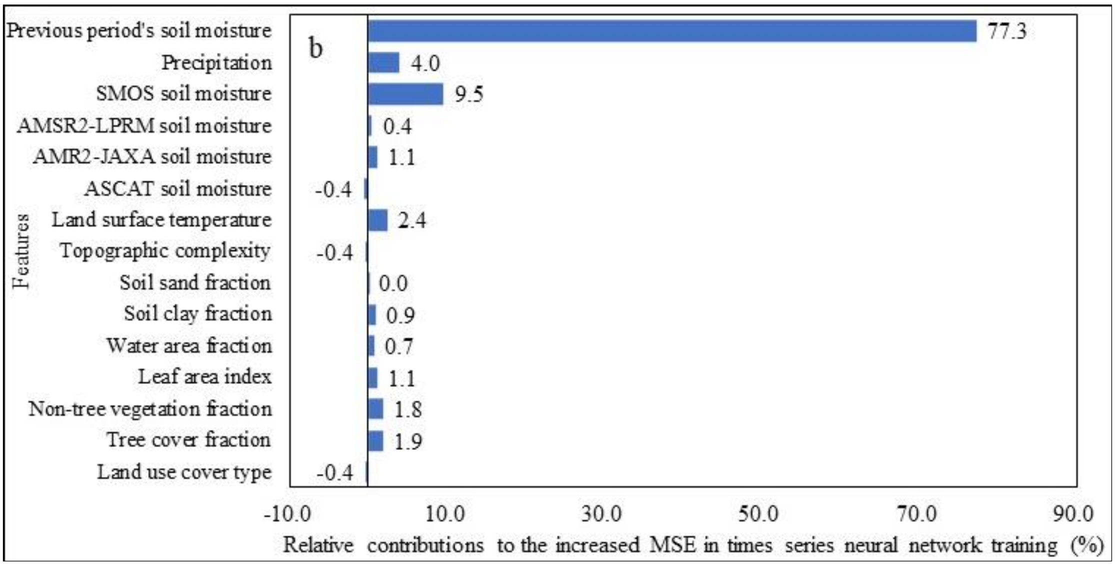



*Figure R1: (Figure S1 in the revised Supplement): The roles of different input features in the soil moisture simulations based on BP neural networks and nonlinear autoregressive with external input (NARX) with microwave soil moisture products incorporated: (a) the contributions of different input features 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 of the above, precipitation data are 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 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).*'

In Qu et al.'s (2019) study, the random forest input features include only microwave Tb products, DEM, IGBP global vegetation classification, latitude, longitude, and DOY.

Therefore, we apologize that the latter half of this sentence was not correct, which we have deleted. The sentences now read: '*Another study rebuilt a soil moisture time series over the Tibetan Plateau by using SMAP data as the reference data for a random forest model (Qu et al., 2019). For the environmental factors, while vegetation cover was 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 only the Tibetan Plateau) and suffer from the temporal gap between AMSR-E and AMSR2 data (October 2011~June 2012).*'.

**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:** We apologize for the unsuitable sentences. We have rewritten them as follows: '*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 phenomenon occurs because although these data 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 small scales, leaf area index (LAI) may have a negative influence on soil moisture due to the variation in evapotranspiration*

*(Naithani et al., 2013) or may not have 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 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 factors that impact the retrieval quality. The detailed explanations are as follows: 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 were not included as a predictor; see the detailed explanations in response to 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 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.

**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 apologize 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 a size of 0.1°×0.1° (approximately 120 km$^2$), there could be a small fraction of water, which may be due to the presence of rivers, streams, ponds, partly inundated wetlands or paddy croplands. If we mask out all 0.1° resolution grids with even 1% 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 decrease the brightness temperature (Tb), while different retrieval algorithms correct the impact of water within the 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 when compared to nearby lands. Moreover, the sensitivity of different microwave sensors to the water fraction within the 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' as both a quality impact factor and an ancillary soil moisture indicator (please also note that the word 'predictor' in this manuscript refers to only the existing soil moisture products that are applied as the neural network inputs). This information has been added to the manuscript. For the '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) proves 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 factor 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. 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, the daily water area fraction derived* *from the Surface WAter Microwave Product Series (SWAMPS) v3.2 dataset (Schroeder et al., 2015) is applied.*' The confusing words 'water body' have been replaced by '*water area fraction*' in other parts of the manuscript as well.

**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' was due to data availability. SMAP uses only GMTED 2010 DEM data to derive quality flags for data retrieved in mountainous areas, while topographic complexity is included as only 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 are not included in the static layers of the ASCAT-SWI product, so we had to obtain these 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 if they come from different soil moisture products. In the revised manuscript, we added this information: '*(topographic complexity data are not available from SMAP Constant; soil texture is not provided by ASCAT Constant)*'.

**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 denoising' 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. Therefore, if there are extreme heat or precipitation events, as you noted, the whole 1°×1° zone will exhibit 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, extreme weather events will not be removed by this 'spatial 3σ denoising' step.

We apologize 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 denoising' (note that denoising is conducted spatially, rather than temporally, so that the extreme events will not be treated).*'

**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:** We apologize 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 from 2012~2013. In this study, the common period for the 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 long-term continuous satellite-based soil moisture mapping, almost full spatial coverage at the global scale, and high data accuracy. Following this comment, we have revised this section in the manuscript to clarify the NN design. It reads as:

'*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, 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 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 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 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 from 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 soil moisture predictor 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 soil moisture predictor 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 soil moisture predictor data entering the neural networks consist of WINDSAT, TMI, AMSRE-JAXA, AMSRE-LPRM and AMSRE-NSIDC (Table S14~S15).*’

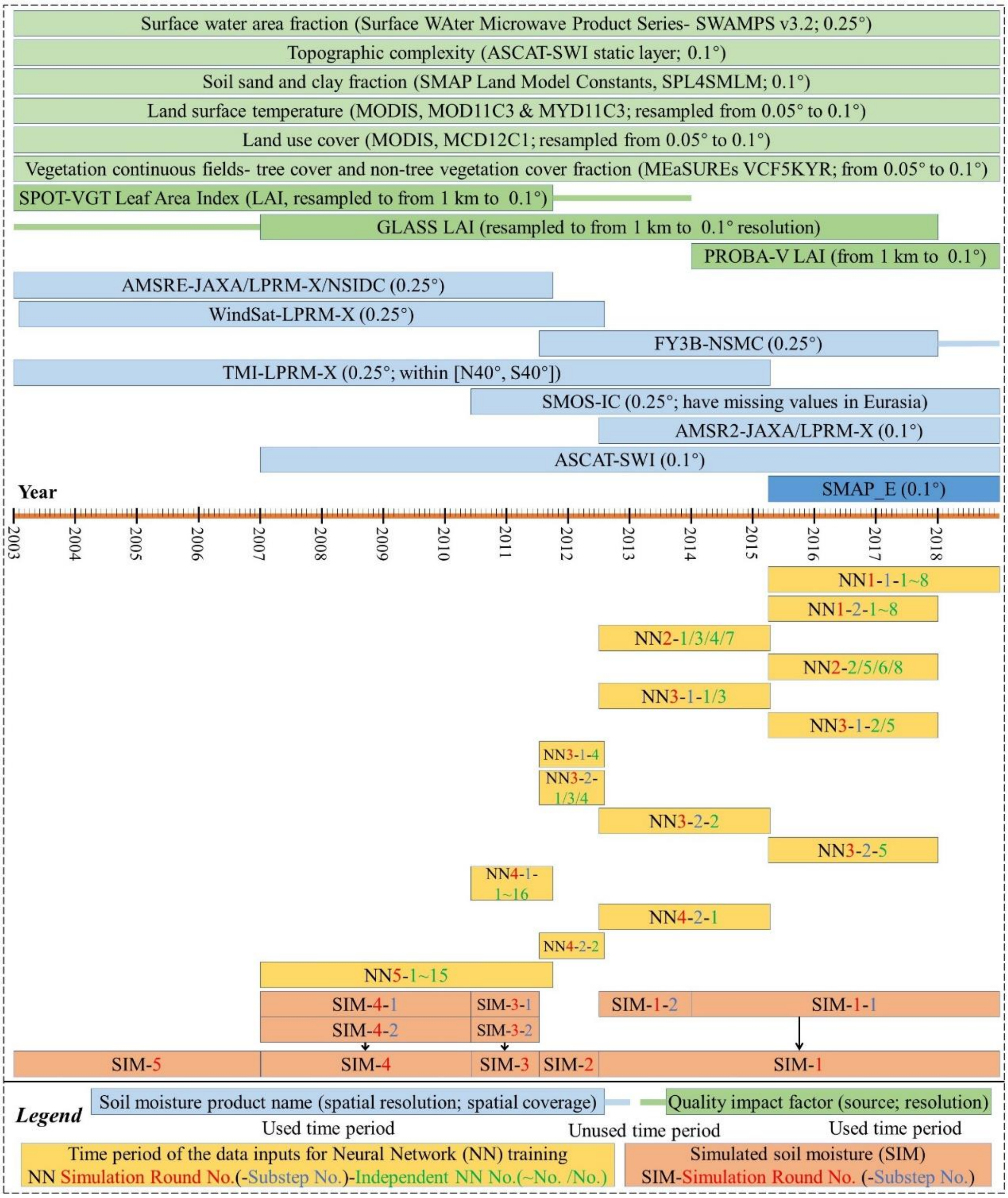



*Figure R2 (Figure 1 in the revised manuscript): The timeline figure showing the time periods of the soil moisture datasets and the 'quality impact factor' products (e.g., LAI dataset) used in this study (listed above the timeline), as well as the periods of data applied for the training of the 67 independent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

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 in each round of the NN. In detail, SMAP soil moisture data are used as the training target of the first round NN (labeled NN1), with ASCAT-SWI, SMOS, AMSR2-JAXA and AMSR2-LPRM-X applied as soil moisture predictor 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 be used for the simulation only 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 the time 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 does 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 is 4+3+2+1=10, 8 of them are valid since the soil moisture retrievals over snow or ice are 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 NN1-1(2)-1 ~ NN1-1(2)-8; for example, NN1-1(2)-1 is trained using all four soil moisture predictor 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. This problem occurs 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), which is denoted by 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 (primary 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 the round 2 NN (see Table S3), while the corresponding simulation output is SIM-2, covering the period from 2011D20~2012D18 since the FY data product has been 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, 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).

**Specific comment 1.** In the abstract, change 'elaborate' to 'elaborated', delete 'various', change 'simulation' to 'simulations'

**Response:** We have corrected them accordingly.

**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 follows: '*This new dataset,* *named RSSSM, is 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 spatial and temporal patterns.*' Please also find the details in the response to Major Comment 1.

**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:** We apologize for the too complicated sentences. We have shortened the sentence as: '*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).*'

**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 follows: '*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).*'

**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 carefully checked this information and determined that the L-band microwave is sensitive to the soil moisture within only <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 microwaves, which is 1 cm (C band) or less (Piles et al., 2018). We agree that the word 'penetrability' is not scientific, while 'penetration depth' is not very accurate as well. Therefore, we have changed the sentence to '*current satellite microwave sensors can detect only soil moisture within the top 5 cm of soil)*' for clarity.

**Specific comment 6.** Line 45: change 'frommicrowave' to 'from microwave'.

**Response:** We have made the revision accordingly.

**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*

*ASCAT product represents the longest continuous record of global surface soil moisture that is derived from only microwave remote sensing*.'

**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 to: '*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).*'

**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.

**Specific comment 10.** Change 'deviationsto' to 'deviations to'.

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

**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 follows: '*The training R-square value ($R^2$) of this product was only 0.45 (or correlation coefficient, r, equals 0.67).*' Following this comment, we have also corrected all the 'R' to '*r*' to distinguish it from $R^2$ in the revised manuscript.

**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, these data are regional (for only the Tibetan Plateau) and suffer from the temporal gap between AMSR-E and AMSR2 data (October 2011~June 2012).*' We have also rewritten Text S1 to more clearly explain why precipitation was not included as an input feature in the neural networks in this study. Please also find the details in the responses to Major Comment 3.

**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 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 phenomenon occurs because although these data 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 small scales, leaf area index (LAI) may have a negative influence on soil moisture due to the variation in evapotranspiration (Naithani et al., 2013) or may not have 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 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 factors that impact the retrieval quality. The detailed explanations are as follows: 1) …*'.

**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.

**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.

**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. This process is used because topographic complexity and soil texture data cannot be obtained from one product. Additionally, the quality for the two data are satisfying (Reichle et al., 2018; Scipal et al., 2005). Please see the response to Major Comment 6.

**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 spatial (zonal) ‘3σ denoising’, which helps to mask out the incorrect retrievals in mountain areas (zonal extreme values); however, it will not mask out extreme climatic events. Please find the detailed explanation in the response to Major Comment 7. Following this comment, we have made the clarification as follows: ‘*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 denoising is conducted spatially, rather than temporally, so that the extreme events will not be treated).*'

**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 from 2012~2013. In this study, the common data period for the 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 differences in the data periods 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 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 GLASS LAI covers the period from the beginning of our study period until 2017, the training period for substep 2 is 2015D10~2017).*' For more details, please see the response to Major Comment 8.

**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 that were utilized as predictors. For example, the simulation period (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 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 from 2011D20~2012D18, which 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.

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

**Response:** For the 3$^{rd}$ 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, SIM-1T (SIM-1T is the postprocessed simulation output of the first round NN), or both SIM-1T and SIM-2T. Detailed information on NN training and soil moisture simulation in Round 3 is provided in Tables S5~S8. We have also revised the sentence in the manuscript as follows: '*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 soil moisture predictor 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

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
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
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