

Note: The modifications are shown in green. The responses to comments are blue colored.

We are very grateful to Reviewer for reviewing the paper so carefully. These comments are very helpful to improve the quality of the manuscript. Please find my itemized responses in below and my revisions will be in the re-submitted files.

### Major comments:

Comment#1-1 Why use the ERA5-LAND dataset instead of weather station data correspondence with the CMA SM Observations? As far as I know, each soil moisture observation station of China Meteorological Administration will have a corresponding meteorological station.

Responds: Thanks for your kind comments and helpful suggestions. The reason using the ERA5-LAND dataset is that it is a relative accurate dataset for all types of land applications and it has seamless spatial distribution. The SM data with seamless spatial distribution can be generated by establishing the relationship between ERA5-LAND dataset and CMA SM Observations. However, the meteorological stations are too sparse to capture adequate spatial coverage, which cannot cover the whole China. Our data generating method follows the current research of most soil moisture products (Sungmin and Orth 2020, Zeng et al. 2019, Carranza et al. 2020, Karthikeyan and Mishra 2021), which use spatial continuous meteorological data as the predictors instead of individual meteorological stations. As the results shown in our work, this way can already provide satisfying performance in the data producing.

However, the valuable opinion of the reviewer also gives us a new inspiration. In the future work, we can first collect the meteorological data from the corresponding meteorological station. To achieve the seamless spatial distribution, we then apply spatial interpolation method to generate the meteorological data, or just use other existing gridded meteorological data. Finally, RF model is used to establish the relationship between the covariates and CMA SM Observations, and the SM data with seamless spatial distribution is further generated.

[1] Carranza, C., Nolet, C., Pezij, M., and van der Ploeg, M.: Root zone soil moisture estimation with Random Forest, *Journal of Hydrology*, 593, 125840, <https://doi.org/10.1016/j.jhydrol.2020.125840>, 2021.

[2] Karthikeyan, L. and Mishra, A. K.: Multi-layer high-resolution soil moisture estimation using machine learning over the United States, *Remote Sensing of Environment*, 266, 112706, <https://doi.org/10.1016/j.rse.2021.112706>, 2021.

[3] O, S. and Orth, R.: Global soil moisture data derived through machine learning trained with in-situ measurements, *Scientific Data*, 8, 170, <https://doi.org/10.1038/s41597-021-00964-1>, 2021.

[4] Zeng, L., Hu, S., Xiang, D., Zhang, X., Li, D., Li, L., and Zhang, T.: Multilayer Soil Moisture Mapping at a Regional Scale from Multisource Data via a Machine Learning Method, *Remote Sensing*, 11, <https://doi.org/10.3390/rs11030284>, 2019.

Comment#1-2: Most of the study areas are not natural areas, but farmland areas that have been affected by agricultural management practices for a long time. However, the random forest model does not consider human management measures, such as fertilization, irrigation, farming measures, etc., which I think is the biggest deviation of the results of this study.

Responds: Thanks for your kind comments. The SM in farmland areas is indeed affected by more factors than that in natural areas. However, the human management measure data are hardly obtained accurately and not available covering the whole farmland areas. We agree that this kind of data will help in improving the SM estimation in crop land as long as they are available. However, we also argue that the SM observations have already contains the effect of human management to some extent, as these measures will directly affect SM itself.

Of course, according to the valuable comment, it also provides an important direction for our future study. Specifically, for a specific farmland area, we will first collect or even developing the agricultural management data, such as fertilization, irrigation, farming measures, etc., and further explore how the agricultural management data affect the variation in SM. It is expected to be an important supplementary scheme to optimize the SMCI1.0 product.

Revise:

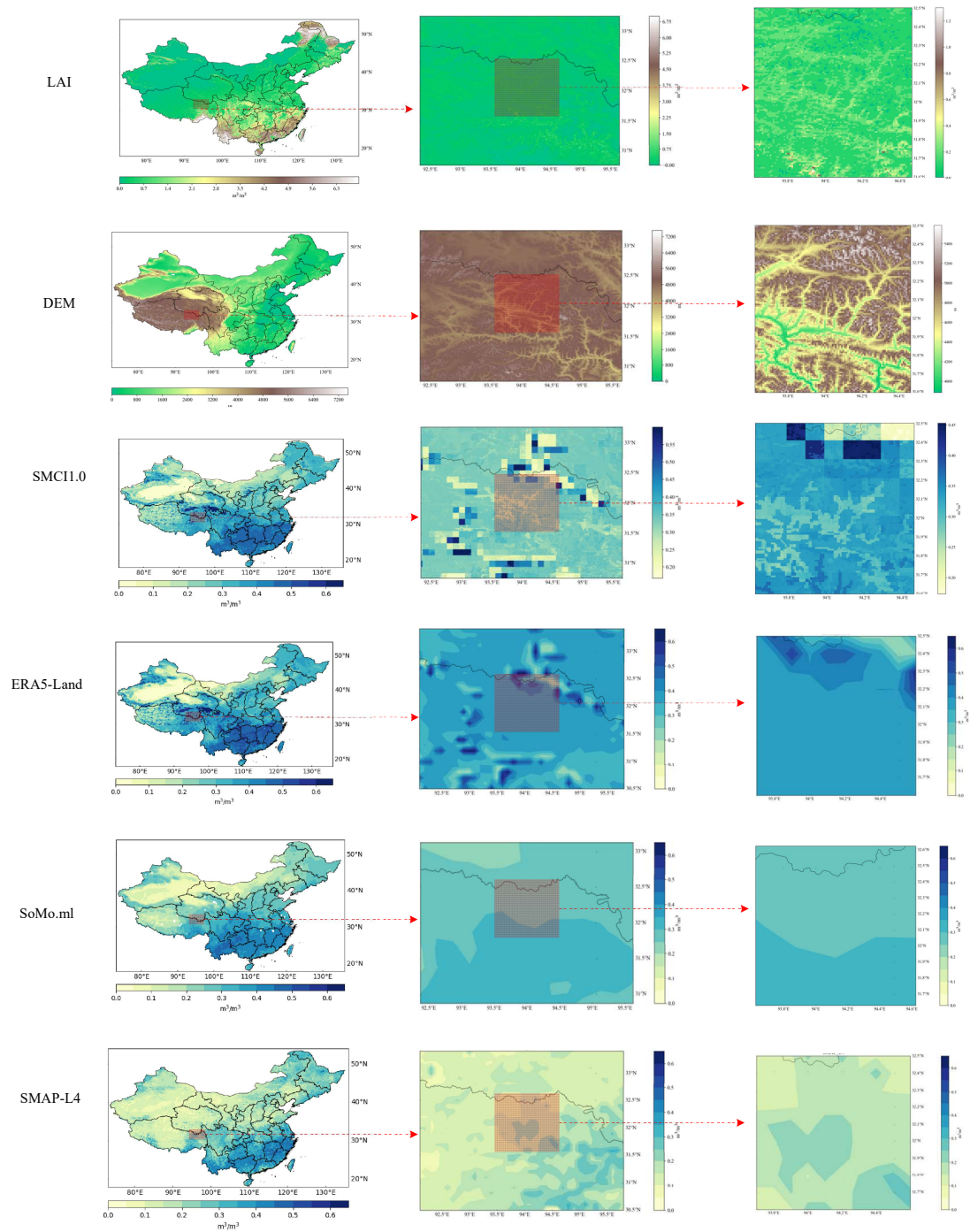
Other predictors should also be explored to improve the SM prediction. For the farmland areas, human management measures such as fertilization and irrigation should be considered in a proper way, even though this kind of data are rarely available in a spatial continuous way for the whole China. Lack of the consideration about agricultural management practices in the SMCI1.0 may lead to some deviation in SM estimation of the crop land.

Comment#1-3: As far as I am concerned, the study area includes the Qinghai-Tibet Plateau and the Northeast Alpine region, which are typical areas of freeze-thaw soil. Is it compared with the results of soil moisture research in these regions in China? Such as "A first assessment of satellite and reanalysis estimates of surface and root-zone soil moisture over the permafrost region of Qinghai-Tibet Plateau" and others.

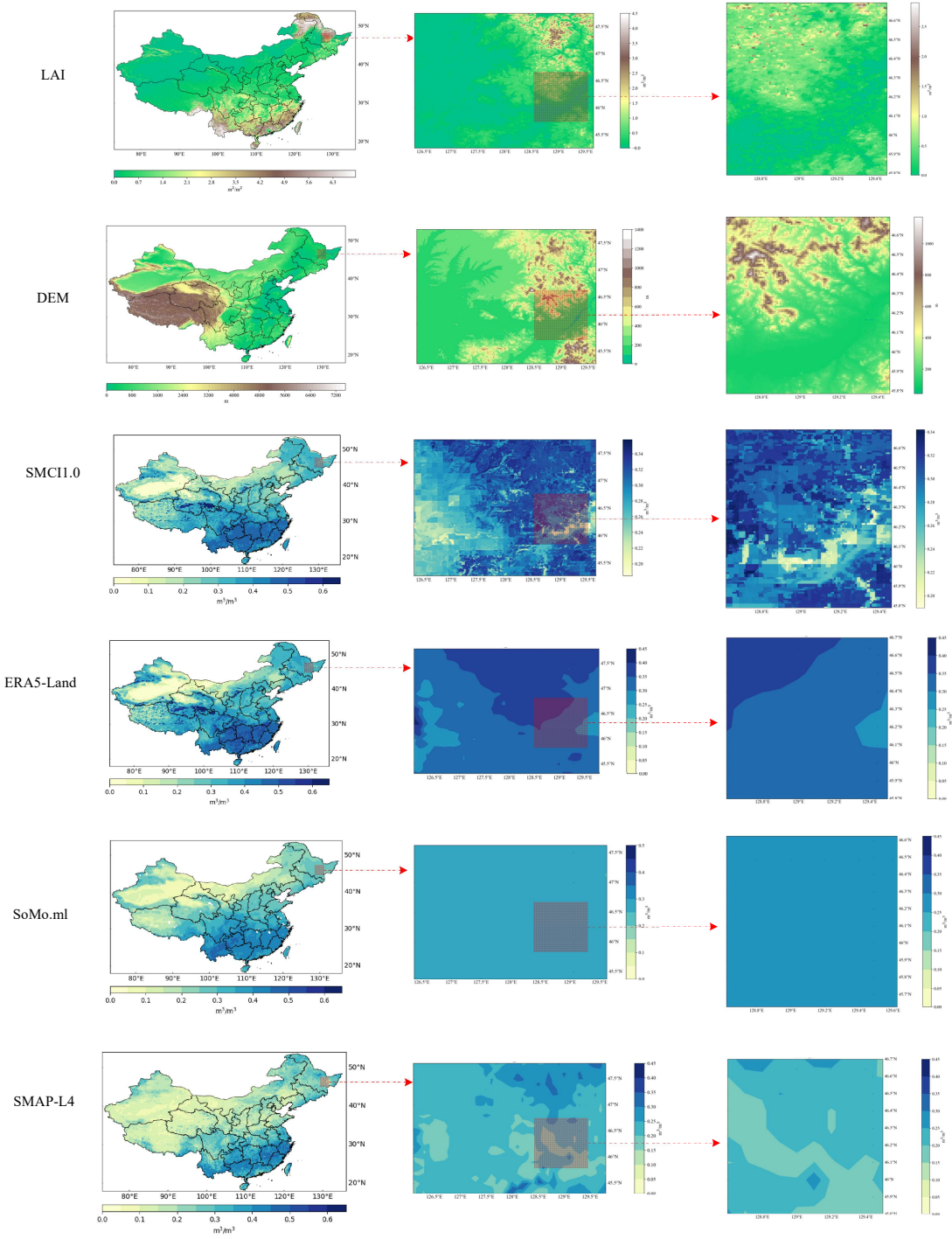
Responds: Thanks for your kind comments and helpful suggestions. We have added the assessment of SM maps from different products over the Qinghai-Tibet Plateau and the Northeast Alpine regions

Revise:

We also compared the SM estimation of the Qinghai-Tibet Plateau (74°00'E-104°00'E, 25°00'N-40°00'N) and the Northeast Alpine region (128°50'E-129°50'E, 45°50'N-46°70'N), as they are typical areas of freeze-thaw soil (Fig. S8 and S9). According to the previous study of Xing et al. (2021), the ERA5-Land often be overestimated compared to in-situ SM (see their Fig. 5). In Fig. S8, the SMCI1.0 SM is underestimated compared to ERA5-Land over Qinghai-Tibet Plateau, which is more closed to in-situ SM. Additionally, over both Qinghai-Tibet Plateau and Northeast regions (128°50'E-129°50'E, 45°50'N-46°70'N), the more details of SM spatial patterns for SMCI1.0 SM can be found than that of ERA5-land, SoMo.ml and SMAP-L4 SM (Fig. S8 and Fig. S9).



**Figure S8. Soil moisture maps from different products on 1st January 2016 over Qinghai-Tibet Plateau region. The resolution is 1km for SMC1.0, 9km for ERA5-land and SMAP-L4 and 0.25 degree for SoMo.ml.**



**Figure S9. Soil moisture maps from different products on 1st January 2016 over Northeast region. The resolution is 1km for SMC1.0, 9km for ERA5-land and SMAP-L4 and 0.25 degree for SoMo.ml.**