This manuscript explores a RF based approach to fill the spatial gaps in satellite SM observations. The paper is extensive and well organized. The literature review is also extensive, nevertheless it explores studies that are immediately similar to the study to a high degree (i.e. predicts SM from satellite sensors) that there might be some lessons learned in some related studies (i.e. predicts reanalysis SM from model outputs using observed predictors similar to those used in this study) that weren't discussed. Other specific comments are below.

Response: Thanks for your valuable time and constructive comments. We will try our best to incorporate them in the new revision. In the modified version, some references (see the last sections) will be enhanced, and some lessons learned in some related studies are added especially in the discussion and conclusion section. In addition, the English language usage in this version will be improved deeply by one fluent English researcher that is suggested by the commercial corporation.

1. L85 ElSaadani et al. 2021 "Assessment of a spatiotemporal deep learning approach for soil moisture prediction and filling the gaps in between soil moisture observations" discussed in detail the following issues that are relevant to this manuscript:

- Filling SM gaps saptio-temporally using a convlstm ML approach.
- The effect of the number of predictors and time step of prediction on the model performance.

In addition, Q Li et al "Improved daily SMAP satellite soil moisture prediction over China using deep learning model with transfer learning" built on the above work to apply it to SMAP observations while improving the ML convlstm accuracy. Please include the above references for completeness of literature review and lessons learned regarding the predictors and their effect on ML model performance.

Response: Thanks for pointing out this issue. We will enhance the references throughout the manuscript to make it more convincible.

2. Table 5 Please separate the two sides of the table properly

Response: Thanks for pointing out this version. The original Table 5 has been adjusted in the new version.

Table Metrics for the gap-filling performance regarding Maqu network for the extended years

Year	R^2		RMSE (cm^3/cm^3)		MAE (cm^3/cm^3)		Bias (cm^3/cm^3)		ubRMSE (cm^3/cm^3)		NSE	
	CCI	gap-filled	CCI	gap-filled	CCI	gap-filled	CCI	gap-filled	CCI	gap-filled	CCI	gap-filled
2008	0.8	0.71	0.11	0.13	0.1	0.13	0.06	0.07	0.06	0.06	0.8	0.81
2010	0.82	0.73	0.1	0.11	0.09	0.11	0.05	0.06	0.06	0.05	0.81	0.83
2011	0.83	0.74	0.09	0.11	0.09	0.1	0.06	0.06	0.06	0.05	0.82	0.84

2012	0.81	0.72	0.12	0.13	0.09	0.12	0.06	0.05	0.05	0.05	0.81	0.82
2013	0.82	0.73	0.09	0.12	0.09	0.13	0.06	0.07	0.05	0.07	0.8	0.82
2014	0.85	0.74	0.09	0.11	0.08	0.09	0.06	0.08	0.05	0.06	0.83	0.85
2015	0.79	0.69	0.12	0.14	0.1	0.12	0.07	0.09	0.07	0.07	0.79	0.81

Note: NSE is from the evaluation with the time series of average 0.25^{o} pixels while the other five metrics are from the evaluation with 1 km disaggregated values.

3. L143 please check grammar and explain why regular DEM wasn't used

Response: Earlier studies [1] found that topographic position index (TPI) correlated best with surface variables such as snow depth and soil moisture amongst several other terrain-derived indices, with the strength of the correlation varying as a function of elevation range and topographic complexity in the designated neighborhood. Larger neighborhoods are more likely to reveal larger-scale terrain features such as valleys and ridges, while smaller neighborhoods more likely identify local depressions/fissures [2]. Meanwhile, our analysis observes that TPI has the higher Pearson's correlation with soil moisture in comparison with DEM (0.317 versus 0.288 in significance percentage, and 0.172 versus 0.153 in significance percentage), which encourages our work for using TPI rather than the DEM. In the new version, we will add the related context.

[1] Cristea N C, Breckheimer I, Raleigh M S, et al. An evaluation of terrain-based downscaling of fractional snow covered area data sets based on LiDAR-derived snow data and orthoimagery. Water Resources Research, 2017, 53(8): 6802-6820.

[2] Shaw T E, Gascoin S, Mendoza P A, et al. Snow depth patterns in a high mountain Andean catchment from satellite optical tristereoscopic remote sensing. Water Resources Research, 2020, 56(2): e2019WR024880.

4. L335 (important comment) please explain the reasoning behind the separation percentage of 90 10 training testing and how this subset was extracted. Is it random undefined subset or is a certain defined period was extracted and why.

Response: The artificial gaps were created to mitigate potential bias in model performance for any particular gap sequence. Model performance was tested by 10-fold cross validation (CV). The gap-filling accuracy was evaluated by a spatiotemporal 10-fold CV, i.e., randomly holding out 10% of pixels, training the model with the remaining 90% of the pixels, making evaluating on the held-out stations, and repeating this process 10 times. Previous studies generally implemented regular spatial or temporal 10-fold CV by randomly choosing 10% subset of each band or 10% period of each grid each time. However, a much stricter spatial and temporal CV procedure [1, 2] was selected in this study by rearranging the pixels of during all studied period into one time series and then dropping 10% samples each time (leave-one-year-out-cross-validation) to test the capacity of gap-filling soil moisture. According to the reviewer's suggestion, we will revise the related context in this version.

[1] Meng X, Liu C, Zhang L, et al. Estimating PM2. 5 concentrations in Northeastern China with full spatiotemporal coverage, 2005–2016. Remote sensing of environment, 2021, 253: 112203.

[2] Zhang D, Du L, Wang W, et al. A machine learning model to estimate ambient PM2. 5 concentrations in industrialized highveld region of South Africa. Remote Sensing of Environment, 2021, 266: 112713.

5. Figure 13: please reword the description of panel (a) regarding the text in the figure, does that describe the decrease in performance as a percentage of the original value?

Response: This denotes the relative percentage of the decreased accuracy in comparison to the baseline accuracy using all seven variables. We will revise the related context in the new version.

6. Figure 15 please explain in writing the effect of the significance level on the accuracy of conclusions and interpretation of the results.

Response: As shown in Fig. 15(d)-(f), the difference in valid SM values participating in trend analysis causes a disparity in calculating SM trend, i.e., bringing a lower SM trend in most wet regions but an increased SM trend in some dry regions when gap-filled values are introduced. As a result of the missing satellite soil moisture retrievals, the 8-days SM tends to be overestimated. In particular, most of the regions owing significant trend demonstrate a lower trend when comparison to the original values. Meanwhile, the confidence level of SM trend is converted from a significance level to non-significance level for a considerable fractional of the grids. This is more pronounced in wet regions such as northeast, northwest and southwest parts of China, which is sensitive to monsoon precipitation and ice melting. Our results are corroborated by earlier studies that revealed an overestimation in trend of missing AOD [1] and albedo [2] when the cloudy conditions prevented satellite retrievals. In general, the variations in SM trend are related to the climate variables (e.g., precipitation) changes and land management activities (Li et al., 2018). According to the reviewer's suggestion, we will revise the related context in this version.

[1] Zhang R, Di B, Luo Y, et al. A nonparametric approach to filling gaps in satellite-retrieved aerosol optical depth for estimating ambient PM2. 5 levels. Environmental Pollution, 2018, 243: 998-1007.

[2] Gunnarsson A, Gardarsson S M, Pálsson F, et al. Annual and inter-annual variability and trends of albedo of Icelandic glaciers. The Cryosphere, 2021, 15(2): 547-570.

7. Figure 16 is difficult to interpret, please make sure to have a proper legend that makes the figures self-explanatory, especially in the lower panels. Reading the figure caption to understand symbols adds difficulty to the interpretation.

Response: Thanks for pointing out this issue. The original Figure 16 and the related caption will be adjusted in the new version.

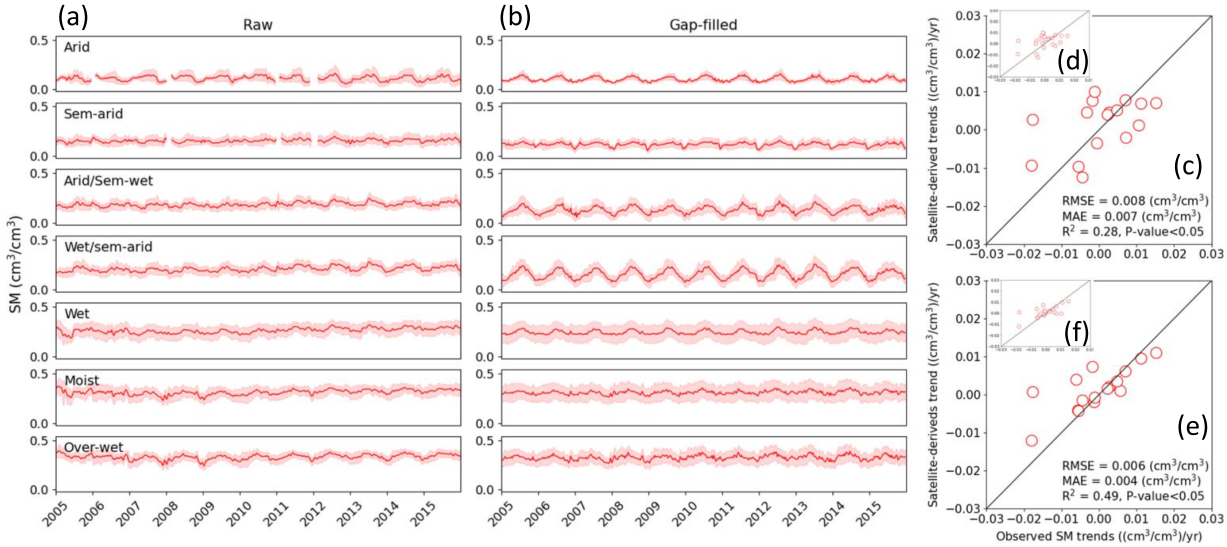

Figure 14: (a) shows the temporal patterns of raw CCI SM regarding different climate regions during 2005-2015, and (b) shows the temporal patterns of gap-filled CCI SM regarding different climate regions. The shaded area in (a) and (b) denotes ±1 standard error. (c) and (d) The scatter plot of 1-km CCI SM-derived trends against in situ measures during 2005-2014, and (c) shows the trends under significance level (P<0.05), while (d) shows all the trends. (e) and (f) The scatter plot of 1-km gap-filled SM-derived trends against in situ measures during 2005-2014, and (e) shows the trends under significance level, while (f) shows all the trends.

References

Li, B., Liang, S., Liu, X., Ma, H., Chen, Y., Liang, T., and He, T.: Estimation of all-sky 1 km land surface temperature over the conterminous United States, Remote Sensing of Environment, 266, 112707, https://doi.org/10.1016/j.rse.2021.112707, 2021.

Li, Q., Wang, Z., Shangguan, W., Li, L., Yao, Y., and Yu, F.: Improved daily SMAP satellite soil moisture prediction over China using deep learning model with transfer learning, Journal of Hydrology, 600, 126698, https://doi.org/10.1016/j.jhydrol.2021.126698, 2021b.

Li, Q., Li, Z., Shangguan, W., Wang, X., Li, L., and Yu, F.: Improving soil moisture prediction using a novel encoder-decoder model with residual learning, Computers and Electronics in Agriculture, 195, 106816, https://doi.org/10.1016/j.compag.2022.106816, 2022b.

Li, L., Dai, Y., Shangguan, W., Wei, N., Wei, Z., and Gupta, S.: Multistep Forecasting of Soil Moisture Using Spatiotemporal Deep Encoder–Decoder Networks, Journal of Hydrometeorology, 23, 337-350, 10.1175/jhm-d-21-0131.1, 2022.

Li, Y., Piao, S., Li, L. Z. X., Chen, A., Wang, X., Ciais, P., Huang, L., Lian, X., Peng, S., Zeng, Z., Wang, K., and Zhou, L.: Divergent hydrological response to large-scale afforestation and vegetation greening in China, Science Advances, 4, eaar4182, doi:10.1126/sciadv.aar4182, 2018.

Long, D., Bai, L., Yan, L., Zhang, C., Yang, W., Lei, H., Quan, J., Meng, X., and Shi, C.: Generation of spatially complete and daily continuous surface soil moisture of high spatial resolution, Remote Sensing of Environment, 233, 111364, https://doi.org/10.1016/j.rse.2019.111364, 2019.

Shangguan, W., Hengl, T., Mendes de Jesus, J., Yuan, H., and Dai, Y.: Mapping the global depth to bedrock for land surface modeling, Journal of Advances in Modeling Earth Systems, 9, 65-88, https://doi.org/10.1002/2016MS000686, 2017.