Letter to Editor

Dear Editor,

Thank you very much for handling our manuscript and providing constructive remarks. This has helped us to improve our manuscript significantly. The manuscript has now been thoroughly revised and strengthened based on all the comments and recommendations made by both reviewers. Please find below our detailed responses to each comment made by the reviewers.

We think that the revised manuscript has appropriately addressed all the reviewers' concerns and we hope that you can consider it for publication in Earth System Science Data.

Sincerely,
Pei Zhang
On behalf of all co-authors

We would like to thank the reviewer for carefully reading our manuscript and providing detailed and constructive comments. This has helped us to improve our manuscript significantly. In the text below we provide our response to each comment point by point.

Reviewer's comments are in **bold**.

Author's responses are in regular.

Author's additions/modifications in the text are in blue.

This manuscript introduces a valuable dataset (i.e., soil moisture and soil temperature at different depth from 5 cm to 80 cm) collected for nearly 10 years (from 2008 to 2019) at three dense networks including Maqu, Naqu, and Ngari built on the Tibetan Plateau. This dataset is a good extension of the surface soil moisture dataset introduced by Zhang et al. (2021, ESSD). The characteristics and trend of the in situ datasets as well as the derived freeze/thaw state were analyzed, and five reanalysis datasets of soil moisture and temperature were also evaluated by these in situ datasets. The paper is generally well written and organized. Prior to accepting this paper, the authors may want to address the following issue.

Thanks very much for your recognition of our work and for your constructive comments. We have carefully considered all the comments into the revision. Please find below the response to each comment for detailed information. We hope that the revised manuscript has appropriately addressed the reviewer's concerns and can be considered for publication in Earth System Science Data.

General comments:

1. **It is good to see the authors conducted a detailed evaluation of five reanalysis datasets, i.e., ERA5, GLDAS CLSM, GLDAS Noah, GLDAS VIC, and MERRA2. But in the current version, the authors only present the results which were not interpreted further. For example, why the ERA5 outperforms the other products in simulating the surface soil moisture? Why the model simulations generally underestimate soil moisture at different depths in Maqu network while generally overestimate soil moisture at Shiquanhe network? Why all the model simulated soil temperature at all depths shown a noticeable underestimation in both Maqu and Shiquanhe networks? Some explanations (or discussion) of these results will be helpful to enhance the robustness of the paper.**

Thanks for the comments and suggestions. We have provided additional explanations to the evaluation results of five reanalysis datasets in Section 4.3 of the revised version as shown below.

On Page 15 Line 451-458:

"In general, the modelling uncertainties may be caused by many factors, such as model structure, model parameterization and parameters, and meteorological forcing data. The underestimation of surface SM noted for the Noah, CLSM and MERRA2 can be related to fact that the impact of organic matter on soil hydraulic parameters is ignored (Yi et al., 2011; Chen et al., 2013; Zheng et al., 2015a). The better performance of ERA5 can be associated with the better estimation of precipitation and assimilation of ASCAT SM product (Shi et al., 2021; Hersbach et al., 2020). In the cold season, the Noah SM product generally captures well the SM_{ups} variations at surface layer (i.e., 5 cm) but overestimates the SM_{ups} at deeper layers (e.g., 20 and 40 cm), and overestimations

are also found for the MERRA2 products at all the depth. The overestimation can be related to the inappropriate parameterization of soil freezing characteristics as shown in Fig. 7 (Zheng et al., 2017)."

On Page 15-16 Line 467-471:

"Similar findings have recently been reported by Ma et al. (2021). The underestimation can be due to the i) underestimation of downward shortwave or longwave radiation (Chen et al., 2011; Wang et al., 2016), ii) inappropriate parameterization of diurnally varying roughness length for heat transfer (Chen et al., 2011; Zheng et al., 2015b; Reichle et al., 2017), and iii) overlook of the impact of organic matter on soil thermal parameters (Zheng et al., 2015b)."

On Page 16 Line 487-488:

"This can be related to the underestimation of ST noted for all the model-based products."

On Page 16-17 Line 500-507:

"As in the Maqu network, the better performance of ERA5 can be associated with the better estimation of precipitation and assimilation of ASCAT SM product (Hersbach et al., 2020; Shi et al., 2021). The overestimation noted to other products can be associated with the overestimations of precipitation (Yang et al., 2020) and uncertainty of soil texture and thus overestimation of soil porosity (Su et al., 2013; Shangguan et al., 2013; Bi et al., 2016). Both the Noah and MERRA2 products also overestimate the SM_{ups} of 5 cm in the cold season, which is related to the inappropriate parameterization of soil freezing characteristics as shown in Fig. 10. For the 20 and 40 cm deeper depths, all the products systematically overestimate the SM_{ups} due to uncertainty of soil texture, among which the ERA5 product shows the lowest bias while the VIC product presents the largest bias."

On Page 17 Line 518-519:

"The reason for the underestimation can be the same as the Maqu network."

On Page 17 Line 531-533:

"Similar as the Maqu network, all the model-based products tend to produce earlier onset of freezing and later onset of thawing at every depth due to the underestimation of ST, leading to underestimation of FSD and overestimation of TED and thus longer F/T duration in comparison to the upscaled dataset."

2. **To match the depths of in situ SMST measurements, the model-based SMST data were resampled across the vertical soil profile using the linear interpolation method. Did the authors find the linear interpolation is the best choice after testing or just follows the same procedure in previous studies? Can the uncertainty of this linear interpolation be quantified?**

Thanks for the comments. Linear interpolation and depth-weighted methods were widely used to resample the SMST data across the vertical soil profile in the previous studies. We tested both methods and found the results were similar (see Figs. R1-R2 shown below). In order to make full use of the valid in-situ measurements, we adopted linear interpolation method in this study. The corresponding text have also been added in revised version on Page 9 Line 263-266:

"To match the depths of in-situ SMST measurements, we compared the linear interpolation method and the depth-weighted interpolation method that are widely used to resample the SMST data across the vertical soil profile in the previous studies (Gao et al., 2017), and the results were found to be comparable to each other (Fig not shown). To make full use of the valid in-situ measurements, the linear interpolation method was thus adopted in this study."

Take the GLDAS-2.1 Noah SM product as the example, the calculating process and results of the two methods are shown as below.

- Depth-weighted method

The SMST derived from the GLDAS-2.1 Noah product for the layer 1 (0-10 cm), layer 2 (0-40 cm), and layer 3 (0-80 cm) are calculated as follows:

$$X_{0-10,Noah} = X_{0-10,Noah}$$

$$X_{0-40,Noah} \approx X_{0-10,Noah} * \frac{10}{40} + X_{10-40,Noah} * \frac{30}{40}$$

$$X_{0-80,Noah} \approx X_{0-10,Noah} * \frac{10}{80} + X_{10-40,Noah} * \frac{30}{80} + X_{40-100,Noah} * \frac{40}{80}$$
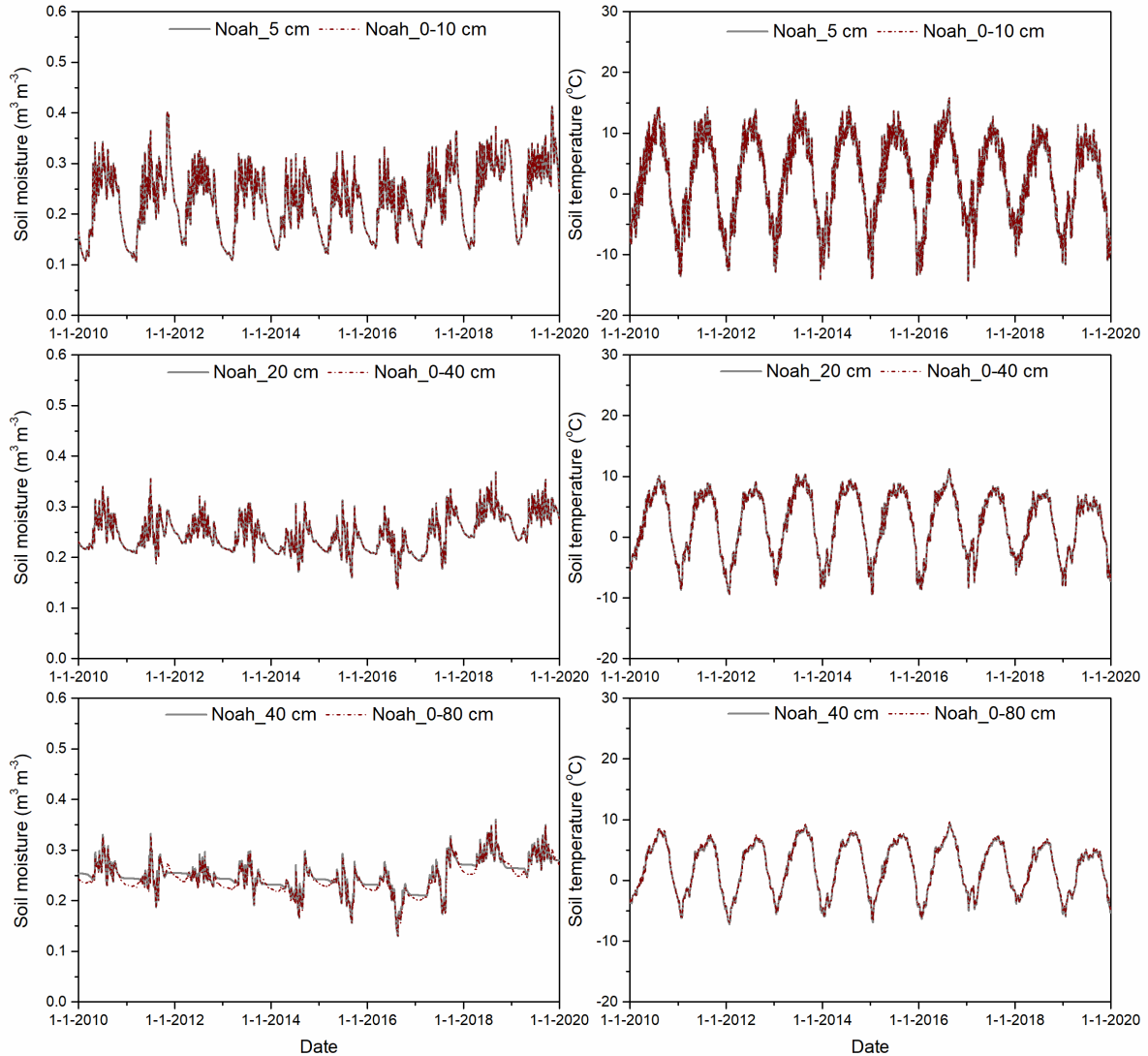
- Linear interpolation method

The SMST derived from the GLDAS-2.1 Noah product for the depths of 5, 20, and 40 cm are calculated as:

$$X_{5,Noah} \approx X_{0-10,Noah}$$

$$X_{20,Noah} \approx X_{0-10,Noah} + (X_{10-40,Noah} - X_{0-10,Noah}) * (20 - 5)/(25 - 5)$$

$$X_{40,Noah} \approx X_{10-40,Noah} + (X_{40-100,Noah} - X_{10-40,Noah}) * (40 - 25)/(70 - 25)$$

Figure R1: Time series of daily SM and ST derived from GLDAS-2.1 Noah product at soil depths of 5, 20, and 40 cm and in soil layers of 0-10, 0-40, and 0-80 cm from January 2010 to December 2019 for the Maqu network.
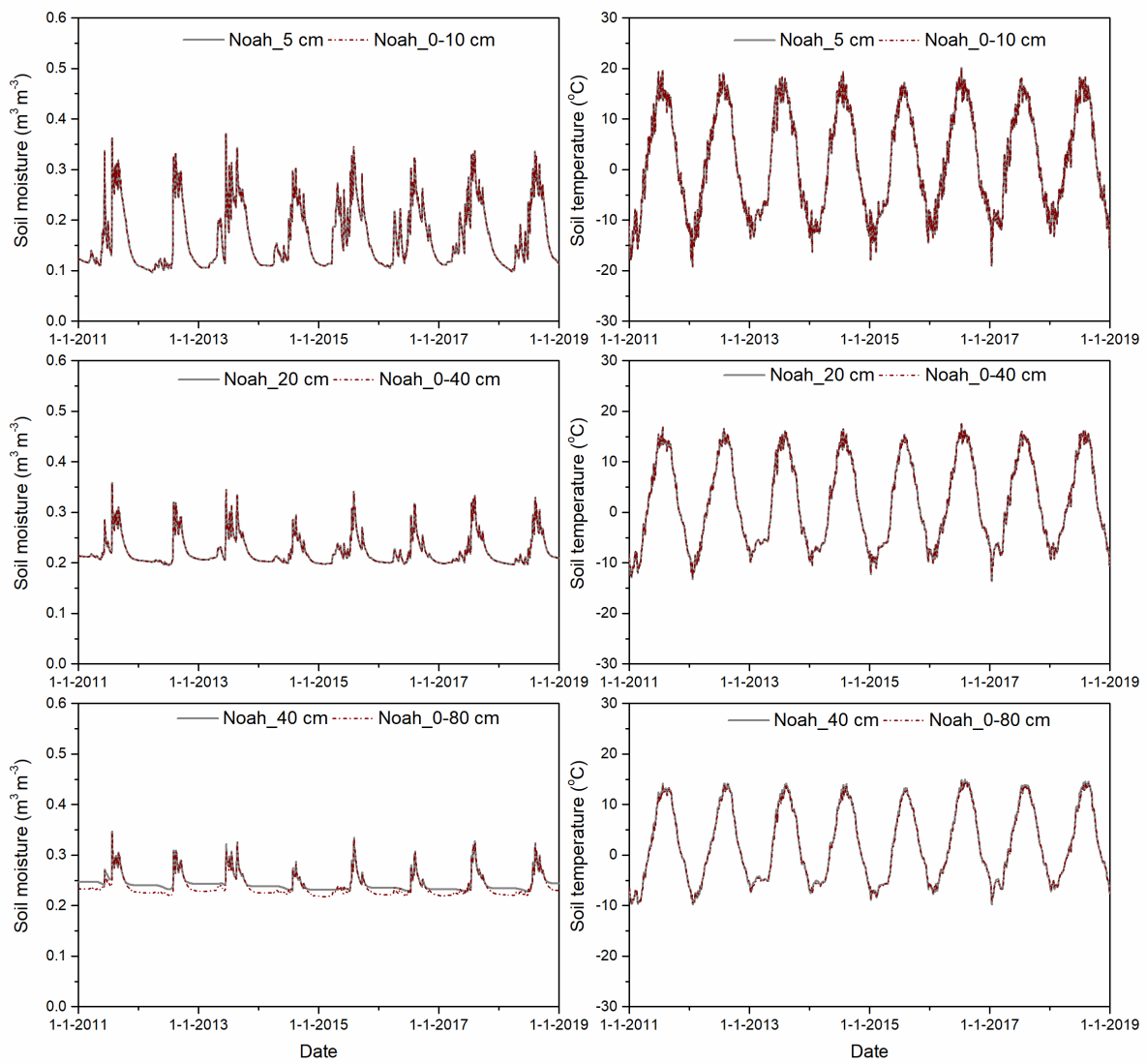
Figure R2. Same as Figure R1 but for the Shiquanhe network from January 2011 to December 2018.

Specific comments:

1.   Line 31: it is better to clarify the names of the reanalysis datasets.

Thanks for the suggestion. The names of the reanalysis datasets have been added in revised version on Page 1 Line 31:

"namely ERA5, MERRA2, GLDAS-2.1 CLSM, Noah, and VIC"

2.   Line 63-64: which climate data and land cover data did you use here?

Thanks for the comment. The climate classification can be found from Beck et al. (2018), and the land cover data was collected through the field work (Su et al., 2011; Zhang et al., 2021), the references have been added in revised version on Page 3 Line 64-66:

"which are respectively located in the cold humid area with cold dry winter and rainy summer covered by grassland, the cold semiarid area dominated by tundra, and the cold arid area dominated by desert (Su et al., 2011; Beck et al., 2018; Zhang et al., 2021)."

**3.   Section 2.2: please give the data portal (where you downloaded the data) here.**

Thanks for the suggestion. The meteorological data portal has been added in revised version on Page 6 Line 167-169:

"Precipitation and air temperature used in this study for the Maqu and Shiquanhe networks are obtained from the meteorological dataset provided by the China Meteorological Administration (http://data.cma.cn/en/?r=data/detail&dataCode=A.0012.0001, last access: 9th September 2022)."

4.   Equation (1): did you conduct regional average vs. regional average validation (that means some averaged model grids did not include in situ data but their values still involved in the spatial averaging)?

Thanks for the comment. Indeed, we conduct regional average vs. regional average validation. Equation (1) is used to produce the regional-scale SMST based on in-situ measurements, and all the model grids falling into the network area are averaged to obtain the corresponding values from the model-based products. Relevant description is provided in revised version on Page 8 Line 259-262:

"All the model grids falling into the scope of in-situ network are extracted from each product. Afterwards, the native grids of each product are downscaled to 0.25°*0.25° sub-grid cells using a bilinear interpolation. Subsequently, the SMST data in all the sub-grid cells falling into the scope of in-situ network are averaged to match the upscaled in-situ SMST data that represent the regional-scale mean values of in-situ network (see Fig. B1)."
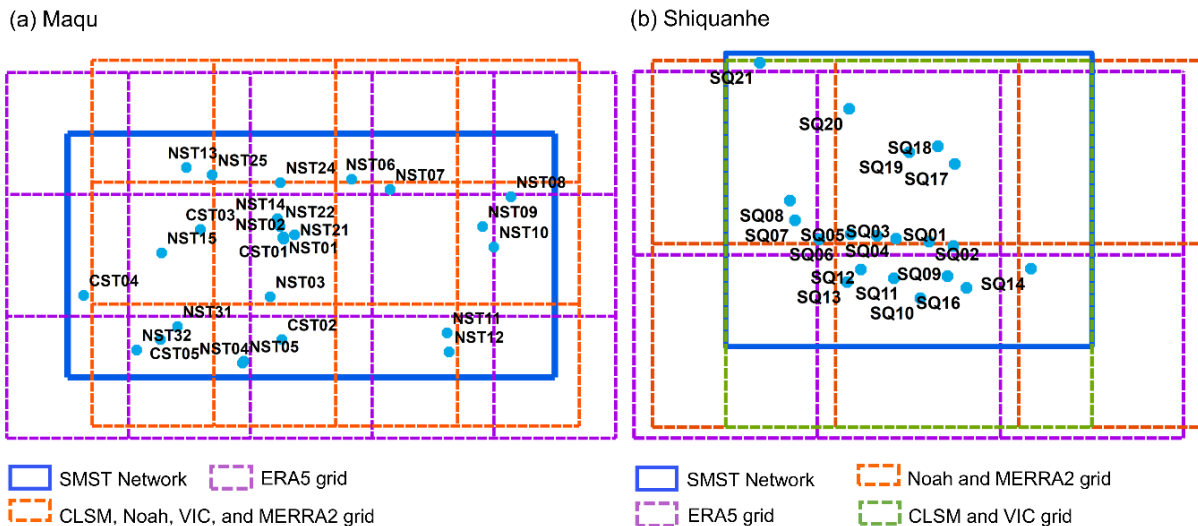
Figure B1: Grids of the model-based products falling into the (a) Maqu and (b) Shiquanhe network areas (denoted by the colourful dashed rectangles).

5. **Line 247: did you already consider the difference of coordinated Universal Time (UTC, and note models provide data at UTC) and local solar time (LST)? BTW, did the Tibet-OBS provides data in UTC or LST?**

Thanks for the comment. The Tibet-Obs provides data at Beijing time, we have already considered the difference between UTC and Beijing time (= UTC + 8).

6. **Line 277-279: did you average the monthly mean values for warm season and cold season per year (i.e., for every grid, one value per year for each season)? If not, how did you distinguish the inter-annual and inter-month influences? Moreover, what do you mean by "and all missing data points are assigned an equal value smaller than existed valid data points", why did you do such step?**

Thanks for the comments. Indeed, we use the monthly mean values for every year to calculate the trend, and the warm season spans from May to October while the cold season includes the rest of the months. Please find the details in the Appendix C.

"All missing data points are assigned an equal value smaller than existed valid data points" is one step of the Mann-Kendall trend test. We referred to the book named "Statistical Methods for Environmental Pollution Monitoring" by Gilbert Richland. (1987) Section 16.4 on Page 221:

"This procedure is particularly useful since missing values are allowed and the data need not conform to any particular distribution. Also, data reported as trace or less than the detection limit can be used (if it is acceptable in the context of the population being sampled) by assigning them a common value that is smaller than the smallest measured value in the data set."

The reference has been added in revised version on Page 10 Line 284:

"The Mann Kendall trend test reported by Richland (1987) is used in this study to determine whether a trend is presented within the long-term SMST time series derived either from the upscaled in-situ measurements or from the model-based products."

7. **Line 294: should be "rainfall" rather "snowfall"?**

Thanks for the comment. The "snowfall" is changed to "precipitation" in revised version on Page 11 Line 336.

8. **Line 300: change "diminish" to "diminishes".**

Thanks for the comment. The "diminish" is changed to "diminishes" in revised version on Page 12 Line 343.

9. **Line 319-339, and 365-373, Figs. 4 and 7: it is nice to see the authors investigated the trend of SM and ST in the TP as well as that of some meteorological data. Are there any conclusions about the influencing factors of SM and ST changes in the TP? Are there any similarities and differences in your results (e.g., trend analysis) compared to some similar studies conducted in TP, e.g., Shi et al. (2021, doi: 10.1175/JHM-D-21-0077.1)? What are the possible reasons for these differences? Some discussion of this issue will be helpful for the readability of these parts.**

Thanks for the comments and suggestions. We have provided additional discussion to the trend analysis results compared to the previous study in revised version as shown below.

On Page 12 Line 371-373:

"The full year trend analysis results are consist with the results reported by Shi et al. (2021) using the ESA CCI SM product, since the precipitation is the dominant drive of SM variation which shows significant negative trend in the humid area on the TP."

On Page 14 Line 420-423:

"The result is slightly different from Shi et al. (2021) that might be attributed to the different time span. Nevertheless, it is in agreement with the conclusion of spatial-temporal trend changes of surface SM generally decreasing from southeast to northwest over the TP comparing to the trend analysis result of Maqu network area."

10. **Line 326 and elsewhere: add unit for Sen's slope values, if applicable.**

Thanks for the comment. The unit for Sen's slope values has been added in revised version.

On Page 12 Line 370:

"the Sen's slopes of -0.004 (m^3 m^{-3}/yr) and -0.002 (m^3 m^{-3}/yr), respectively"

On Page 13 Line 375:

"with a Sen's slope of -0.08 (℃/yr)"

On Page 14 Line 420:

"with a Sen's slope of 0.001 (m^3 m^{-3}/yr)"

11. **Line 417-419: was only liquid soil water content in frozen soil simulated by the GLDAS Noah and MERRA2? Any references to illustrate this issue?**

Thanks for the comment. The references has been added in revised version on Page 15 Line 439-440:

"and the MERRA2 and Noah SM products can provide liquid soil water content (Gelaro et al., 2017; Zheng et al., 2017). "

The information of MERRA2 SM data can also refer to https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/FAQ/# "FAQ1: WHAT ARE THE MERRA-2 SOIL MOISTURE VARIABLES AND THEIR UNITS?"

12. **Table 4 and elsewhere (e.g., figures, and texts): add unit for Bias, RMSD, and ubRMSD.**

Thanks for the comment. The units for Bias, RMSD, and ubRMSD have been added in Tables 5 and 6 in revised version.

13. Fig. 3(a): any possible reasons that the SM at 40 cm is the lowest (lower than that at 80 cm)?

Thanks for the comment. A possible reason has been given in revised version on Page 12 Line 338-340:

"The soil layers below 20 cm are dryer than the upper layers in the warm season, whereas the soil at depth of 80 cm is wetter than 40 cm that might be attributed to absence of evapotranspiration and existence of shallow groundwater (Li et al., 2021)."

14. Figs. 5 and 8: the abscissa ST is negative, and the reader may be confused that how did you obtain the thawing results at such situation. Try to make them clearer if possible.

Thanks for the comment. We have added an explanation of freezing and thawing period in revised version on Page 13 Line 379-381 to make it clearer:

"The freezing period defined in this study spans from the first date of ST falling below zero to the date of lowest ST occur, whereby the SM value is generally decreasing in this period. Later on the thawing period starts and ends when the ST rise above zero, whereby the SM value is increasing during this period."

We would like to thank the reviewer for carefully reading our manuscript and providing detailed and constructive comments. This has helped us to improve our manuscript significantly. In the text below we provide our response to each comment point by point.

Reviewer's comments are in **bold**.

Author's responses are in regular.

Author's additions/modifications in the text are in blue.

1.   Page 2 Line 58: The authors cited this reference for at least 10 times. How many similarities are there between the cited paper and this paper?

Thanks for the comment. The similarities between the cited paper and this paper are the adopted methods including spatial upscaling and trend analysis. In addition, the datasets presented in both papers are produced based on the Tibet-Obs measurements. The novelty of this paper in comparison to the cited paper is provided on Page 3 Line 81-85:

"In this paper, we present a long-term (~10 years) SMST profile dataset collected from the Tibet-Obs, which expands the surface SM dataset introduced by Zhang et al. (2021) to include both SM and ST measurements collected at multiple depths. As such, analysis of freezing and thawing characteristics become possible. The analysis of seasonal dynamics and trend changes as well as validation of model-based products are also extended to multiple depths for an approximately 10-year period. In addition, more model-based products are evaluated in this paper."

2.   Page 2 Line 64: What is short grasses?

Thanks for the comment. The "short grasses" has been changed to "grassland" in revised version that is well known by readers.

On Page 1 Line 18:

"located in the cold humid area covered by grassland"

On Page 2 Line 65:

"in the cold humid area with cold dry winter and rainy summer covered by grassland"

On Page 4 Line 117:

"with a land cover dominated by grassland"

On Page 18 Line 548:

"in the cold humid area covered by grassland"

3.   Page 4 Line 112: There should have a chart to clearly summarize the basic properties of each networks.

Thanks for the comment. A new table, Table 2 in the revised version is added to provide the basic information of the Tibet-Obs networks.

Table 2. Information of the Tibet-Obs networks

Networks	Climate zone	Land cover	Altitude (m)	Annual Precipitation (mm)	Monitoring sites
Maqu	Cold humid	Grassland	3400-3800	600	26

Shiquanhe	Cold arid	Desert	4200-4700	100	20
Ali					4
Naqu	Cold semiarid	Tundra	Around 4500	400	11

4.   **Page 5 Line 131: What's the meaning of a.s.l. ?**

Thanks for the comment, it was mentioned on Page 4 Line 117-118: "above sea level (a.s.l)".

5.   **There are numerous of model-based SMST products. Why do you select these ones?**

Thanks for the comment. The reason to select the five model-based products is provided in revised version on Page 6 Line 178-179:

"The reason to select these products is due to the fact that they are more widely adopted and extensively assessed."

6.   **There are many upscaling methods. Why do you use the classic arithmetic averaging approach other than elaborate methods (such as DISPATCH, machine learning, and assimilation). Please explain it.**

Thanks for the comment. The reason to use the classic arithmetic averaging approach is provided in revised version on Page 8 Line 219-223:

"Zhang et al. (2021) demonstrated the better performance of the arithmetic averaging approach in upscaling the surface SM of the Tibet-Obs network in comparison to the voronoi diagrams, time stability, and apparent thermal inertia methods that are widely adopted in existing literatures (Qin et al., 2015; Colliander et al., 2017). Therefore, the arithmetic averaging approach is also adopted in this study to obtain the regional-scale SMST profile data for Maqu and Shiquanhe."

7.   **What's the difference between the soil moisture derived from this paper and Zhang et al. (2021)?**

Thanks for the comment. In Zhang et al. (2021), only the surface SM dataset is presented. In this paper, both profile SM and ST data are presented. Relevant description is provided on Page 3 Line 81-83:

"In this paper, we present a long-term (~10 years) SMST profile dataset collected from the Tibet-Obs, which expands the surface SM dataset introduced by Zhang et al. (2021) to include both SM and ST measurements collected at multiple depths. As such, analysis of freezing and thawing characteristics become possible."

8.   **Page 8 Line 245: How to convert the unit?**

Thanks for the comment. The equation of unit convert has been added in revised version on Page 9 Line 249-253:

"The units of SM data from the GLDAS-2.1 CLSM, Noah, and VIC products is converted from "kg m^{-2}" to "m^3 m^{-3}" following Eq. (2), and the units for the ERA5 and MERRA2 SM data is already with "m^3 m^{-3}".

$$SM = SWC/(L * \rho_{H_2O}) \tag{2}$$

where SWC represents the soil water content (kg m^{-2}), L (m) represents the layer thickness, ρ_{H_2O} represents the soil water density (kg m^{-3})."

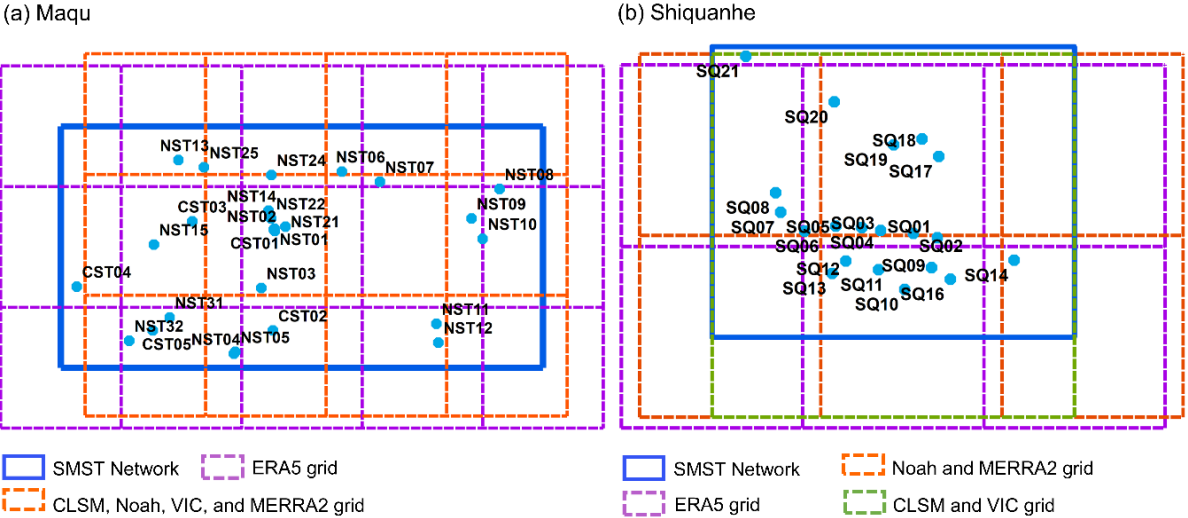**9.**   **The reliability of upscaled datasets should be verified at first before trend analysis.**

Thanks for the suggestion. We have adjusted the order of Section 4.1 and Section 4.2 in revised version, the corresponding text has been revised on Page 10 Line 293-297:

"Section 4.1 gives the uncertainty analysis results for the upscaled SMST profile data. Section 4.2 presents the upscaled SMST profile data for the Maqu and Shiquanhe networks spanning the 10-year period from 2010 to 2019 (see Section 3.1), as well as the analysis results for the SMST seasonal dynamics, trend test, detection of F/T state and soil freezing characteristics at different depths. Application of the upscaled data to evaluate the performance of model-based products is presented in Section 4.3 to demonstrate its suitability for the evaluation of readily available SMST profile products."

10.   To prove the uniqueness and superiority of the upscaled dataset for validation, there should have a comparison between the accuracy degrees achieved from the validation against the upscaled and the original dataset, respectively.

Thanks for your comment. The uniqueness of the SMST dataset from the Tibet-Obs is in terms of its long-term period and multiple soil depths on the Tibetan Plateau. The SMST data used in validating the model-based products is the upscaled dataset, which can generally represent the regional value of monitoring network (i.e. the Maqu and Shiquanhe network areas as shown in Figure B1). Correspondingly, all the model grids falling into the network area are averaged to obtain the regional values as well. This is a common method widely adopted to validate the model-based products to reduce the uncertainty related to the spatial variability of SM. In other words, both of the upscaled in-situ SMST data and averaged model-based SMST data represent the regional condition of network, while the original data is at point-scale for each monitoring site that may exist significant uncertainty for assessment. It is thus pointless to make a comparison between point data and the regional data.

Figure B1: Grids of the model-based products falling into the (a) Maqu and (b) Shiquanhe network areas (denoted by the colourful dashed rectangles).