

Dear Editor,

Thank you very much for providing detailed comments regarding to the manuscript and dataset. The data publication and manuscript have been revised based on the comments.

Please find below our detailed response to each comment.

We think that the revised manuscript and data publication have appropriately addressed all the 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

In the text below we provide our response to each comment point by point.

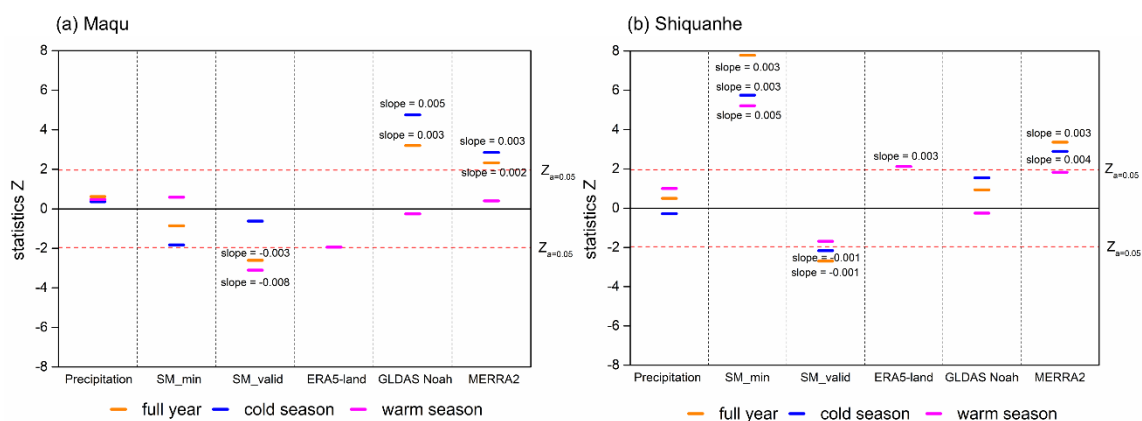
Editor's comments are in **bold**.

Author's responses are in regular.

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

**The description of the hourly data still lacks metadata, additionally it is not clear which data exactly was used for the analysis. Could you please update the User guide in your repository and describe the data acquisition procedure in similar detail as you did for the soil moisture products (how to choose the station you used, which time frame, etc...).**

Thanks for the comments and suggestions. The precipitation data were mainly used for the trend analysis by the Mann Kendall trend test and Sen's slope estimate (see Section 3.2) based on the monthly cumulative values, and the results are shown in Fig. 9 of Section 4.2, while daily value is merely plotted in Fig. 8 as a reference for the soil moisture variation.



**Fig. 9. Mann Kendall trend test and Sen's slope estimate for precipitation,  $SM_{AA-min}$ ,  $SM_{AA-valid}$ , and model-based SM derived from the ERA5-land, GLDAS Noah, and MERRA2 for a 10-year period for (a) Maqu and (b) Shiquanhe networks.**

The precipitation data part in user guide has been updated in the revised version. It should be pointed out that the precipitation data is not part of the upscaling soil moisture dataset, the objective of providing precipitation in this study is for the reference against the soil moisture seasonal dynamic and trend changes.

#### “4.4 Precipitation data

The precipitation data is from two weather stations, i.e. Maqu (34°00'N, 102°05'E) and Shiquanhe (32°30'N, 80°05'E), operated by the China Meteorological Administration (CMA) which provides the near-surface meteorological data of about 700 weather stations in China. The daily precipitation data can be downloaded from [https://data.cma.cn/data/detail/dataCode/SURF\\_CLI\\_CHN\\_MUL\\_DAY.html](https://data.cma.cn/data/detail/dataCode/SURF_CLI_CHN_MUL_DAY.html). The monthly precipitation data that actually used in this study is contained in the “Supplementary data” folder.

Downloading precipitation data through CMDC (China Meteorological Data Service) website interface:

1. Register an CMA account.
2. Go to the [https://data.cma.cn/data/detail/dataCode/SURF\\_CLI\\_CHN\\_MUL\\_DAY.html](https://data.cma.cn/data/detail/dataCode/SURF_CLI_CHN_MUL_DAY.html). The language of this web is Chinese.
3. The interface include the following parts:
  - **Product Summary:** It gives a description of the dataset (e.g. name, keywords, begin time, end time, frequency, share level, quality, etc.).
  - **Retrieve:** It is a retrieve web form for your require data.
  - **Metadata:** It gives the metadata of the dataset.
  - **Documentation:** It provides links to details documentation about the dataset.
4. Click the button of **Retrieve** to select the required data.
  - **TimeScope:** 2009.5.1-2019.4.31 (2010.8.1-2019.7.31)
  - **Select Station:** Select Station list, Gansu, [56074] Maqu (Select Station list, Qianghai, [55228] Shiquanhe)
  - **Element:** 20-20h precipitation
5. Click the bottom of **Retrieve** to submit the required data. The data is only available to agreement users.”

**Manuscript section 3.2: You wrote in your response that the trend calculation for the precipitation data was done similarly to the soil moisture data. However, in the manuscript, there is no mention whatsoever about precipitation data in section 3.2. Please revise this so that it is clear that you did the trend analysis for both soil moisture and precipitation. An obvious way would be to describe the equations in general terms (x instead of SM as a variable) and then describe for which variables you did the analysis.**

Thanks for the comments and suggestions. The corresponding parts in Section 3.2 are revised:

Page 7 Line 220-222:

“The Mann-Kendall test and Sen’s slope estimate (Gilbert, 1987; Mann, 1945; Smith et al., 2012) are adopted to analyze the trend of the 10-year time series for the upscaled SM, model-based SM products (i.e. ERA5-land, GLDAS Noah, and MERRA2), and precipitation.”

Page 7 Line 226-230:

“For month  $i$  (e.g. January), the statistics  $S_i$  can be computed as:

$$S_i = \sum_{k=1}^9 \sum_{l=k+1}^{10} \text{sgn}(X_{i,l} - X_{i,k}) \quad (2a)$$

$$\text{sgn}(X_{i,l} - X_{i,k}) = \begin{cases} 1 & X_{i,l} > X_{i,k} \\ 0 & X_{i,l} = X_{i,k} \\ -1 & X_{i,l} < X_{i,k} \end{cases}$$

where  $k$  and  $l$  represent the different year and  $l > k$ ,  $X_{i,l}$  and  $X_{i,k}$  represent the monthly value of the variable for the month  $i$  of the year  $k$  and  $l$ , respectively.”

Page 8 Line 247-249:

“If the trend shows upward or downward, we will further estimate the slope (change per unit time) with Sen’s method (Sen, 1968). The slopes of each month can be calculated as:

$$Q_i = \frac{X_{i,l} - X_{i,k}}{l - k} \quad (2f)''$$

**Monthly precipitation data: It is a good idea to include a slightly aggregated data product if you can't supply the original data that you used in the analysis. However, it is not clear if monthly data was used at all and where, or if it is simply provided as an informative dataset because the daily values are not accessible. Please make this very clear in the text.**

Thanks for the comments and suggestions. The description for the utilization of the monthly precipitation data application in this study is revised on Page 5 Line 170-171:

“The monthly precipitation data for the period between 2009 and 2019 is used for the trend analysis by Mann Kendall trend test and Sen’s slope estimate in this study (see Section 4.2).”

**Furthermore, there is only very little metadata about this dataset except the coordinates. What kind of instruments were used (maybe there's a standard in the CMA, but it would be good to have more information, similar to the level of detail you provide for soil moisture)?**

Thanks for the comments and suggestions. The specific information of the instruments are not provided from the official website, and all available information for the weather station is included to Section 2.2 in the revised manuscript on Page 5 Line 159-165:

“Precipitation data is available from the dataset of daily climate data from Chinese surface meteorological stations. This dataset is maintained by the China Meteorological Administration (CMA) and based on the measurements from 756 basic and reference surface meteorological observation and automatic weather stations (AWS) in China from 1951 to present. The online dataset mainly includes seven meteorological variables such as air pressure, air temperature, relative humidity, wind speed, evaporation, sunshine duration, and precipitation. The precipitation data from two weather stations (see Fig. 1), i.e. Maqu (34°00'N, 102°05'E) and Shiquanhe (32°30'N, 80°05'E) are used in this study.”

**Additionally, it is not clear how you aggregated the data to monthly values. It seems like they are averages of each month (because the numbers are too low compared to your annual precipitation amounts in section 2.1), which would be very unusual for precipitation data. In any case, you should explain why you have done monthly values, give sufficient metadata (comparable to the soil moisture metadata), describe what kind of aggregation you used and why, and where you use this data.**

The monthly value is monthly cumulative precipitation that is the sum of the daily precipitation of each month. It has been updated in the dataset, and please find the details in the response to previous comment.

The corresponding description is also included in the revised manuscript on Page 5 Line 165-171:

“The available daily precipitation is the cumulative value for the period between 20h of previous day and 20h of current day at Beijing time, which is available from [https://data.cma.cn/data/detail/dataCode/SURF\\_CLI\\_CHN\\_MUL\\_DAY.html](https://data.cma.cn/data/detail/dataCode/SURF_CLI_CHN_MUL_DAY.html) (last access 11 March 2021). The daily precipitation is summed up for each month to obtain the monthly cumulative value in this study, which can be found at <https://doi.org/10.4121/uuid:21220b23-ff36-4ca9->

a08f-ccd53782e834 (last access 16 April 2021). The monthly precipitation data for the period between 2009 and 2019 is mainly used in this study for the trend analysis (see Section 4.2).”

The reason for using the monthly precipitation value to conduct trend analysis is described in Section 3.2:

“The Mann-Kendall test and Sen’s slope estimate (Gilbert, 1987; Mann, 1945; Smith et al., 2012) are adopted to analyze the trend of the 10-year time series for the upscaled SM, model-based SM products (i.e. ERA5-land, GLDAS Noah, and MERRA2), and precipitation. Specifically, the trend analysis is based on the monthly data, and all the missing data is regarded as an equal value smaller than other valid data. The test consists of calculating the seasonal statistics  $S$  and its variance  $VAR(S)$  separately for each month during the 10-year period, and the seasonal statistics are then summed to obtain the  $Z$  metric.”

**Also, while skimming through the manuscript again, I noticed a number of typos and phrasing issues. Please consult an English native speaker to check for those so that the manuscript can be read easily.**

Thanks for the suggestion. The revised manuscript has been checked by the co-author Rogier van der Vdelde who have stayed in the USA for about three years. We think the typos and phrasing issues has been addressed.