Reviewer #1 Comment 1: (hereafter referred to as R1C1, R1C2…) *Based on data fusion and numerical modeling, the authors reconstructed a long-term (1981-2017) 0.1° daily dataset of total precipitation (P), liquid rainfall (Rain), snowfall (Snow), snow water equivalent (SWE), snowmelt (Melt), and soil moisture (SM) in China. Reconstructing these hydrological components is a very challenging topic, particularly for west China including the Northwest and Tibetan Plateau. The reader may expect some progress in these regions when seeing the title, but the results of this study show limited progress in these regions. This is certainly not surprising, and it looks like there is a long way to go. Nevertheless, the authors took advantage of two favorable conditions to extend the recent data with better accuracy to the 1980s, which is a clear progress. That is, (1) meteorological data and satellite data in the last decade or so are more abundant, higher resolution and more accurate; (2) Satellite remote sensing data can be used to calibrate a land hydrological model. They are the innovative points of this paper and support its publication.*

A: Thank you for your constructive comments. We have carefully considered your suggestions for the revision of the manuscript. We would like to reaffirm the contribution of this study. Recent state-of-the-art satellite-based data have high spatial resolutions and abundant observational information with full spatial coverage. The objective of this study is to extend the satellite-based data of P, SWE and SM to the 1980s in China so that hydrological studies focusing on a longer time span can take advantage of those high-quality data. The innovations you summarized match the contribution of this study except for one point, i.e., we do not aim to create a dataset with "better accuracy" compared with the satellite-based data used for model training or calibration. The satellite-based data are the benchmarks of the reconstruction dataset we produced. So the accuracy of the reconstruction dataset is not able to exceed that of the satellite-based data.

According to the objective of the study we pinpoint above, we do make some progress in the Northwest and Tibetan Plateau but not in the aspect you expected. Improvements are made in those regions not for the latest years but for the historical period when multiple-sources satellite-based observations are not available. It is challenging to comprehensively evaluate any new dataset in the Northwest and Tibetan Plateau since the "ground truth" information, i.e., observations from ground-based stations, are very limited in those regions.

My main comments are as follows.

R1C2: *Research focus. From the results, the reconstructed P data are of high quality and may be used by other researchers. However, the significance of the snow data and soil moisture content data is weaker. Especially for snow data, this paper uses the data of Che et al. as training data and the evaluation is also based on the data and there is no validation based on independent data. Given that the time series of Che et al. data is even longer (1979-2020), the snow data presented in this study is not very necessary. I suggest that the snow-related section be greatly weakened, and its validation section could even be removed.*

A: Thank you very much for your critical comment. We agree with you that the significance of the reconstructed *P*, SWE and SM data varies based on the quality. However, the reconstructed SWE and SM data have their unique values. As we stated in the introduction section, existing long-term high-resolution SWE and SM data in China are the byproduct of runoff modeling without direct calibration. Therefore, this study takes a step further to calibrate hydrological models with satellite-based data so that the reconstruction data become the historical extension of the recent satellite-based data. The reconstructed SM data are accurate with the median KGE=0.61 across the country at the daily scale (Table 6). Although the reconstructed SWE data fail to match the satellite-based snow data in the Tibetan Plateau, they are relatively accurate in the Songhua and Liaohe River Basin with median KGE=0.55 across the basin at the daily scale (Table 7). In addition, the reconstructed SWE data have certain skills in matching the satellite-based SWE in northern Xinjiang (Figure 7) which is another snow hydrology hotspot region.

The SD-CN (Che and Dai, 2015) dataset is indeed long enough to cover the reconstruction period of this study. However, SD-CN has a coarse spatial resolution (25km), which does not satisfy the 0.1° resolution we intend to reach. Therefore, we prefer to keep the validation section of snow data.

R1C3: *The abstract needs to describe that (1) the constructed P is used to drive HBV, which generates snow and SM data, and (2) the reconstruction algorithm and data evaluation use the same data sources, only the data periods are different (if my understanding is correct). The current description will make the reader struggle to find the difference between the training data and validation data.*

A: Thank you for your important comment. Yes, for reconstructed *P* data, the reconstruction algorithm and data evaluation use the same data sources but with different periods. We will adopt those two suggestions you proposed in the revised manuscript.

R1C4: *Data description. (1) The use of CMPA_1km is not clearly stated. The authors just say upscale to 0.1 degree, without stating how the 1km-resolution data is used. (2) It should be made clear that the data in Table 2 are both for training and for validation. If this is not made clear, it is difficult to understand the structure of Figure 1. This once made the reviewer confused. (3) The data introduction section suddenly mentions air temperature and net radiation without introducing the usage of these data. This may confuse the reader.*

A: Thank you very much for your valuable comment. (1) The CMPA_1km data in 2015-2017 were upscaled to 0.1 degree to act as the temporal extension of the original CMPA data in 2008-2014. We will add the detailed use of CMPA_1km in the revised manuscript. (2) We agree. We will clarify that the CMPA precipitation data are used both for training and validation in the revised manuscript. (3) The temperature and net radiation are used to drive HBV model. As we stated in Line 191-193 "PET was calculated using Priestley-Taylor equation (Priestley and Taylor, 1972) with interpolated T and daily-aggregated ERA5-land Rn." We will add the usage of temperature and net radiation data in the revised manuscript.

R1C5: *Methodological aspects. (1) It is not clear why the HBV model is used, and what are the advantages in reflecting SM and SWE. (2) It needs to be clarified what exactly is meant by 5-fold cross validation, e.g. whether this fold is temporal or spatial, which time periods (or spaces) are used for training and which time periods (or spaces) for validation. This information must be clearly, perhaps in Table 2.*

A: Thank you very much for your comments. (1) The HBV model is used because of its low computational complexity and general applicability in various climate conditions (Beck et al., 2020; Seibert and Bergström, 2022). In addition, HBV models show high capability in simulating soil moisture in various regions of the world (Beck et al., 2021). We will add the advantages of HBV in the revised manuscript. (2) We performed a spatial and temporal mixed 5-fold cross-validation, i.e., we put the data of one hundred 0.1° grids in a 1° tile on all days together to form a training dataset and then split it into 5 folds randomly. Each sample corresponds to the data of one 0.1° grid on one day. We will clarify the strategy of cross-validation in the revised manuscript.

Beck, H. E., Pan, M., Lin, P., Seibert, J., van Dijk, A. I., & Wood, E. F.: Global fully distributed parameter regionalization based on observed streamflow from 4,229 headwater catchments, J Geophys Res-Atmos, 125(17), 10.1029/2019jd031485, 2020.

Beck, H. E., Pan, M., Miralles, D. G., Reichle, R. H., Dorigo, W. A., Hahn, S., ... & Wood, E. F.: Evaluation of 18 satellite-and model-based soil moisture products using in situ measurements from 826 sensors, Hydrol. Earth Syst. Sci., 25(1), 17-40, 10.5194/hess-25-17-2021, 2021.

Seibert, J., and Bergström, S.: A retrospective on hydrological catchment modelling based on half a century with the HBV model, Hydrol. Earth Syst. Sci., 26(5), 1371-1388, 10.5194/hess-26-1371-2022, 2022.

R1C6: *Validation issues. (1) The data sources used in the current evaluation are the same as those used for training, but validations based on independent data are more convincing. For example, precipitation data need to be validated at stations of pre-CMPA era and SM data need to be validated based on intensive SM observation networks (e.g., the widely used Maqu and Naqu soil moisture measuring networks). (2) Figure 6 shows that overall the reconstruction of P is good, but why are its errors significantly larger in several watersheds in 2017?.*

A: Thank you very much for your comment. (1) As we stated in R1C1, we do not aim to create a dataset with the best accuracy compared with all existing data products. Instead, we intend to extend the state-of-the-art but short-term data of P, SWE, and SM to the 1980s. Therefore, validations should present the similarity between the reconstruction data and the target data. Validations based on independent station-based data are out of the scope of this study. (2) In fact, a small discrepancy between reconstructed P and CMPA data occur since 2015 according to Figure 6. The reason is that the CMPA data used in this study are not perfectly consistent across time. As introduced in section 2.2, the CMPA data after 2015 come from the upscaling CMPA_1km data. Compared with CMPA, CMPA_1km has additional observational data sources including more automatic weather stations and radar estimations. Therefore, system bias may occur between CMPA and CMPA_1km.

Other comments:

R1C7: *The terms "ground-truth" or "raw observational data" or "observed SWE" are mentioned in the text, but please avoid using them in this way, because in fact they refer to fused data or remotely sensed data.*

A: Thank you very much. We will change the description into "target data" accordingly in the revised manuscript.

R1C8: *It is difficult to understand the Continental Basin, suggest to change to NW Continental Basin*

A: Thank you very much for your comment. The name "Continental Basin", also known as the "Inland River Basin", is often used in hydrological research. In addition, the name "Continental Basin" is used in the official boundary data of nine major basins in China (https://www.resdc.cn/data.aspx?DATAID=141). Therefore, we prefer to keep this name.

R1C9: *P3: "For P, we merged CGDPA and MSWEP to reconstruct the P from CMPA using machine learning techniques; for SM, we used the reconstructed P to drive a hydrological model to reconstruct SM from SMAP level 4". This description (reconstruct the P from CMPA, reconstruct SM from SMAP) is quite confusing.*

A: Thank you very much. In the revised manuscript, the sentence will be revised as "We merged CGDPA and MSWEP to reconstruct the *P* benchmarked by CMPA using machine learning techniques. We used the reconstructed *P* to drive a hydrological model to reconstruct SM calibrated by SMAP level 4."

R1C10: *P8: "For SM, the 1 m root zone SM …". Although I know what it refers to, for most readers it may not be clear that it is SMAP-L4.*

A: Thank you very much. We will specify the SM data as SMAP-L4 in the revised manuscript.

R1C11: *P10: Start a new paragraph from "The validation metric for SWE and SM is KGE in Eq. 3".*

A: Thank you very much. We will modify the paragraph in the revised manuscript.