

Reviewer #2 Comment 1: (hereafter referred to as R2C1, R2C2...) *This manuscript reconstructed a long-term precipitation data set using a stacking machine learning model according to MSWEP and CGDPA, and at the same time reconstructed long-term soil moisture and snow water equivalent data sets using the HBV model. The reconstructed data ranges from 1981-2017 and with a high spatial-temporal resolution, i.e. a 0.1 degree spatial resolution and daily temporal resolution. The results show that the data has a high reliability, especially the precipitation data. It provides valuable products across China for meteorological and hydrological domains, such hydrological modeling, assessing the hydrological response under climate change and etc. The manuscript is well-organized and well-written. It is suitable for Hydrology and Earth System Sciences. Therefore, I recommend a minor revision.*

A: Thank you for your constructive comments. We have carefully considered your suggestions for the revision of the manuscript.

Detailed comments:

R2C2: *1. The spatial data used in this manuscript has different units, namely degree and km. Especially, the variables with different units, including precipitation, soil moisture, snow cover, and snow depth, were used for calibrating and validating HBV model. It isn't clear which geographic and projected coordinate systems were chosen and which unit were used for calibrating and validating.*

A: Thank you very much for your critical comment. The HBV model is calibrated and validated at a 0.1° resolution under the WGS 84 latitude/longitude coordinate system (EPSG:4326). Therefore, the forcing, auxiliary, and validation data for HBV were pre-processed to be 0.1° under EPSG:4326. We mentioned the way to unify the spatial resolutions of different datasets in the original manuscript. For example, “The calibration targets were SCA, SWE, and SM, which should be pre-processed from the raw observational data before calibration. The 0.05° SCA data were first aggregated to be 0.1°.” (Line94-95); “For SWE, we first resampled the 25 km snow depth into 0.1° using bilinear interpolation.” (99-100); “For SM, the 1 m root zone SM was resampled from 9 km to 0.1° using bilinear interpolation and aggregated from 3-hourly to daily.” (Line 103-104). In the revised manuscript, we will specify the coordinate system we use and clarify the requirement of spatial resampling for different datasets before hydrological modeling.

R2C3: *2. Two types of precipitation, namely CMPA from 2008 to 2014 and CMPA_1km from 2015 to 2017, were used for machine learning model. The uncertainty due to the precipitation inconsistency should be discussed.*

A: Thank you for your important comment. Yes, the inconsistency of CMPA data over time may raise uncertainties. As the successor of CMPA, CMPA_1km retains most of the observation sources that are used in CMPA, including more than 30000 automatic weather stations and the CMORPH product. In addition, The main deriving methods of CMPA and CMPA_1km are similar, including optimal interpolation and local bias correction. Therefore, the difference between CMPA and CMPA_1km is not significant in theory. Considering that data sizes contribute a lot to the accuracy of machine learning predictions, it is reasonable to combine CMPA and CMPA_1k which can trade consistency for more data samples.

R2C4: *3. The abstract stated “the short-term 0.1o CMPA”, and I suggest clarifying time frame on the short-term.*

A: Thank you very much for your valuable comment. The “short-term” phrase refers to the period of CMPA (2008-2017) compared with the long period of the reconstruction data (1981-2017). We will specify the period instead of using “short-term” in the revised manuscript.

R2C5: *4. On the names of the nine major river basins of China, I suggest replacing the Southwest Basin with the Southwest Basins, and replacing the Southeast Basin with the Southeast Basins, since either of them includes more than one basin.*

A: Thank you very much for your comments. A large basin consists of many sub-basins. The names “Southwest Basin” and “Southeast Basin” are 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 these names.