Earth Syst. Sci. Data Discuss., referee comment RC2
https://doi.org/10.5194/essd-2020-361-RC2, 2021

Comment on essd-2020-361

Anonymous Referee #2

Referee comment on "1 km Monthly Precipitation and Temperatures Dataset for China from 1952 to 2019 based on a Brand-New and High-Quality Baseline Climatology Surface" by Haibo Gong et al., Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2020-361-RC2, 2021

Some major concerns on the manuscript titled "1 km Monthly Precipitation and Temperatures Dataset for China from 1952 to 2019 based on a Brand-New and High-Quality Baseline Climatology Surface" are listed below:

First, except for combining satellite-based precipitation and tempeature data, it is hard to tell the novelty of this study in terms of methodology for creating the 1-km monthly datasets, especially given there are already some datasets available for ecological, hydrological studies etc.

Second, the method for generating "ChinaClim_timeseries" is questionable or, to some degree, wrong. It is unthinkable that the authors used "ChinaClim_baseline" to obtain the anomaly time series. Why not use the 30-years mean normal from weather stations to derive the anomaly time series? The method described in Section 3.2 and Fig. 3 is incomprehensible.

Third, the evaluation on the accuracy of ChinaClim_baseline and ChinaClim_timeseries is also questionable. Because the weather stations used to evaluate ChinaClim_baseline and ChinaClim_timeseries are different from those used to evaluate the WorldClim2 and CHELSA, it is hard to infer that the quality of ChinaClim_baseline and ChinaClim_timeseries is better than those of compared datasets, respectively.

Fourth, the determination coefficient R2, MAE and RMSE are used to compare the accuracy of ChinaClim_baseline and ChinaClim_timeseries with those of other datasets. Even thought the values of R2 is slightly higher while the values of MAE and RMSE are slightly lower for the newly-created datasets, are the differences statistically significant? If not, it just suggests there are no significant differences between the newly-created datasets and those existing ones.

Fifth, the colors used for creating Figure 5, 7, 9, and 11 are bad. The divergent or sequential colors had better be used correctly to map the data. For example, red color is good for high values and blue color is good for low values.

Sixth, time-series of daily temperature and precipitation data are highly valuable for hydrological and ecological studies. From the current version of the newly-created monthly datasets, it is difficult to see the significance of the datasets, at least for hydrological studies.

Seventh, the method for creating ChinaClim_baseline is not very clear. The step 5 (on page 11), i.e.,"(5) Repeat steps 2 to 4 for 10 times, and final baseline climatology surface (ChinaClim_baseline) was created by averaging ten surfaces" means that the nine-folds weather station data used as training data will vary as the process repeats. For each repeat, did you evaluate the accuracy of model formulations for each month?

Eighth, is there any overfitting problem when creating the so-called brand-new datasets?