Reply to Reviewer 3

Manuscript ID: essd-2018-150
Title: WHU-SGCC: A novel approach for blending daily satellite (CHIRP) and precipitation observations over Jinsha River Basin
Journal: Earth System Science Data
Type: Article

Dear Reviewer,

Thank you for your insight comments and suggestions. We have modified the manuscript accordingly. We trust that all of your comments have been addressed accordingly in the revised manuscript. If you have further suggestions for changes, please let us know. The detailed corrections are listed below point by point:

The manuscript presents a new method for combining high-resolution daily satellite precipitation estimates with rain gauge observations. The method is applied and evaluated over the Jinsha River Basin for the summer period in 2016 (June, July August). The performance of the method is compared to already existing satellite datasets CHIRP, which is also the base for the new dataset, and CHIRPS. The evaluation reveals an improvement in accuracy of precipitation estimates with rain rates of less than 20 mm per day compared to CHIRP and CHIRPS, however, the chosen time period of just 3 months seems to be rather short for this somewhat general conclusion. For heavy precipitation, however, no improvement could be found. The dataset and the blending method are described and the data is available for free.

The manuscript fits in the scope of ESSD, but some issues need to be addressed. I recommend taking the following suggestions and comments into account:

1.

(1)- It is not quite clear to me what exactly is the reference dataset in this study. On page 6, line 170 the authors state that 70% of the total gauged stations and gridded points were used as the training dataset and the remaining 30% serve as reference dataset. How was decided which station / grid point was used for training and which station / grid point was used for evaluation? As I understand it is a mixture between actual station measurements and gridded, i.e. interpolated, station data. Is the ratio for both data types also 70% training and 30% reference data points? Is there a difference in performance metrics when only one of the two datasets is used for evaluation? Direct measurements from stations might be even more accurate than the interpolated data.

Answer: In the previous experiment, the 30 rain gauge stations and 170 gridded points were used as the “true” precipitation values. However, the gridded precipitation data was from China Meteorological Data Service, interpolated from 2472 rain gauge stations, which was less accurate than the direct measurements from stations, for example, daily precipitation was more than 1000 mm at one interpolated grid point. So

only the rain gauge observations were used to the new experiments. What's more, selecting 30% of the stations for validation was not an appropriate validation method, while the leave-one-out cross validation step was a better instead for using all the stations in WHU-SGCC correction algorithm

Change: We have only used 30 rain gauge stations as the reference precipitation values to conduct the WHU-SGCC method. We changed from "The proposed approach was evaluated for the Jinsha River Basin for JJA 2016. From that data, the training samples represented 70% of total gauged stations and gridded points, and the remaining data were used to verify the model performance." to "The proposed approach was evaluated over the Jinsha River Basin based on 30 gauge stations and CHIRP satellite-based precipitation estimations during JJA from 1990 to 2014. The leave-one-out cross validation step was applied to computing the out-of-sample adjusted error with gauge stations."

(2)- A more detailed description of the reference dataset and decision making process is desirable, e.g. a map with the mean or the sum of precipitation during the observation period at the reference grid points and stations.

Answer: Thanks. Done.

Changed: We added a map with the multi-year (1990-2014) average annual precipitation (Fig. 2). The multi-year average annual precipitation increases from north to south and the spatial distribution of precipitation is uneven, with an average annual precipitation ranging from less than 250 mm to more than 600 mm during the summer seasons over the Jinsha River Basin.

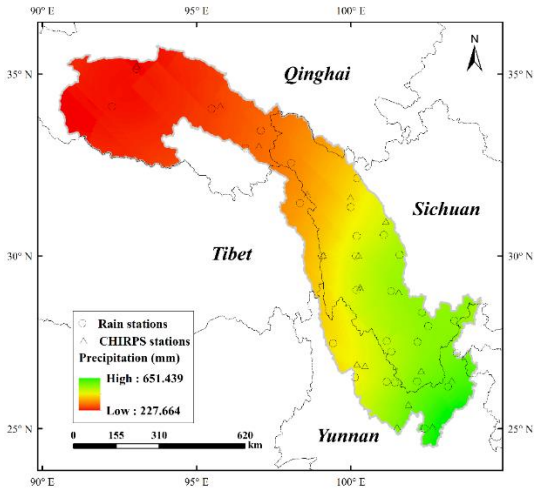

Figure 2 The multi-year (1990-2014) average annual precipitation during JJA over the Jinsha River Basin. 30 rain stations were provided by the China Meteorological Administration stations, the other 18 CHIRPS fusion stations were provided by the Climate Hazards Group UC Santa Barbara online at ftp://ftp.chg.ucsb.edu/pub/org/chg/products/CHIRPS-2.0/diagnostics/global_monthly_station_density/tifs/p05/ (last access: 10 December, 2018).

(3)- As far as I understand, using this evaluation dataset implies that only C1 and C2 grid points are evaluated, because they contain either a rain gauge station or a grid point of the interpolated station data. Is that correct? Can the authors give an assessment on the quality of the method at C3 and C4 pixels?

Does the selection of the stations and grid points for training have an influence on the model performance? Depending on the location of the points for adjustment the quality of the blended dataset may vary. An ensemble study using different compositions of the pool of training stations / grid points would give statistically more robust results.

Answer: Thanks. In the new experiment, the leave-one-out cross validation step using all the stations was used to evaluate the performance of the WHU-SGCC algorithm. The training set was used to establish statically relationships when conducting the WHU-SGCC method, and the remaining one gauge station was used to evaluate. The adjusted process shown that the adjustment method for C2 pixels was derived from C1 pixels, the adjustment method for C3 pixels was derived from C2 pixels, and the adjusted values for C1 and C4 pixels were interpolated by IDW with C2 and C3 pixels. There were statistically relationship among C1, C2, C3 and C4 pixels. Thus, the performance of WHU-SGCC method would be evaluated on the overall accuracy, not on a certain class of pixels.

2.

(1)- CHIRP data is used as basis for the WHU-SGCC dataset and it is shown that the blending approach leads to better (light and moderate rainfall) or similar (heavy precipitation) results compared to measurements. CHIRPS, however, seems to perform much worse than the original CHIRP dataset although it is also adjusted to rain gauges. Can the authors give an explanation for that?

Answer: The CHIRPS was derived from blending in-suit precipitation observations and the CHIRP data, with a modified inverse-distance weighting algorithm at a quasi-global area (land only, 50° S-50° N). The blended data (CHIRPS) has an effective performance on a large scale region according to existing studies, such as at the national scale, but there are still large discrepancies with ground observations at the sub-regional level, especially at the river basin scale. The performance and applicability of CHIRPS at the sub-regional level still need to be validated. What's more, the interpolation performance from the limited and sparse rain gauge stations will be affected by more errors which was evaluated with rain gauge stations shown in Table 5.

As such, due to the poor performance of CHIRPS data at the sub-regional scale and the shortcomings of the modified inverse-distance weighting algorithm, the aim of this article is to offer a novel blending approach to improve the precipitation estimated accuracy at the river basin scale.

Change: We changed the sentence from "As such, the aim of this article is to offer a

novel approach for blending daily precipitation gauge data, gridded precipitation data and the Climate Hazards Group Infrared Precipitation (CHIRP) satellite-derived precipitation estimates over Jinsha River Basin." to "As such, due to the poor performance of CHIRPS data at the sub-regional scale and the shortcomings of the existing blending algorithms, the aim of this article is to offer a novel approach for blending daily liquid precipitation gauge data, gridded precipitation data and the Climate Hazards Group Infrared Precipitation (CHIRP) satellite-derived precipitation estimates developed by the UC Santa Barbara, over the Jinsha River Basin." for better explanation.

(2)- It would also be desirable to expand the investigated period to get more robust results, e.g. add more summer seasons from other years.

Answer: Thanks. Done.

Change: we changed the study period from summer of 2016, JJA to a longer study period during June-July-August from 1990 to 2014, to evaluate the model performance more reasonably.

Specific comments

(1) - P.1, L.37: There is twice "without adjustment" in the sentence

Answer: Thanks. Deleted one "without adjustment".

Change: We changed "Without adjustments, inaccurate satellite-based precipitation estimates without adjustment will lead to unreliable assessments of risk and reliability" to "Without adjustments, inaccurate satellite-based precipitation estimates will lead to unreliable assessments of risk and reliability".

(2) - P.2, L.63 and 65: Remove the brackets at Bai et al. and Trejo et al.

Answer: Thanks. Done.

Change: We removed the brackets at Bai et al. and Trejo et al.

(3) - P.3, L.89: Section 5 is about data availability. Section 6 presents conclusions

Answer: Thanks. Done.

Change: We changed the section order that Section 5 is conclusions and Section 6 is data availability.

(4) - P.3, L.102-103: I'm a bit confused here. Does "average annual precipitation", "annual precipitation" and "total annual precipitation" mean the same thing? Or is the total (for me this refers to the sum) of the precipitation north of Shigu almost four times smaller than the mean annual precipitation in the whole Jinsha River Basin?

Answer: Thanks. Done. We have used the spatially averaged annual accumulation of precipitation as an indication of precipitation climatology for the study region. The reference (Yuan, Z., Xu, J. J., and Wang, Y. Q.: Projection of Future Extreme Precipitation and Flood Changes of the Jinsha River Basin in China Based on CMIP5 Climate Models, Int. J. Environ. Res. Public Health, 15, 17, 10.3390/ijerph15112491,

2018.) can support the average annual precipitation statistic.

Change: We changed "Average annual precipitation in the Jinsha River Basin is approximately 3433.45 mm, the total annual precipitation north of Shigu is 937.25 mm, while south of Shigu annual precipitation is 2496.20 mm." to "The average annual precipitation of the Jinsha River Basin is approximately 710 mm, the average annual precipitation of the lower reaches is approximately 900-1300 mm, while the average annual precipitation of the middle and upper reaches is approximately 600-800 mm (Yuan et al., 2018)."

(5) - P.6, L169: I would remove the numbering here, as it doesn't seem to be another part of the method, but refers to the overview of steps 1-4.

Answer: Thanks. Done.

Change: We added this sentence into the first phase in section 3.1, as the reference to the overview of steps 1-4.

(6) - P.11, L.309: Nash and Sutcliffe(1970) is missing in the references

Answer: Thanks. Done.

Change: We have added the Nash and Sutcliffe(1970) in the references.

(Nash, J. E., Sutcliffe, J. V.: River flow forecasting through conceptual models, Part I - A discussion of principles, Journal of Hydrology, 10, 282–290, doi.org/10.1016/0022-1694(70)90255-6, 1970.)

(7) - P.14, Table 4: How is the accuracy assessment of C3 pixels done? What is the reference here? Why is SCC < 0.5?

Answer: Thanks. In the previous experiment, the number of C3 pixels accounted for 62.18% of the total pixels inside the river basin and the major of the C3 pixels had the same location with the 30% testing data. So we evaluated the C3 pixels with part of testing set (rain gauge stations and gridded points). While, in the new experiment, due to the leave-one-out cross validation step using all the stations, the performance of WHU-SGCC method would be evaluated on the overall accuracy, not on a certain class of pixels. So we didn't evaluate the C3 pixels separately.

Change: Deleted the statistical analysis about the C3 pixels.

(8) - P.17, Fig.10: It might be helpful to present the percentage deviation from the observations for clarification of the model performance. It seems that at some days, all three datasets deviate more than 70% from the observations.

Answer: Thanks. Because the daily precipitation of rain stations may be no rain, the percentage deviation from the observations cannot be obtained (the denominator is 0). Due to the leave-out cross validation method, and the days of each rain gauge was different which made difficult to redrawn the figure (the daily precipitation difference between WHU-SGCC, CHIRP, CHIRPS and observations). We evaluated the model

performance based on overall accuracy and rain events for WHU-SGCC.

Change: Now the section 4 was divided into 4 parts: 4.1 Spatial Clustering from the FCM method, 4.2 Model performance based on overall accuracy evaluations, 4.3 Model performance based on the spatial distributions and 4.4 Model performance for rain events.