Responses to reviewer #1

Dear reviewer,

Thank you very much for your constructive comments and suggestions. In the texts below, we will try to answer all questions addressed by the reviewer. If you feel more explanations or revisions are needed. Please do not hesitate to contact with us.

Best regards,

From the authors

This is an interesting and timely study. Reducing simulation uncertainty has always been an important issue of hydrological modeling. Traditionally, hydrological models are calibrated and validated using only runoff data, which not only leads to parameter equivocality and failure to obtain reasonable and true parameters, but also leads to large uncertainties in other elements of the simulation such as evapotranspiration, soil water. This study attempts to explore how incorporating RS-ET data into the calibration could improve hydrological modelling. I think there is potential in this manuscript, but there are several sections that need to be more clearly explained My comments are as follows,

1. There are a large variety of ET products, why choose GLEAM data and how accurate are GLEAM data in this basin? A set of ET model data may have large uncertainties both in magnitude and in spatial and temporal distribution, and direct use without calibration may introduce greater uncertainty. It is recommended that a water balance analysis, which analyzes relationships between the precipitation, runoff, and ET data used in the study, be added to determine the overall confidence in the data and thus improve the credibility of the article's results.

Response:

Thank you very much for this suggestion. To evaluate the accuracy of GLEAM ET data, a water balance at the studied basin was conducted at annual scale. For the upstream area of the Ganzi Gausing station, which is the outlet of hydrological modelling in this study, the runoff ($Q_{est,1}$) was estimated from the area averaged precipitation (derived from MSWEP dataset as input for SWAT model; the analysis about accuracy of MSWEP data is described in the responses to the following comment) minus GLEAM ET and compared with observed value at the Ganzi Station (Q_{obs}). Figure R1 is the scatterplot of Q_{obs} and $Q_{est,1}$. Absolute bias (ABIAS) was computed to evaluate the accuracy:

$$\text{ABIAS} = \frac{\sum_{i=1}^{n}|Q_{obs,i} - P_{sim,i}|}{\sum_{i=1}^{n} G_{obs,i}} \quad (1)$$

where $Q_{obs,i}$ and $Q_{sim,i}$ represents observed runoff and estimated runoff from water balance analysis at time step i, respectively, n is the number of samples. The ABIAS for $Q_{est,1}$ is 0.20.

Huang et al. (2020) used bias-corrected PML-AET data to calibrate Xinanjiang hydrological model. Compared with the accuracy evaluation of PML-AET in the whole Yalong River Basin (Huang et al., 2020), the ABIAS of GLEAM ET is much lower than the uncorrected PML-AET (ABIAS: 0.55). As a high-attitude region, the snow-melting process has influences on runoff in the simulated basin of this study. We further corrected $Q_{est,1}$ by subtracting annual snow-melting amount (derived from Monthly Snowmelt Dataset in China during 1951-2020 (Yang et.al, 2022)) from $Q_{est,1}$ (mentioned as $Q_{est,2}$) and compared with Q_{obs} (Figure R1). The ABIAS of $Q_{est,2}$ is 0.12, which is also lower that the bias-corrected PML-AET data (ABIAS: 0.18, Huang, et al., 2020). It is indicated that partly the error of $Q_{est,1}$ comes from ignoring the snow-melting process and cannot contribute to the GLEAM ET data completely. Based on literature review, the GLEAM ET has been wide used worldwide for analyzing changes in regional water cycle (e.g., Bennour, et al., 2022, Ding and Zhu, 2022), rainfall-runoff modelling (e.g. Dembélé, et al., 2020, López López, et al., 2017). Also, from the results of model calibration in this study, incorporating GLEAM ET data into SWAT model calibration did improve the accuracy of runoff simulation. Based on these facts, we are confident using the current version of GLEAM ET data in this study.

The above-mentioned analysis will be added to the results section of the manuscript.

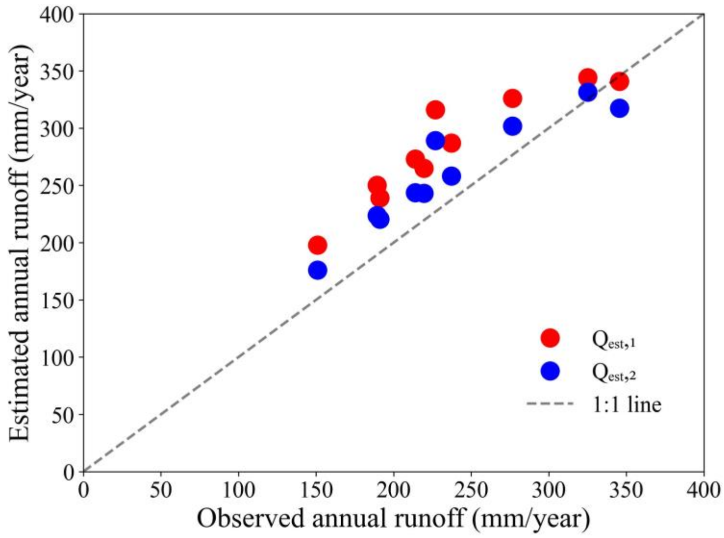

Figure R1. Observed annual runoff versus the runoff estimated from the area averaged precipitation (MSWEP) minus GLEAM ET ($Q_{est,1}$), and versus the ones obtained from precipitation minus GLEAM ET and annual snow-melting ($Q_{est,2}$)

1.Similarly, is it possible to validate or document the regional applicability of climate-driven data?

Response:

Precipitation data from two meteorological stations operated by China Meteorological Administration within the upstream region of Ganzi gauging station are available to evaluate the MSWEP precipitation data the we used to drive the hydrological model. The scatterplots for daily, monthly and annual precipitation between ground and satellite precipitation data at the pixels corresponding to the two meteorological stations (2001 to 2010) are shown in Figure R2. The correlation coefficient for daily, monthly and annual is 0.52, 0.94 and 0.85, respectively. The accuracy is highest for the monthly scale, which is the temporal scale for the hydrological model simulation.

The above-mentioned analysis will be added to the results section of the manuscript.

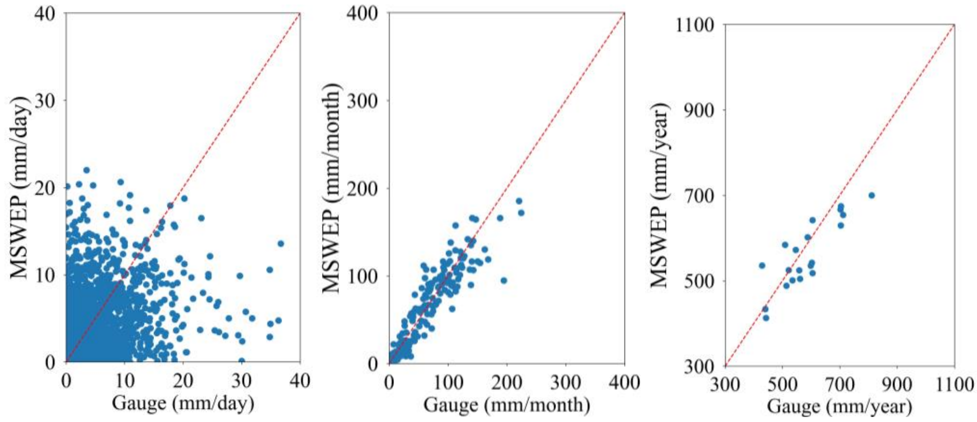

Figure R2. Comparison between observed precipitation and MSWEP precipitation at (a) daily, (b)monthly and (annual) scale

2.In section 2.1, what are the values for the percentage of runoff sources roughly, this could be crucial information. It is also recommended that the percentage of area of major soil types and LULC types be given. Although this may have been shown in the figure, it would be easier for the reader to understand if specific values were given.

Response:

Based on the analysis of MSWEP precipitation data and streamflow data for the period of 2001 to 2010, the rainfall-runoff ratio is 0.59. The percentage of percentage of area of major soil types has been shown in the following table (Table R1). The major land use and land cover type are grass land, forest and bare land, which occupies 78.9%, 13.2% and 12.4% of the basin area, respectively.

These information will be added to Section 2.1 Study area.

Table R1 The percentage of area of major soil types

Soil Type	Grey Brown Soil	Swamp Soil	Grass Felt Soil	Thin Grass Felt Soil	Brown Grass Felt Soil	Black Felt Soil	Permafrost Soil
Percentage of Area	3.36%	2.64%	56.23%	7.69%	4.43%	15.78%	9.87%

3.In Section 3.1, it is desired to make a multidimensional comparison of the analysis of information in figures and tables, e.g., a comparison of NSEQ and NSEET in the same experiment.

Response:

Thank you very much for the comment. The accuracy and uncertainty of streamflow and ET simulation will be compared in the revised paper through the information provided by Table R2.

For Experiment I using streamflow data solely for calibration, the $NSE_{50\%}$ for streamflow simulation in the calibration and validation period is much higher than those of ET, which is all lower than 0 and means that the performance of ET simulation is unsatisfactory. The P_factor corresponding to the streamflow estimation is higher than that for ET simulation, which means more observation are embraced by the uncertainty band. Meanwhile, The R_factor for streamflow estimation is lower than ET, indicating the width of uncertainty band is narrow. For modeling accuracy and simulation uncertainty, streamflow estimation all outperforms ET simulation. For Experiment II using RS-ET data solely for calibration, the $NSE_{50\%}$ for ET simulation in the calibration and validation period is higher than those of streamflow. The P_factor corresponding to the ET estimation is similar with the value for streamflow simulation. The R_factor for streamflow estimation is higher than ET, implying uncertainty of streamflow estimation is higher than ET simulation. For Experiment III combing both of streamflow and ET data for model calibration, the $NSE_{50\%}$ for streamflow simulation is slightly higher than that for ET simulation. P_factor is in the same level for the simulation of the two hydrological variables. The R_fator of ET is a little bit higher than that for streamflow simulation. All these factors demonstrated that the streamflow simulation performs better to a small degree.

Table R2. Accuracy and uncertainty for streamflow and ET simulations

		Experiment I		Experiment II		Experiment III	
		Calibration	Validation	Calibration	Validation	Calibration	Validation
$NSE_{50\%}$	*Q*	0.71	0.75	0.52	0.54	0.81	0.84
	ET	-0.22	-0.09	0.84	0.88	0.79	0.79
P_factor	*Q*	0.75	0.75	0.88	0.88	0.85	0.87
	ET	0.58	0.52	0.93	0.88	0.88	0.88
R_factor	*Q*	0.86	0.92	1.42	1.51	1.1	1.17
	ET	1.42	1.35	1.2	1.14	1.18	1.18

4.In Section 3.2, why is the number of behavioral parameter sets different in the three experiments? Is it because the number of parameters sensitive to evapotranspiration processes in a hydrological model like SWAT is much smaller than the number of parameters sensitive to runoff processes?

Response:

We agree with the reviewer that, a complex model like SWAT, the number of parameters related to evapotranspiration processes is lower than the ones connected with runoff processes, because as the integrated output of water cycle at basin scale, the runoff is determined by many hydrological processes with a basin, which also include evapotranspiration processes. For the three experiments, the data used for calibration is different, which is observed streamflow data, remote sensing evapotranspiration, and the combination of the streamflow and evapotranspiring data. The automatic calibration process tried to minimize the difference between observation data and model simulation by searching the parameter space. As the calibration data are different among the three experiments, the parameter sets being gained by calibration are also different, which explains why is the number of behavioral parameter sets is different.

5.The headings of sections 3.1 and 3.2 are the same.

Response:
The heading of section 3.2 will be changed into "Comparison of parameter posterior distributions among three experiments"

6.There should be an error in the x-axis in set (b) of Figures 2-7.

Response:
The years showed in the x-axis of Figures 2-7 will be corrected.

7.Figures 2-7 could be merged into one or two figures. And Tables 3-4 could be merged into one table. In Figure 6, "b" was mislabeled as "d".

Response:
The figure 2 to 4 are merged in to one figure. Similarly, the Figure 5 to 7 are merged into one figure. The merged figures are shown as Figure R3 and R4. Tables 3 and 4 are merged into one table as Table R2. The sequence number of each figure has been corrected.

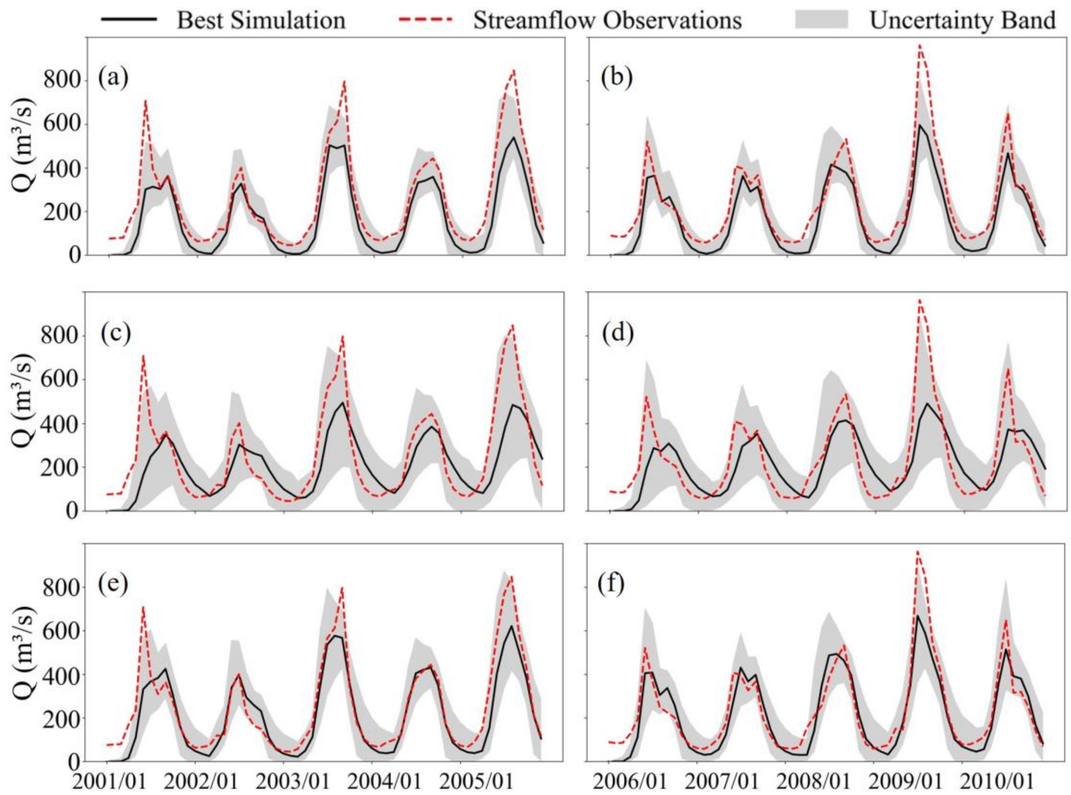

Figure R3. Streamflow observation, best simulation (50% quantile of ensemble simulation) and uncertainty band for the calibration and validation period corresponding to Experiment I (a) and (b), Experiment II (c) and (d), and Experiment III (e) and (f)

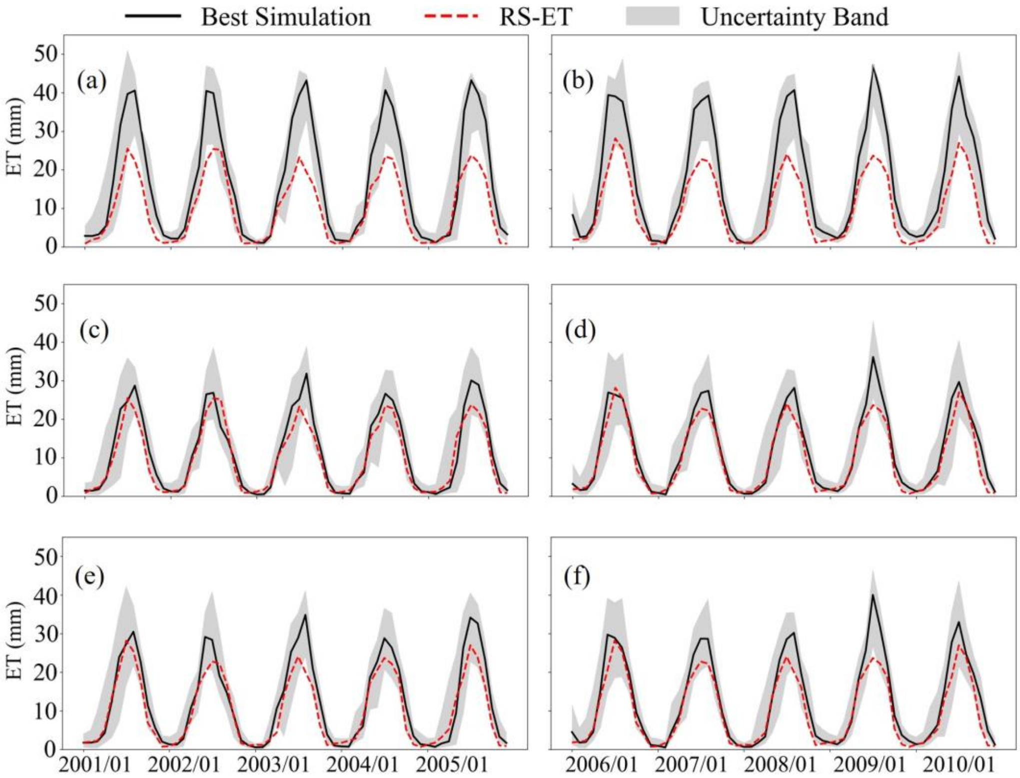

Figure R4. Evapotranspiration observation, best simulation (50% quantile of ensemble simulation) and uncertainty band for the calibration and validation period corresponding to Experiment I (a) and (b), Experiment II (c) and (d), and Experiment III (e) and (f)

8.Figure 12 is missing (d) in the title

Response:

This typo will be corrected in the revised manuscript.

References:

Bennour, A., Jia, L., Menenti, M., Zheng, C., Zeng, Y., Asenso Barnieh, B., Jiang, M.: Calibration and Validation of SWAT Model by Using Hydrological Remote Sensing Observables in the Lake Chad Basin. Remote Sensing, 14(6): 1511. doi:10.3390/rs14061511, 2022.

Dembélé, M., Hrachowitz, M., Savenije, H.H.G., Mariéthoz, G., Schaefli, B.: Improving the Predictive Skill of a Distributed Hydrological Model by Calibration on Spatial Patterns With Multiple Satellite Data Sets, Water Resources Research, 56(1), doi:10.1029/2019WR026085, 2020b.

Ding, J., Zhu, Q.: The accuracy of multisource evapotranspiration products and their applicability in streamflow simulation over a large catchment of Southern China. Journal of Hydrology: Regional Studies, 41: 101092. doi: 10.1016/j.ejrh.2022.101092, 2022.

Huang, Q., Qin, G., Zhang, Y., Tang, Q., Liu, C., Xia, J., Chiew, F.H.S., Post, D.: Using Remote Sensing Data-Based Hydrological Model Calibrations for Predicting Runoff in Ungauged or Poorly Gauged Catchments, Water Resources Research, 56(8), doi:10.1029/2020 wr028205, 2020.

López López, P., Sutanudjaja, E.H., Schellekens, J., Sterk, G., Bierkens, M.F.P.: Calibration of a large-scale hydrological model using satellite-based soil moisture and evapotranspiration products, Hydrology and Earth System Sciences, 21(6), 3125-3144, doi:10.5194/hess-21-3125-2017, 2017.

Yang, Y., Chen, R., Liu, G., Liu, Z., Wang, X.: Trends and variability in snowmelt in China under climate change, Hydrology and Earth System Sciences, 26: 305-329. doi: 10.5194/hess-26-305-2022, 2022.