Dear editors,

Included here is my review of the manuscript submitted to the *Hydrology and Earth System Sciences*.

Manuscript #: hess-2023-200
Title: How Remote-Sensing Evapotranspiration Data Improve Hydrological Model Calibration in a Typical Basin of Qinghai-Tibetan Plateau Region

The hydrology processes in the Qinghai-Tibet Plateau are important, and the hydrological models are useful tools for simulating these processes. In this work, the authors combined stream flow data and the RS-ET data from GLEAM in the calibration to improve the accuracy of SWAT model in Yalong River Basin of the Qinghai-Tibet Plateau. The authors did a large amount of calculations and the data analysis is solid and convincing. However, in my opinion, the findings in this manuscript are not beyond the general understanding, thus, provides few new knowledge to the science public. I feel sorry, but I need to reject this manuscript to keep the high quality of this journal.

The detail comments and suggestions are listed as following:

General comments:

1. Evaporation is an important component in the water balance. It is not surprising that adding another calibration target of evaporation can improve the accuracy of the hydrological models. In the introduction section, the authors cited some previous studies (e.g., Immerzeel and Droogers (2008); Huang et al. (2020)) which have already evaluated the efficiency of such improvement in India, the Yalong River basin, and other basins. What's the advance of this study compared with those previous studies? This model (SWAT in Immerzeel and Droogers (2008)) and this region (the Yalong River basin in Huang et al. (2020)) have already been investigated. Someone can write dozens of papers by combining different hydrological models and different RS-ET datasets applied in different hot zones about this topic, which has fewer contributions to the academic society. Thus, I did not see the innovations of this study. I strongly recommend the authors change the primary purpose of this manuscript. Elaborating the story from another perspective using the calculation they already did will be a wise choice.

2. I did not see any correction about the evaporation data from GLEAM 3.5a. The authors seemed to extract the grid data from the GLEAM dataset directly without any local bias correction, which is not proper. I suggest that the authors calibrate and correct the evaporation data from GLEAM first, since Huang et al. (2020) have already evaluated the different performance in the hydrological models between bias-corrected and nonbias-corrected evaporation data.

3. The authors spent a lot of effort on the sensitive analysis. I really respect the extensive work and understand that they tried to provide understandings of the driving mechanism, which has the potential to be regarded as innovations. However, this manuscript only listed the statistical results with a few simple discussions, which cannot elaborate on the driving mechanism. The authors should offer additional deep analysis of these sensitive results with physical meanings.

Other Specific comments:

1. The abstract is too long, and the authors can simplify it for the convenience of the reader.

2. Line 35, the improvement of NSE (e.g., from 0.71 to 0.81) is not significant compared with the previous studies (e.g., from 0.41 to 0.81 in Immerzeel and Droogers (2008))

3. Lines 56-56, any connection with this study's topic?

4. Lines 276-277, why?

5. Figure 4b, what's the reason for the underestimation of Q between 2004 and 2005

6. In Figure 8, the authors can try only to provide the key parameters in the figure to reduce reader interference.

7. Line 346, "when combing streamflow and RS-ET data for model calibration, the accuracy of simulated streamflow and ET are all higher." I remember that the accuracy of evaporation in experiment two is higher than in experiment three (Lines 276-277).