This manuscript presents different calibrations and evaluations based on different climatic conditions spatially and temporally intended to explore how the model performance are sensitive to drought condition and the real potential causes data, and test that a drought included in calibration period would improve model transferability or not. The authors designed two calibration experiments and three evaluations under three different wetness conditions. They mainly found that a drop in performance of Q indeed happened in their study area based on Continuum model, and was related to the representation of ET anomalies rather than TWSA, and including a moderate drought in the calibration did not lead to an improvement in Q and ET simulation during a severe drought. The research makes a contribution to understanding the application of LSASAF product, and model performance under different climatic conditions.

However, I have some major concerns that need to be addressed prior to reconsideration:

1) Based on the results, I don't think authors can say that "the drop in Q modelling performances during the severe 2022 drought event can be related to the mis-representation of ET anomalies, among other factors". I just observed that a drop happened in ET performance from moderate droughts to severe droughts which was similar with that for Q performance. I didn't see any other evidence to prove that drop in Q performance can be related to ET simulation. Please design more experiments to give audience more evidence.

2) I don't think the authors really solved the research question: would a drought in calibration period improve model transferability or not? Authors just have two calibration experiments(one with normal period and another one with moderate drought) and then compared the evaluated results in severe drought. The results in this study were very different from that in Yang et al. (2021). But this may result from the different model which Yang used and this reference can not prove your result is correct. I suggested that author can design more experiments including different type of droughts based on different models and compare their results so that make your results more reliable.

3) how do you define wet, normal, moderate drought, and severe drought? I didn't see detailed clarification or an indicator in this manuscript.

4) could you please add the evaluation performance results in supplementary?

5) please make your paragraph format consistency.

6) what does the grey shade represent in Figure 5? Please add that in the text blow the figure.

7) what does the river basin really look like? When I see the river basin in figure1 and 4, the river basin looks well, however, the river basin in figure6-8 looks like that it was stretched vertically. Please make the river basin consistency in your figures. And please organize your figure6-8 better.

8)Line 203, what is “a climatology”? please clarify it in details.