Comment on hess-2022-204
Anonymous Referee #1


I enjoyed reading the manuscript. This study combines several remote sensing products to improve the hydrologic model’s physics together with streamflow performance.

I only have several concerns regarding the presentation of the work and the framework.

-L159: USBR dataset needs a reference (url/doi)

-L239: why monthly and not daily streamflow performance was targeted in calibration? Daily water balance is key for hydrologic models. Monthly fit is easier and reducing the value of baseline simulation.
-L248-Fig3: in this section (3.3) I read what has been done but I couldn’t find answers for the question “how”. Framework needs elaboration. Baseline simulation is clear but other steps are not clear.

-Most importantly, model calibration is an exercise of fine tuning of the model parameters. Before calibration a robust sensitivity analysis (SA) must be applied for such sophisticated models with many parameters to reduce the search space. Did authors apply SA in their study?

-The authors followed a stepwise approach but sensitive analysis (sobol’s, LHS O-A-T, Morris etc) may reveal parameter interactions which can be important to consider during calibration. The authors should discuss the implications of parameter interaction in their framework.

-It would be good to simultaneously use LST and snow RS data on uncertainty reduction via model calibration.

-My biggest concern is about the spatial structure of the selected hydrologic model (VIC) which is a semi-distributed model. In such model parameters get the same value in the same subbasins which inevitably leads to uniform parameter fields and resultant uniform flux maps. One way to avoid this, is using fully distributed models with parameter regionalization tool based on pedo-transfer functions using soil and vegetation properties.

-The authors used bias-sensitive error metrics (rmse, bias) and CC as bias insensitive metric. CC must be used with cautious it can be affected by outliers in the sample. High CC values are not always informing. Instead spatial metrics (SSIM, FSS etc) could be
preferred.

-Fig6: the readers can be curious why median night time bias for baseline is usually less than other 3 cases.

-Fig7 should be better explained. How correlation between parameters is assessed?