Comment on hess-2020-669
Anonymous Referee #1

In the recent decade, the availability of spatially distributed remote sensing-based evapotranspiration, soil moisture and biophysical variables have made it possible to constrain and evaluate ecohydrological models reasonably in space and time. Lee et al. have presented a work that attempted to highlight the usefulness of spatially distributed remote sensing-based LAI and ET coupled with conventional streamflow data to enhance the SWAT model predictability at watershed, sub-watershed and HRU level. The topic raised is interesting and relevant to HESS; however, the manuscript has several critical limitations that need to be addressed before accepting for publication. I have provided below my major and minor comments back-to-back:

Major issues:

**Model calibration**: I think the model calibration part of the manuscript is weak and needs considerable improvement. The manuscript seems to present a kind of multi-variable SWAT model parameterization using in-situ streamflow and remote sensing-based ET and LAI. However, firstly, the organization of the calibration framework is not clear and difficult to follow/understand. Second, the model calibration exercise involves around 27 model parameters that could vary at HRU and /or sub-watershed scale, and yet the optimization was done for 3000 model runs only. Given the number of parameters and their interaction, and the multiple calibration variables consideration, I think this is not sufficient and probably the results presented are suboptimal. This can be well marked from the results presented in the manuscript. For example, the calibrated parameter values in general showed wide range, including CN (+/- 19%), which is the same as the calibration range setup (+/-20%). The collective poor model calibration effort can also be observed by looking at the range of the KGE values (0.65-0.87) for streamflow using selected behavioral solutions. There are several published studies that deals with SWAT model multi-variable optimization, and the authors should clearly explain and justify their calibration approach. Additionally, the model evaluations run should be increased at least by a few fold given the number of parameters and variable considered.
**Poor model configuration/ill-posed**: The manuscript setup the SWAT model to simulate corn and Soybean LAI and ET because information is available with regard to crop rotation, fertilization application and planting and harvesting calendar. On the contrary, the ET and LAI simulation for forest and pasture areas are not configured and/or calibrated reasonably based on literature values for the region. I believe this limitation significantly influences the water balance estimations in the basin because such cover type accounts about 40% of the watershed. Although the calibration that was done is not sufficient enough, it is likely the uncertainty from the uncalibrated vegetation parameters (as mentioned in Line 280-282, line 418-426) contributed to the overestimation in streamflow at the watershed outlet. The manuscript also discussed underestimation of SWAT simulated ET at watershed scale. Besides, I would argue that comparing model performance in simulating ET and LAI at sub-watershed and HRU level without reasonable calibration of parameters that control the state(LAI) and flux(ET) variables spatially could undermine the benefit of RS-based ET and LAI products. The study by Yang and Xuesong (2016) could help to address part of this problem.

"Improving SWAT for simulating water and carbon fluxes of forest ecosystems. Science of the Total Environment. DOI 10.1016/j.scitotenv.2016.06.238"

**Improving the data and method section**: the manuscript failed to include critical information that are useful to understand the model configuration, forcing data and calibrated model output evaluations. For instance, the manuscript mentioned daily precipitation is used to force SWAT model, however, it is not clear whether the data from one gauging station or all were used. Apparently, there is no gaging station in the watershed, how is the precipitation magnitude and seasonality in those three nearby gauges differ? and how the available information is related to the observed streamflow? I think such primary information are necessary to provide a good background for the reader. There are several missing pieces of information across the manuscript that needs further improvement. What kind of runoff generation method used in SWAT? No mention at all, but we see CN being calibrated. How many HRUs the model has? What kind of method was used to determine HRUs?

I think it is not a good idea to mix model calibration and validation period (see line 277). I hope this is a typo, otherwise all the calculated indices wouldn’t reflect a reliable model verification. I would also suggest to add a couple of statistical indices in addition to KGE, at least for post-calibration evaluation if not during calibration.

One of the interesting part of the manuscript I would say is the effort to shed light on the spatial consistency of ET and LAI after calibrating SWAT model using streamflow, ET and LAI at watershed level. I would suggest merging section 2.5 and 2.6 into one section as spatial consistence of ET and LAI and then describe as subsetion sub-watershed level and HRU level.
Minor comments/questions

Ln 92-95: It is good also to add the fact that ET is in general the biggest flux from the water balance components.

Ln 103-105, briefly include more information on the improvement, any statistical measures used to highlight the improvement

Ln 122-143: Part of this text should move to the methodology part.

Ln 154-155: add the source for the soil classification system: USDA NRCS. Please only highlight B and D as they dominate clearly the watershed.

Ln 158: I assume you’re referring here daily average temperature not max or min temperature, please specify for better clarity

Ln 189-190: what factors? Please specify

Ln 192: please add reference

Ln 205: what about transpiration? Isn’t this the main component of the study in relation to cropland and ET?

Ln 206-216: Please try provide important information when you describe concepts. It is better to describe LAI development from the onset of the growing season to crop maturity, as your investigation tries to uncover the benefit of RS-based ET and LAI to constrain SWAT. Equation 3 is not really relevant, and I would replace it with the logistic development curve before maturation. Mind you after maturity, the plant ceases to transpire and take up water and nutrients.

Ln 219: In the Fig.1 there are three precipitation and temperature monitoring stations around the watershed. Which station data was used to force SWAT? Or did you use all of them? If so how do you aggregate them or assign to each sub-area? You mentioned in the manuscript precipitation varies spatially, I think it is useful to provide the reader a brief
information about the dominant factor influencing the precipitation pattern while describing the study area. Elevation information might be useful too.

Ln 221:-222: do you mean observed daily solar radiation, relative humidity and wind speed are not available for the watershed? If so mention it explicit. Then we know why the weather generator is used.

Ln 223: modify “…(DEM) data were..” to (DEM) data was

Ln 236: add source for RS-LAI

Ln 271: These maps are not well integrated in the text. Why do you want to keep them here?

Ln 277 why do you use the calibration period again in the validation period? See also comments in major section

Ln 273: Please mention explicitly LAI and ET calibration is at a watershed level

Ln 273: why are you not adding one or two statistical indices to improve the calibration results?

Ln 298: i) Please add a footnote to the table for some of the parameters that vary by landuse/soil type. ii) where is the sensitivity?

Ln 383: Given the small parameter sampling effort, I would not say there is a reduction in degree of equifinality.

Ln 412-424: To what extent do you think biases in ET simulation with forest cover affected the calibration? Do you think the results would improve if you use ET and LAI per cover in the calibration than watershed level ET? I think this at least help to constrain reliably those sub-watersheds and HRUs covered by type#4 to #8

Ln 428: why do you think SWAT seems to simulate well consistently (as shown in Fig. 5)
LAI than ET?

Ln448: Use the same value range (color key) for both LAI and ET.