Interactive comment on "SWAT-CUP for Calibration of Spatially Distributed Hydrological Processes and Ecosystem Services in a Vietnamese River Basin Using Remote Sensing" *by* Lan T. Ha et al.

Anonymous Referee #2

General Comments

The manuscript undertakes a calibration of SWAT with the SUFI-2 procedure using remote sensing data from precipitation, ET and LAI at the HRU level for the Day Basin in Vietnam. The paper anticipates that this method is better for the quantification of eco-hydrological processes in un-gauged basins. Generally, while the authors use a novel and interesting method, they confuse the terminology between SWAT CUP and SUFI-2. As well, an ensemble of ET products was applied, and often unclear or vague process descriptions are provided (i.e. for SUFI calibration and in the results). Therefore, for

these and the reasons listed below, my recommendation for this manuscript is "major revisions".

The paper is written in good English. However, often the term SWAT-CUP is used, when really SUFI-2 is being referred to. The authors need to specify which programs are in SWAT-CUP and clarify which was used in this study (i.e. SUFI-2). Please make sure to make the distinction between SWAT-CUP and SUFI-2 throughout the paper. Also, the SUFI-2 calibration procedure is not described, i.e. how does it function, what are the main selection processes for parameter sets, etc.?

The authors should clarify how the sensitivity of parameters was carried out for the basin: information on page 7 is different than on page 16. I believe the authors only relied on literature studies of sensitive parameters. If this is the case, I strongly suggest to conduct an analysis for the sensitivity of the ET and LAI and other parameters in their basin to determine the most sensitive parameters. Also, the Nash-Sutcliff was chosen as the main objective function, I would suggest to also add the KGE to put less emphasis on the peaks and include other processes leading to lower values.

The authors have combined the ET satellite products to compare the calibration to, but they should have undertaken a comparison to each individual product. Even though they justify the use of the ensemble mean by comparing it to observed data, the SSE-Bop performed just as well. Furthermore, the mean of the ET products is meaningless because they are based on different physical processes and equations and the differences are subdued due to averaging. Using the suite of ET products individually provides a more interesting range of uncertainty, showing the peaks and troughs, much like using an ensemble of climate change simulations and is more helpful to the reader.

The authors do not specify which HRUs had data from the satellite data for the calibration to be undertaken? In which subbasins were these located? Surely not all 7909 HRUs were calibrated individually!

Specific comments:

In the abstract, provide the size of the watershed. P1, L29. Remove the word non-significance.

P2, L3: what is mean by micro-climate cooling?

P2, L24: why is SWAT recommended for un-gauged basins?

P2,L33: what is meant by "It is designed and applied to field measurements only"?

P3.L1: do the authors mean SWAT instead of SWAT-CUP? Because SWAT-CUP is in fact the model SWAT.

P3.L25: It would be interesting to add a paragraph on the uncertainties of the ET products in the introduction.

P5.L13: Are there any inter-basin withdrawls annually?

P5,L14: The "core engine of SWAT" is what exactly?

P6. L2: how does SWAT simulate the sequestration of carbon?

P6, L16: what is meant by ecological change? Provide examples.

P6, L23: SWAT-CUP does not have any intelligence. Only living beings have this.

P6, L23-25: confusion with SWAT-CUP and SUFI-2, they are used interchangeably.

P8, L10-13: how were the performances of the satellite products determined?

P10,L5-7: explain what a hot and cold reference value is.

P10, L12-14: Taking an average of the ET products is not recommended! Rather, each individual ET product should be compared to the simulated outcome so that a range of uncertainty can be gleaned.

P10, L15-17: the downscaling of the products requires clarification and more detailed information, especially for CMRSET. This is an important step that is not well described. A year reference to Guerschman is also missing.

P10, L20: which 10 consecutive years were used?

P10, L24: it is unclear which "analysis" is being referred to? The one from Simons et al., 2016 or this study? This section should also go in the results (if it belongs to this study) about comparing the ET products.

P10, L23, I can only count 4 ET products mentioned (MOD16, SEBS, CMRSET, SSE-Bop), not five.

P10, L25: I believe the "ensemble mean" is meant, not the ensemble.

P11, L13: How was the ET ensemble tested further before being implemented?

P11, L14: SUFI-2 is meant, not SWAT CUP.

P12, L8-9: Please list the typical crops found in the watershed, if there are any other than rice.

P15,L13-14; Describe the LAI process in SWAT. What is meant by "the empirical LAI parameters prescribed in SWAT use an internal database"?

L14: Again SWAT CUP is not meant here. The "fine tuning" needs to be explained (brief mathematical description) how in the methodology under calibration section.

P16,L5: remove this sentence

P17,L6: SUFI-2, not SWAT CUP. Explain why Abbaspour (2015) recommended 1500 simulations.

P20,L11-13: this is a very poor explanation for what is being observed.

The conclusion section needs to be reworked to reflect the results and "take home message" of the study.