**Interactive comment on** “Evaluation of global terrestrial evapotranspiration by state-of-the-art approaches in remote sensing, machine learning, and land surface models” by Shufen Pan et al.

**Anonymous Referee #2**

Received and published: 12 September 2019

This paper was already well-written, especially for the detailed discussion of limitations and possible next steps of different ET products. The reviewer thus has a few minor suggestions for the authors to consider. 1. The remote-sensing based, machine learning, and LSMs ET were comprehensively intercompared. However, how is the performance of ET outputs from the Earth system models (e.g., those from CMIP5 and CMIP6) and the reanalysis? There must be a reason why the authors did not include them. But please clarify this or add these comparison results. 2. In lines 245-246, you indicated the benchmarking products are from the machine learning and physical-based satellite datasets. It seems confusing both here and in Figs 3, 5, and 7. For example, in Fig. 7, if the benchmark product is the simple combination of the two data sources, the variation of benchmark product (yellow line), which is the averaged value, should be largely in the middle of the remote sensing (green line) and machine learning (blue line) results. Please double check this. 3. The Abstract seems quite long. Please double check if the Abstract length fits this particular journal. 4. In line 483, Fig. 5 does not have subfigures.

---