

The authors have improved the manuscript quite a lot. The discussion now contains much more explanations to bring the results into context. I still have some more remarks, and would say that revisions are still necessary.

1. In the authors' response to my general comment 2 (i.e., (2) The innovativeness of the present version cannot be observed and was not highlighted in the Introduction.), I think the authors still did not point out the novelty of the study. Indeed, assessing model uncertainty is already common knowledge from the point of view of input data and model structure. This does not seem to be a point of innovation. The authors still need to state the innovations of this study in detail.

2. In the authors' response to my general comment 3, ok, the authors have provided some description on JULES-ES model in the revised version. Indeed, these basic descriptions can be found through the original references. However, the reader is more interested in how the authors applied the model to this study. For example, the authors did not provide details on model parameter and their settings in your study area; how to determine some important parameters relevant to this study and how the model was validated by site-based observations. Despite the fact that the authors cite relevant publications, this is still hard to believe the results of their study, and I strongly request the authors to provide the results of the validations based on site observations (i.e., Flux Tower Eddy-Covariance Measurements). Although JULES-ES model is one of DGVM and is widely used, the authors need to further test it using observations when applying it to Indonesia, which is directly related to the reliability of this study.

3. The authors mentioned that 'JULES-ES is one of Trendy V12 and has been widely used'. If the 0.5 degree resolution was adopted in this study, it appears that the results of this study can be obtained directly from Trendy v12. This study appears to look like repetitive work. In addition, 0.5° is fine for large-scale assessments like the globe, while caution is needed when using it for regions. The authors have introduced a lot of uncertainty by resampling some of the higher resolution data (30m) to 0.5° , even though the authors state that it is for better comparisons. I would strongly recommend that the authors add the assessments of 30m to the supplementary to make the results more reliable.