Comment on essd-2021-113

Anonymous Referee #2

Referee comment on "Refined burned-area mapping protocol using Sentinel-2 data increases estimate of 2019 Indonesian burning" by David L. A. Gaveau et al., Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2021-113-RC2, 2021

The MS has a clear structure and a good description of reference data collection. However, my major concern is the design of the mapping framework. The most confusing part is the integration of NBR and RF classification. As stated by the authors, NBR and dNBR were used to "detect the day when a pixel's vegetation was disturbed by fire". If this is valid, what was the point of using RF for burned area mapping? On the other hand, RF can be directly used for classification. What was the benefit of combining NBR and RF? Another concern is that the authors put the research in the context of national burned area mapping in Indonesia only, but there is a lack of clarification on the scientific contribution of the work. There is no literature review of the state-of-art on large-scale burned area mapping. There is no scientific objective (what scientific issue was addressed), and no comparisons with similar studies on the topic (expect for the comparison with the Official and the MODIS results).

Other comments follow:

Introduction: The background info about wildfire is too long, especially for the first four paragraphs. The emphasis should be burned area mapping.

L108: Isn't visual interpretation more accurate than machines? Many field data are from visual interpretation, including yours. It is important to point out the issues in their visual interpretation strategy/method/data.

L109: It is surprising that such a national campaign does not have protocols for accuracy assessment.

Section 2.2.: The purpose of getting the pre- and post-fire composites with NBR is confusing. Why did you use RF classification since you already identified the burned pixels with NBR? Or why not directly using RF to extract burned areas?

L163-165: The description of implementing your method is vague. What NBR and dNBR thresholds did you use? How did you know the variation of NBR was caused by wildfire, not other events (e.g., plant disease)? Do you also need to have a vegetation baseline map?

Section 2.3: How did you tune RF? What parameters did you use?

L177: Aren't the training samples too small for national scale mapping?

Results: You compared your overall results with the Official and the MODIS results. However, it is also important to pick sample locations to demonstrate what types of areas had high agreements and what areas caused discrepancies.