The authors seek to generate a new global LC map by fusing the existing ones using Dempster-Shafer theory of evidence. The manuscript is easy to follow and language is always understandable. The technique is routine and lack of innovations. By examining their result, I accidently found an error which suggests the data and method they used may not robust (see my comments below). The new map achieves higher accuracy but also shares the common problems in the existing data, such as unstable performance for specific LCs or in certain areas. As such, I am not convinced that this dataset and manuscript could be candidates for ESSD.

The GlobeLand30 in 2010 has a 5-year temporal gap between the other datasets. The LC changes in this 5 five years will bias you result. How did you deal with it? Please clarify this in the manuscript.

To my knowledge, the GSW data has multiply layers and its historical data is provided monthly. Therefore, how did you derive the water bodies from GSW for 2015. This should be explicitly clarified in the manuscript.

For single-class datasets (e.g., GSW), how did you deal with the background? Did you just ignore it or treat it as non-water? I think the latter is more useful.

Why different number of blocks were chosen for patch-based samples while the number of pixel-based samples seems to follow an equal allocation? Grids with more blocks will have more weights in the validation.

When I try to download your result, I found it was labeled by grid id. It would be better to label it with latitude and longitude (e.g., upper left corner), which is a straightforward and common way.

Echo to my comment above, I download a small file and load the smallest tile into my computer. I found a volcano (5.048 S, 151.330 E) in the Papua New Guinea was misclassified into water bodies. So, I further check the datasets you used. It turns out the error comes from the GLC_FCS30 and FROM-GLC (check the figure below). It indicates that your approach, despite the additional training samples, failed to correct such error. This may be a small problem when it comes to global mapping, and accidently found by me. But it's also a reminder for the authors to check their data and methods.
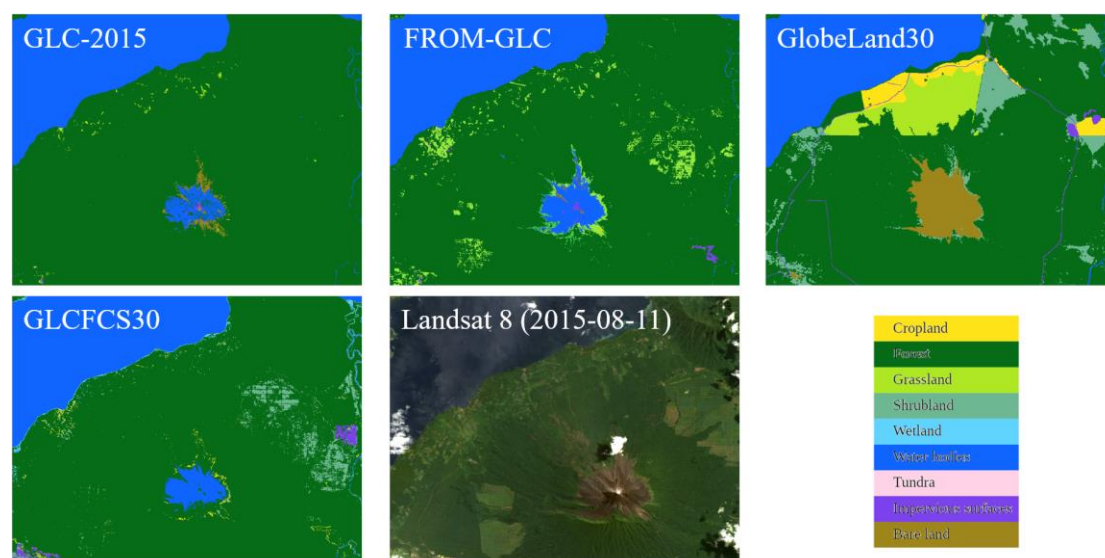
GLC_FCS30 adopted a detailed classification system (level-2) only in some places (seems to inherit from the ESA' CCI_LC). Therefore, I think this may lead to geographical accuracy biases even after you remap the level-2 LCs to yours. How did you deal with it, could you clarify?

The accuracy assessment of your results shows the same pattern with the existing ones, where some LCs (e.g., shrub and wetland) always possess lower accuracies. Geographically, both your results and existing ones exhibit poor performance in areas with more disagreements (Table 7). I don't see much contribution and improvements in this dataset.