The manuscript "High-resolution and Multitemporal Impervious Surface Mapping in the Lancang-Mekong Basin with Google Earth Engine" describes a new IS map for Mekong basin. To this aim, the authors applied visual-interpreted samples to multi-source features derived from Sentinel-1/2 images. The whole framework is routine and some key parameters should be further clarified or discussed. Current accuracy assessment showed better performance than some datasets. However, comparison with the state-of-art datasets is missing, the superiority to the existing works remains unclear. Moreover, the test samples seem to be located mostly in urban and sub-urban areas, leading to incomplete accuracy assessment over rural regions. The structural of manuscript is clear but the writing should be carefully revised. As such, I would suggest a rejection.

Line 70. Many 10-m global datasets have been released, such as the GHS-BUILT-S (Corbane et al., 2021) from EC JRC and the GISA-10m (Huang et al., 2022) from Wuhan University.

Line 100. I would delete the "support" here.

Line 125. Did you just stack the Sentine-1 data from both "ascending" and "descending" orbit together? This will lead to distortions over mountain areas if you do so.

Line 195. I know these metrices make sense, but you may explain why they were chosen.

Line 199. Could you explain why temporal metrices were only derived from Sentinel-1? There are much more Sentinel-2 images (Figure 2).

Table 2."Number" ->"Dimension".

Line 205. Could you show us the distribution of training samples?
Line 233. How this threshold (SAD > 0.125) was determined?

Line 235. Why a pixel can be regarded as water if its MNDWI is greater than 0.12? It's not a robust method.

Line 240. Without seeing your training samples, I presume many pixels identified as changed may not change from NIS to IS if they are far from IS. Could you do a quick ablation experiment to better demonstrate the effectiveness of the sample migration?

Line 265. The methods (and figures) you described here are similar to that in Li et al., (2015). You should at least cite it.


Line 272. If your results follow the assumption that transition from IS to NIS is rare, you may reduce the results to a single band where pixel indicates the time when IS was first detected, instead of putting annual data into separate bands.
Figure 6. It seems that most of your IS test samples locate in or near cities. Could you provide the ISA density around the IS samples? Courtesy of higher spatial resolution, buildings and roads in rural regions can be better delineated in Sentinel images. Therefore, accuracy assessment over rural regions for 10-m IS mapping is important.

Line 302. Are there new (or interesting) findings from your results?

Line 328. "0:05" ->"0.05"?

Table 5. It's would be more interesting to compare your results with GHS and GISA-10m (I mentioned above). They are both 10-m thematic mappings, same as you did here. I strongly suggest you to do so.

Line 377. In fact, according to your results (Table 5), the ESA-2020 achieved higher PA in all cities, indicating that ESA-2020 has less omissions. This is contradictory with the statement here that ESA-2020 ignores buildings in rural areas. Could you explain it?