Comment on essd-2021-313
Anonymous Referee #1

Comments to the authors
This paper proposes a very interesting spatiotemporal gap-filling framework to generate seamless global 1 km daily (mid-daytime and mid-nighttime) LST products from 2003 to 2020. The manuscript is clearly written and well-structured. The products provide global coverage and better accuracies for long-term time-series LSTs. After incorporating my suggested edits, I recommend acceptance.

Major comments
1. The novelty of this work is not comprehensively presented throughout the study. The global significance and necessity of this work should be clarified further. It can be more clearly presented in the abstract, introduction, and conclusion sections.
2. A more detailed explanation of the methodology is necessary. I suggest that SI and S2 should be moved to Section 3.2 (spatiotemporal fitting).
3. More comparisons to previous methods are necessary for the discussion section to show the validity of your method, except for Li et al. (2018).
4. The method may help reconstruct seamless global daily LST products based on MODIS datasets, but the study is limited on clear-sky conditions. How about the LSTs under cloudy conditions? More discussions may be needed.

Minor Comments
1. Line 21 and Line 237-242: Compared with other gap-filling methods, further details should be provided as to how to quantify the effectiveness and efficiency of the gap-filling framework.
2. Line 94-95: confusing sentence, please rephrase.
3. Line 111: What are the criteria for identifying the PVD at 5%.
4. I understand Section 3.1 is from previous studies, but it is necessary to provide more basic information about this section.
5. Fig 2 should be revised.
6. Put more details for Fig 3.
7. What are the advantages of using the smoothing spline function?
8. I did not see any consideration about PVD<5% of four observations from the two satellites.
9. Line 138-139: Not clear of how you manually introduced under three scenarios gaps here? Are the three gaps randomly distributed or continuously missing? How was the reconstruction accuracy in different distributions?
10. Why does the number of pixels (N range from 9300 to 9350) vary approximately among different excluded areas (25%, 50%, and 75%)?
11. Line 155: I consider the selection of 10 pixels per image seems to be not reasonable for cross-validation analysis.
12. Line 214 does not provide a citation?
13. Line 219-224: I saw this very similar sentence in the introduction.