The manuscript "High-resolution all-sky land surface temperature and net radiation over Europe" has been reviewed. The authors presented a methodology to combine the advantages of geostationary observations at high temporal resolution with observations from polar-orbiting satellites at high spatial resolution, resulting in a gap-free all-sky LST and net radiation dataset at 1-km spatial resolution and daily frequencies for 2018-2019 across Europe. This dataset is important for hydrological modelling and as input to models dedicated to estimating evaporation and surface turbulent heat fluxes. However, more comprehensive analysis on this dataset is required before further consideration.

We thank reviewer #2 for the time and effort put into reviewing our manuscript which, we believe, will result in a marked improvement of this study. We are happy to share our point-by-point responses highlighted in blue. In general, we agree that a more comprehensive analysis is required and will expand on the highlighted sections.

1. Lines 223-225. As the dataset includes all-sky land surface temperature, I think it is necessary to implement accuracy assessment to tell us the uncertainties of the produced LST data.

   Response: We agree with the reviewer.

   Action: We will expand on the LST validation add attempt to validate LST directly. The issue so far has been the lack of access to LST ground truth data and therefore we performed the validation for the resulting radiation datasets. We generally will include more validation statistics for the generated products, also per land cover type and seasonally, as also reviewer #1 commented on expanding the validation.

2. A discussion section is required to explain the results and to compare against existing datasets. For example, lines 217-218, why there are worse accuracy in Belgium for *SWin* and around the Alps for *Lwin*?

   Response: We agree with this. Both SWin and LWin are not produced in this study but are input products obtained from LSAF. Nevertheless, more context can be provided and will be added to the manuscript: It is fair to consider that the temporal variability of cloud cover determines to a large extent the variability of SW and LW. Furthermore, that is also the main information provided by satellite data (clouds and cloud optical depth via top-of-atmosphere reflectances). So the generally high R values for both SW and LW corroborate that satellite products follow reasonably well the in situ time-series.
   LW estimates require screen variables (LW is more indirectly linked with top-of-atmosphere observations than SW), which are derived from NWP - therefore it is not surprising that R and RMSE are not as good as those for SW. The accuracy of screen variables may also explain the worse performances of LW in the Alps - although some orographic corrections are performed, the uncertainty is likely larger in mountainous regions.

   Action: We will expand on the discussion of the results along these lines, including an accuracy assessment of these products in the respective literature as well as an analysis of the time series at the mentioned locations with worse performance. This again is also in line with reviewer #1 asking for a more extensive validation.

3. Pearson's correlation coefficient and RMSE are not enough for validation. Examples of comparison of temporal patterns between estimated values and in-situ observations at typical stations are suggested. Meanwhile, the impact factors on the estimated variables can also be analyzed. For example, how does the RMSE change across seasons? Do land cover types significantly affect the accuracy of estimated variables? How about the accuracies in areas with and without missing satellite observations?

   Response: We again agree that the validation requires more detail.

   Action: We will add time series plots to the manuscript and systematically discuss and show differences in performance across the seasons, geographic areas and land cover. This again is in line with above comments and comments from reviewer #1. More extensive validation and discussion is indeed required.

   Preliminary analysis showed deteriorating performance throughout the winter months due to increased cloud and snow cover which prevents the retrieval of clear-sky LST. The resulting dataset thus relies more on modelled all-sky estimates for LST and the incoming radiation products.

   The emissivity dataset used in the study also relies on clear-sky observations and days with no observations are estimated by linearly interpolating between available data. The uncertainty thus also increases during winter and in regions with more frequent cloud cover. While the availability of the clear-sky estimates varies throughout the seasons there are no areas with no data at all.

   We will also look into the uncertainty of in situ obs and representativeness issues, e.g., in cases when the station is far higher/lower than the pixel's average height).

4. Lines 42-57. A comprehensive summary of existing studies/datasets (including advantages and drawbacks) may help to emphasize the novelty of this study.

   Response: We agree fully with this statement.

   Action: We will expand this section and include a comprehensive overview of similar products and novel LST merging methods.

5. Lines 58-61. What research gaps have the authors solved? It is better to describe it here.

   Response: The main research gap is the availability of high-resolution gap-free LST and net radiation datasets at at least daily resolution which can be either used for analysis or the forcing of hydrological/land surface models. This is addressed by developing a suitable methodology to a) combine LST estimates from polar-orbiting and geostationary satellites in order to combine their advantages in temporal and spatial coverage, and b) combine the resulting merged LST product with other datasets to obtain a novel net radiation dataset.

Action: We will include this information in Sect 1 at the proposed position.

6. Lines 108-111. What is the overpass time for clear-sky LST estimates from Sentinel 3A and 3B, respectively? Why do the authors only use the data from Sentinel 3A.

   Response: For this initial study focusing on 2018-2019 only Sentinel 3A data was used. Sentinel 3B was launched in April 2018 and was flown in tandem with Sentinel 3A from June to October of the same year after which it was moved to its nominal orbit, see e.g. https://www.mdpi.com/2072-4292/12/17/2668. The local overpass time of Sentinel 3A and Sentinel 3B thereafter is the same (ca. 10:30 am/pm) with the precise time depending on the latitude and taken into account in the merging methodology.

   Action: When expanding the merged LST and net radiation dataset to more recent years, Sentinel 3B data will be included making the merged product more robust. However, in this pilot study focused on the 2018–2019 period we would like to focus on Sentinel 3A only.

7. Section 3.3. The performance of the merging method needs to be evaluated.

   Response: Some of the main benefits are gained through the bias-correction steps (1-3). The subsequent assimilation step is to obtain a more “Sentinel-like” product with a more marginal impact.

   Response: We will include more validation statistics and evaluation criteria, in line with previous comments of expanding the validation and also in response to reviewer #1.

8. Line 199. More details on the Kalman Filter can be added to make an easier understanding by readers.

   Response: We agree.

   Action: More detail will be added and when necessary, suitable references to the Appendix with a more in-depth description of the assimilation step will be added.