Author's response: High-resolution all-sky land surface temperature and net radiation over Europe

Firstly, we thank the reviewers again for their very helpful feedback on improving the manuscript. We have taken into account their advice and adjusted the manuscript accordingly.

We'd also like to sincerely apologise for the delay in this review process.

Main changes to manuscript:
Most notably, both reviewers agreed on the fact that they were missing validation statistics, e.g. seasonal and for different land cover types. We have added this to the manuscript. Additionally, we have added a comparison between the here presented dataset and ERA5-Land as well as a table listing the in-situ stations used for the validation.

Below we list the point by point responses to the reviewer comments as well as changes which we have implemented. A few of the responses have been modified compared to the original ones when working on the revision of the manuscript.

Reviewer 1:

1. the high resolution LST product is merged from LEAF (all sky) and Sentinel 3 LST (clear sky). The two LSTs have different spatial and temporal resolutions. While doing the merging, if any cloud effect is considered? If any cloud product is involved? If yes, please indicated it.

   Response:

   We have not included any cloud product although the prevalence of clouds is considered in the merging procedure. Cloud cover is assumed when no LSAF clear-sky data are available and then we fall back to the all-sky dataset. Therefore:

   1) Cloud over is taken into account by only using clear-sky LSAF data for the computation of the bias between LSAF and Sentinel-3, both daytime and nighttime.

   2) When computing the diurnal cycle effect on the misalignment of the Sentinel-3 observation to the full hour, this is only done if the two required LSAF samples on which the linear-interpolation is performed are clear-sky observations.

   3) Due to step 1 and 2, In the Kalman assimilation scheme a higher uncertainty is assumed for LSAF all-sky products than for the clear-sky equivalent.

   Action: This is described in the LST merging methodology section and we have added e.g. "The all-sky LSAF product, which contains modelled LST when cloud cover prevents the direct retrieval, enables the merged gap-free LST product with Sentinel-3 resolution." for further clarification.

In the conclusion we have added “Here, we presented a methodology to combine the advantages of geostationary LST and radiation observations, enhanced with modelled data when cloud cover inhibits the direct retrieval, 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 for 2018–2019 across Europe.” for further clarification.

2. while downscaling the LST product, if any edge effects (coast lines, cloud edges) are considered?

   Response:

   Due to the coarser spatial resolution of LSAF coast lines are not as well represented as in the 1 km Sentinel-3 data, as pixels with predominant open water coverage are masked out. The merging procedure and calculation of all-sky LST and net radiation at 1 km resolution can only be carried out where all input datasets are available and therefore the same rugged coast lines are visible in the LST/net radiation product.

   Action: A simple solution, albeit introducing some uncertainties, is to extrapolate from the closest pixels with valid data on a daily basis. We have implemented this for data visualisation and will apply the same correction to the dataset prior to its public release.

3. Line 220, it is said “Extensive validation of the LSAF and Sentinel 3 LST products has already been performed (see below). Both have an average accuracy below 1.5 K, although it varies across space and time. Our goal is to combine their individual strengths in terms of spatial and temporal resolution to obtain an enhanced representation of landscape heterogeneity”. Although there are extensive validations of the LSAF and Sentinel 3 LST products, the validations are based on different spatial and temporal resolutions. It does not mean that the merged product could also has a good performance. It is good to give the statistics.

   Response: We agree with the reviewer. However, we argue that the generated LST product is in essence an intermediate product for the final merging of the net radiation dataset as the final output. The validation is therefore conducted indirectly by validating outgoing longwave radiation at the available in-situ sites located throughout Europe. This is valid as LST is the only dataset that is modified, thus any changes relate to a change in LST.

   Action: For the outgoing radiation, as well as the other radiation components, we have added validation metrics as well as seasonal and land cover analysis. We have generally expanded on the validation making it more insightful. We hope that the reviewer is content with this option (see section 4 as well as appendix B, C, D.)

4. The paper is lack of statistics. e.g. figure 1, any overall statistics could be summarized in a table? And the absolute RMSEs are given in Figure 1. The percentage-wise is worth known. And so does the validations of outgoing raditaions and SNR. Please summarize the overall

statistics (R, bias, RMSE (including percentages)), degree of freedom) in tables.

Response: We agree that they should be included.

Action: We have significantly extended the validation by adding more performance metrics as well as seasonal and land cover analysis with figures showcasing this. We have added a table in the appendix containing the overall and seasonal validation statistics.

5. More detailed information of in-situ sites could be given or summarized.

   Response: We agree.

   Action: We have added the table listing the available in-situ sites (see Apendix A).

6. Figure 2, 3, 4 and A1, A2, A3 could also give the bar chart distribution.

   We generally agree that this could be useful but given that we have added many figures to the new version of the manuscript we would like to skip this, if the reviewer agrees. Otherwise we would be happy to add this in the final version.

7. Please explain the reasons for case selections. e.g. 30 June 2018 in Figure 4 and 30 Sep 2018 in Figure A2.

   Response: There was no specific reason for the case selections. They are representative for other time steps.

   Action: This is now clearly mentioned "This day was chosen for no particular reason and is representative for other dates."

8. If the LST and SNR products are compared with any other reanalysis or satellite products?

   Response: Yes, good point.
   Action: We have added a comparison of the final daily net radiation product to ERA5-Land.

Reviewer 2:

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: The issue is that there is a lack of access to LST validation sites. Given that the purpose of our LST is to be an intermediate product for the calculation of outgoing longwave radiation we have focused our attention on the latter. We argue that, for the intent of this specific study, this is more useful.

Action: The validation of all radiation components has been extended with more validation metrics as well as validation per season and land cover type. The only dataset which is modified for outgoing longwave radiation is LST, so any changes directly relate to this.

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: Both SWin and LWin are not produced in this study but are input products obtained from LSAF and well validated in the literature. Nevertheless, we agree and more context could be provided.

Action: This text has been 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 numerical weather prediction models – 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."

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 agree that the validation requires more detail.

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 have added some time series plots to the manuscript.

Action: We have added additional performance metrics 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.

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: Thank you for the suggestion. We agree this would be interesting, but this could perhaps be for a systematic review paper. If the reviewer disagrees, we are happy to add some more examples.

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 and generic 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 have modified the introduction to better describe the rationale of this study, e.g. "The novelty of this study lies in systematically exploiting the advantages, and mitigating the disadvantages, in terms of spatial and temporal resolution of available observations, which are well validated, in a physical and consistent manner and assembling a net radiation dataset based on the individual incoming and outgoing radiation components."

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. We have clarified the use of Sentinel-3A and the above information in the manuscript. [quote...]

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, and has a more marginal impact. Generally, as the LST product is essentially a useful intermediate product to obtain SNR, and LST in situ measurements are limited, we have validated outgoing longwave measurements instead..

Action: We have included more validation statistics and evaluation criteria, in line with previous comments of expanding the validation and also in response to reviewer #1.We have also clarified the approach taken in the validation of LST, for instance in the discussion section: "It is to be noted that while a gap-free LST datset was developed within this study, the validation of the dataset was carried out indirectly based on LWout measurements. This served the purpose of the study to ultimately create a SNR dataset."

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

Response: Thank you for the suggestion

Action: We have modified the description in the Annex F to make it clearer.