Editor decision: Publish subject to revisions (further review by editor and referees)

Your manuscript "Assimilating ESA-CCI Land Surface Temperature into the ORCHIDEE Land Surface Model: Insights from a multi-site study across Europe" has been subjected now to review by two reviewers. The two reviewers recommend minor revisions. The paper is of good quality and presents an interesting analysis. However, some improvements should be made. This concerns the discussion of the results, which should be in an independent section and take a broader perspective. He authors should also evaluate whether they can quantify the impact of assimilation on the characterization of vegetation processes. I recommend a minor to moderate revision of the manuscript, before it can be published in HESS.

Reply on Editor decision

First of all, we would like to thank the Editor and Reviewers for their valuable suggestions and insightful comments. We have done our best to clarify all the issues raised by the Editor and Reviewers. We included a new version of the manuscript and the point-by-point responses (highlighted in bold) to all comments. Additionally, we have revised the manuscript to improve the overall grammar and readability.

In this new version of the manuscript, we included an independent Discussion section taking a broader perspective, where we addressed different points raised by the reviewers, such as: i) the impact of the selection of the optimization period, ii) the assessment of the impact of the assimilation on vegetation processes and soil water availability, and iii) perspectives about assimilation on a regional scale and joint assimilation. The discussion section together with the related appendices of the new version of the manuscript are detailed below.

"*4 Discussion*

According to our results, we expect that using the PFT-specific parameters derived from the 34 evaluated sites will enhance ORCHIDEE simulations of LST and energy fluxes on a regional scale. While improvements are anticipated for croplands, grasslands, and deciduous broadleaf forests, they are less likely for evergreen needleleaf forests, especially in boreal climates. It is important to note that in this study, only 7 out of the 15 ORCHIDEE PFTs are represented by these sites, with some, like boreal evergreen needleleaf forests, being underrepresented. This study seeks to improve the simulations of LST and surface energy fluxes on a regional scale. However, ongoing work aims to assimilate CCI-LST data across various pixels for all PFTs, ensuring identical representation (i.e., the same number of pixels per PFT).

4.1 Impact of the optimization period

We acknowledge that the selection of a particular year of optimization may impact the selection of parameters to optimize and the performance of the assimilation. Regarding the selection of parameters for optimization, we conducted two sensitivity analyses at the 34 sites for both 2017 and 2018, separately. The selected parameters for optimization were generally consistent between the two years, with the exception of the parameters controlling the water stress curve (α) and the critical soil moisture above which transpiration is maximal ($\theta_{crit,rel}$). In 2017, α and $\theta_{crit,rel}$ were selected for optimization at 7 and 6 sites, respectively, whereas in 2018, both parameters were selected at 13 sites. This difference is attributed to the drought conditions in 2018, which increased the relevance of these water stress-related parameters. Properly representing these parameters is crucial for future projections of climate and water resources (Fu et al., 2022; Fu et al., 2024), highlighting the importance of considering appropriate conditions for accurately optimizing the processes we aim to improve.

In the twin experiment at the ES-Abr site, we previously assessed the impact of selecting a specific year (2018) versus the entire available 6-year period (2015-2020) on the performances of half-hourly LST and turbulent fluxes during 2017 (see Fig. E1 in Appendix E). We chose 2017 to ensure a more independent evaluation of both calibration periods (2018 and 2015-2020). The results showed no significant differences in improving the fluxes in 2017 between using the entire period (2015-2020) and a single year (2018) for calibration. Although using the entire period resulted in a slightly higher RMSD reduction for the three variables with the GA method, the BFGS method yielded superior performances when using only 2018 for calibration. This may seem counterintuitive, as additional information typically creates extra constraints, helping to smooth the cost function and making local minima less likely for BFGS. However, our findings can be explained by the fact that drought periods in 2018 are less predominant in a 6-year period, resulting in a less optimal solution for the calibration of the water stress parameters with the BFGS method.

4.2 Impact on phenology and soil water

We recognize that assimilating LST alone has its limitations and cannot enhance the model-data fit of all variables controlling water, energy, and carbon fluxes. To better understand the performance of the LST assimilation procedure on other variables less directly linked to energy fluxes, we assessed the impact of LST optimization on soil water availability and gross primary productivity (GPP). Since the number of sites with available soil moisture data are limited and measurement depths vary among sites, we evaluated the impact on the top 10 cm of soil moisture in our twin experiments. The soil moisture showed a clear improvement (positive median RMSD reduction) when assimilating the 3-hourly LST alone using the GA method. The median RMSD reduction for this experiment (3h-LST) represented an enhancement in soil moisture of 10.4%, although some runs among the 16 different first-guesses resulted in a deterioration of soil moisture. The fact that the 3h-LST

DA showed an overall improvement in soil moisture confirms the chosen strategy for the assimilation of the CCI-LST data.
Regarding the GPP, we assess the impact by assimilating the 3-hourly CCI-LST time series over the 34 WarmWinter sites in 2018. Assimilating the 3-hourly LST data improves the GPP in 14 sites, showing improvements in diverse conditions such as the grassland CH-Cha and Mediterranean ES-LS2 sites (see Fig. F1 in Appendix F). However, considering all the 34 sites, LST DA results in an overall degradation in GPP, with a median RMSD reduction across sites of -7.4%. Nevertheless, the improvements observed in the 14 sites (i.e., 41% of the sites) is a promising outcome, especially considering the challenge of enhancing model variables that are not closely linked to LST. In fact, assimilating a single data stream may even degrade the model simulations of other variables, as shown in Kato et al. (2013) and Bacour et al. (2015; 2023). In our study, since we calibrated only parameters impacting LST and kept the carbon-related parameters previously optimized without LST observations, a degradation of the carbon fluxes is not surprising. Nonetheless, the overall improvement of the energy fluxes, such as the 20.6% enhancement in LE and 9.6% in H, is significantly more impactful than the observed degradation in GPP.

Although improvements in soil moisture and phenology were not expected by assimilating only LST data, the enhancements found in both the twin experiment for soil moisture and some sites for GPP with real data are very encouraging. These results support ongoing efforts to jointly assimilate LST with satellite-derived products of leaf area index, albedo, and soil moisture into ORCHIDEE. Such an approach is expected to better constrain a wider range of energy, water, and carbon parameters, enhancing the overall performance of the model."

In the conclusion of the new version of the manuscript, we also added a paragraph considering the discussion added as follows:

"However, assimilating LST alone has limitations and cannot improve all variables controlling water, energy, and carbon fluxes. Nevertheless, our findings reveal promising outcomes, such as the clear improvement in soil moisture in the twin experiment and the enhancement of GPP in 41% of the studied sites. Despite the challenges, these results indicate that LST data can positively influence variables less directly linked to energy fluxes. This underscores the potential of combining LST with other satellite-derived products, such as leaf area index, albedo, and soil moisture, to better constrain and improve the overall performance of the ORCHIDEE model."

The corresponding appendices for the discussion in the new version of the manuscript were added as follows:

Appendix E – Impact of the optimization period

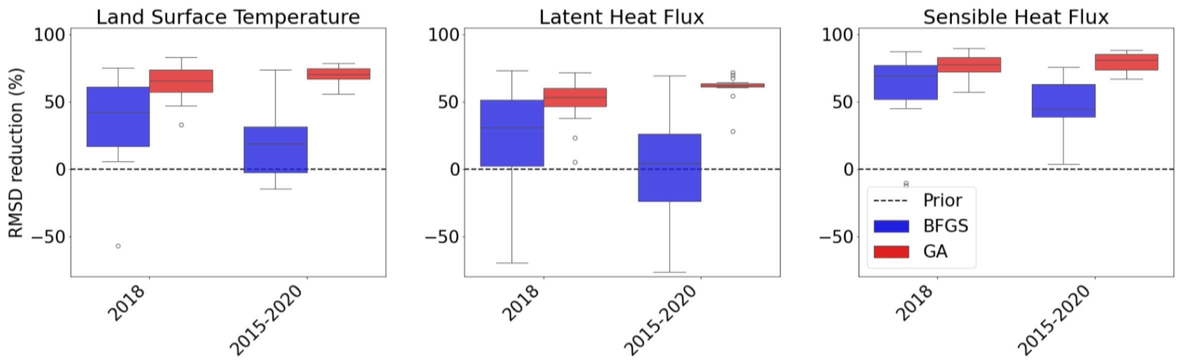

Figure E1. Comparison of model performance in 2017 for 30-min LST, LE and H when parameters are calibrated in 2018 only or for the entire period (2015-20220) over the selected site in Spain (ES-Abr). Boxplots obtained within 16 optimization tests with random first-guess parameter values for the DA experiment using the gradient-based (in blue) and genetic (in red) methods in terms of model–data RMSD. The DA experiment assimilates the daily mean, amplitude and maximum LST.

Appendix F – Impact of assimilating LST on phenology

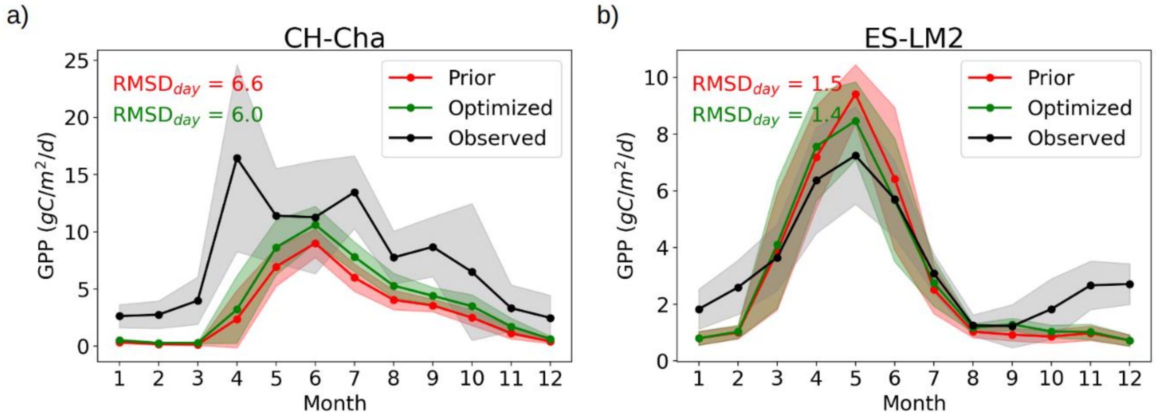

Figure F1. Annual cycle of GPP modeled for 2018 over a grassland (a: CH-Cha) and cropland (b: ES-LM2) site. The mean (dot) and standard deviation (shaded area) are represented for in situ observations (black), Prior (red) and Optimized (green) ORCHIDEE simulations. The RMSD on the daily basis ($RMSD_{day}$) against in situ observations is shown for Prior (red) and Optimized simulations.