Review #4

The manuscript presents a computationally efficient approach for representing crop canopy temperature in LPJmL using polynomial emulators derived from a stability-corrected energy-balance model. The topic is relevant to large-scale crop modeling, and the study is generally well organized and clearly presented. The emulator shows good agreement with the parent model while adding only limited computational cost. However, I have several concerns that should be addressed before publication.

Thank you very much for the positive and thorough evaluation. In the following we give a point by point response including suggested changes for improvement of the manuscript.

1. The coupling between the emulator and LPJmL requires further clarification and validation. The final canopy temperature is obtained by interpolating between the upper and lower temperature limits using a water-stress factor derived from LPJmL canopy conductance. However, the canopy conductance calculated by LPJmL and the optimum conductance used to define the lower temperature limit may not be fully consistent in their definitions or scales. Since this scaling directly controls the final canopy temperature, the authors should provide a clearer justification for this formulation and evaluate its sensitivity. It would also be helpful to clarify which physiological processes are driven by canopy temperature and which remain driven by air temperature.

Thank you very much for this helpful comment. We agree that the water-stress scaling factor K_WS needs further evaluation and we will conduct a respective analysis (see also response to reviewer 2, point 2.). Furthermore, we will describe the underlying idea behind the formulation of the water stress factor more thoroughly and discuss potential limitations. Also, we will provide more information about the temperature-related processes implemented in LPJmL and whether they take 2 m air temperatures or canopy temperatures as inputs.

2. The conclusions regarding the improved representation of high-temperature conditions appear stronger than the current validation evidence supports. ERA5 skin temperature is not a direct observation of crop canopy temperature because it integrates the thermal characteristics of multiple land-cover components within each grid cell. Therefore, better agreement with ERA5 skin temperature does not necessarily demonstrate improved realism in simulated crop canopy temperature. The authors should interpret this comparison more cautiously and clearly distinguish between an improved representation of land-surface temperature extremes and a direct validation of crop canopy temperature. In addition, the practical relevance of the new module would be clearer if the authors demonstrated whether canopy-temperature forcing affects simulated heat stress, phenology, biomass accumulation, or yield.

We appreciate this valuable comment. In our study, we aim to reduce the mismatch by only selecting cells with certain minimum shares of the respective crops being evaluated. Nevertheless, as you point out, we introduce an error that is weakening the strength of the evaluation. We agree that this is not sufficiently discussed in the current version of the manuscript and we will provide a clearer discussion and a more careful interpretation of the respective results in the revised version (see also responses to reviewer 1, point 1., and to reviewer 2, point 6.).

We also agree with your second point that assessing effects of canopy temperature forcing in temperature-dependent processes, such as heat stress, would be very interesting. That is why we are already working on a follow up study that builds upon this one. We think that including such a respective assessment would be beyond the scope of this manuscript, particularly as it would require more model description regarding recent LPJmL developments (see also responses to reviewer 1, point 2., and to reviewer 3, point 1.).

3. The workflow contains several sequential sources of uncertainty, including the parent EBSC model, polynomial emulation of the temperature limits, interpolation between these limits, and the LPJmL-based water-stress scaling. At present, these uncertainties are evaluated mainly through their combined effect, making it difficult to identify which step contributes most strongly to the final prediction error. A quantitative error-attribution or sensitivity analysis would substantially strengthen the methodological evaluation. Such an analysis would help readers assess the robustness of the full workflow and identify priorities for future model development.

Thank you very much for the valuable comment. The first component, the EBSC model, is an independent model that computes the energy balance at the crop surface based on the inputs it receives. In this study, it is not embedded in a crop model, which is why it lacks any information on canopy conductance or transpiration rates. The same holds true for the emulators. Hence, neither of these two components can compute a water stress factor, which is needed to compute final canopy temperatures and to analyse the cumulative error throughout the model chain. That is why the workflow does not provide a basis to analyse how and when errors accumulate. Instead, we will evaluate the errors of the single components more comprehensively. Particularly, we will conduct a sensitivity analysis to challenge the assumptions we use to compute daily maximum canopy temperatures with the EBSC model (first component, see also response to reviewer 1, point 4.) and we will evaluate the water scaling factor K_WS in LPJmL (third component, see also responses to reviewer 2, points 2. and 3., and to reviewer 3, point 3.).