Review of Ensemble-based global fire modelling as a tool to characterize extreme wildfire events by Ribeiro et al.

Reviewer: João Teixeira

This study investigates the use of large ensemble climate simulations to drive the LPJmL-SPITFIRE dynamic global vegetation-fire model to better characterize extreme wildfire events under present-day climate conditions. The authors compare fire simulations driven by a 40-member Single-Model Initial-Condition Large Ensemble (SMILE), a single ensemble member, and reanalysis data, and assess their ability to capture extremes in fire danger index (FDI), burned area (BA), and fire carbon emissions (FireC). The results demonstrate that large ensembles sample a broader range of physically plausible wildfire extremes than single climate realizations and provide new insights into the relationship between extreme fire weather and fire impacts.

This manuscript addresses an important challenge and presents a valuable contribution to wildfire risk assessment and presents a novel application of large ensemble climate-impact modelling to process-based global fire simulations. The methodology is clearly described, and the manuscript is generally well written and well structured. However, the discussion would benefit from a deeper interpretation of the mechanisms driving the simulated extremes and a clearer assessment of the limitations associated with the experimental design, particularly regarding the treatment of fire ignitions and the interpretation of regions where deterministic simulations exceed the ensemble maxima. Subject to these revisions, the study would provide a useful contribution to the literature on wildfire extremes and large ensemble climate-impact modelling.

Major Comments

The manuscript clearly demonstrates that the single-fire and reanalysis-fire simulations generally undersample the most extreme FDI, BA, and FireC compared to the SMILE-fire simulations. However, the discussion focuses primarily on that undersampling occurs rather than what drives it. While the broader sampling of internal climate variability is presented as the main explanation, the manuscript also shows that extreme BA and FireC are not solely controlled by FDI. I encourage the authors to further discuss the processes driving the undersampling, particularly given that BA and FireC exhibit substantially larger differences than FDI. Similarly, the regions where the single-fire or reanalysis-fire simulations exceed the SMILE-fire maxima (e.g., NHAF, CEAM, and BONA) deserve further discussion, as these exceptions may provide additional insight into the processes controlling extreme fire events and the limitations of the ensemble framework.

The experimental design prescribes lightning ignitions from the OTD/LIS climatology while only the meteorological forcing varies across the ensemble members. Consequently, natural ignitions are not dynamically consistent with the simulated weather, and climate-driven variability in lightning occurrence is not represented within the ensemble. I encourage the authors to discuss the implications of this assumption more explicitly, particularly regarding its influence on the interpretation of the simulated extremes and the extent to which the ensemble captures climate-driven variability in wildfire occurrence.

Minor Comments

L22–23 – Consider replacing "global warming" with "climate change", as the latter is more appropriate in this context given that wildfire activity is influenced by multiple climatic drivers, not only increasing temperatures.

Section 2.3 – Given that the manuscript focuses on extreme wildfire events, it would be useful to briefly discuss the ability of LPJmL-SPITFIRE to reproduce the upper tail of the observed BA and FireC distributions, in addition to the mean variability presented in Figure 2.

Figures 3 and 4 – The information shown in the upper maps is largely redundant with the regional boxplots below, as both present the regional values of $\Delta_{\max}$. Instead, I suggest displaying the grid-cell values of $\Delta_{\max}$ in the maps. This would better illustrate the spatial variability within each GFED region and provide additional insight into where the largest differences between the single-fire and SMILE-fire simulations occur.

Figures 3 and 4 – The current map projection makes some regions difficult to visualize due to the wrap-around at the map boundaries. Consider using a projection such as Robinson, which would improve the readability of the global spatial patterns.

Figure 4 – The caption states "Same as Figure 4 but for the reanalysis-fire." This should read "Same as Figure 3 but for the reanalysis-fire."

L242–245 – The discussion of the p97.5 and p99 thresholds would benefit from a brief reminder that these return periods are estimated from the pooled ensemble and should not be interpreted as true return periods.

L322–329 – The discussion of the limitations of the VPD-based FDI is useful. It may also be worth briefly acknowledging that it does not explicitly account for wind, despite wind being one of the primary controls on fire spread.

L352 – "furter contribute" should be "further contribute."