

## Review 1

Thank you for submitting the revised manuscript. The reviewer is generally satisfied with the revision, and appreciates the effort. They provided one comment for you to consider for your final submission: "My only remaining suggestion is that the authors add a brief discussion in the Introduction, perhaps around Line 77, on the importance of physical alignment between saliency patterns and the synoptic factors controlling local extreme weather. This would further highlight the authors' effort to push saliency-based XAI beyond purely diagnostic sensitivity analysis and toward physically interpretable evaluation of learned relationships. It would also help clarify why the added case-wise and physically grounded saliency analyses are critical to the manuscript's contribution."

*We thank the reviewer for the positive assessment of our revised manuscript. We are also grateful to both reviewers for their constructive comments throughout the review process, which we believe have substantially improved the quality and scope of this work. To express our gratitude, we have added the following sentence to the Acknowledgements section:*

*"We are also grateful to the two anonymous reviewers, whose constructive comments substantially improved the quality and scope of this manuscript."*

*Regarding the specific comment raised in this round, we agree that the Introduction would benefit from explicitly motivating why physical alignment between saliency patterns and the synoptic factors driving local extremes is a meaningful criterion for evaluating learned relationships. Accordingly, we have added the following paragraph to the Introduction:*

*"To assess the consistency and interpretability of the fine-tuned models, we analyze the learned relationships using eXplainable Artificial Intelligence (XAI) techniques (Adadi and Berrada, 2018; Arrieta et al., 2020; Minh et al., 2022). Beyond quantifying which predictors a model relies on, in this work we also emphasize whether the learned relationships aligns with the synoptic processes that drive relevant local events (such as extreme precipitation), for which there is a physically motivated expectation of where a skillful model should attend. This alignment provides a more demanding test of the trustworthiness and transferability of the learned relationships since a model can achieve competitive metrics while relying on spurious relationships that are especially problematic under the extrapolation to future climate (González-Abad et al., 2023). In this sense, our analysis pushes saliency-based XAI beyond diagnostic sensitivity analysis toward a physically interpretable evaluation of the learned relationships, motivating the case-wise saliency analyses presented in this work."*

*We believe this addition clarifies the motivation behind the case-wise analyses, framing them as a step toward a physically interpretable evaluation of the learned relationships rather than a purely diagnostic sensitivity analysis.*