I sincerely thank the reviewers for their valuable work. The results are clearer and have improved the overall quality of the paper. However, I still have major comments that were already raised in my previous review but have not been addressed, even though I believe they are important to consider.

1. Regarding the notion of extreme events, I maintain the same view on the use of this term in the text. I acknowledge that defining EPEs based on the 90th or 95th percentiles has been widely used in Antarctic literature, and I have also contributed to studies that adopted this terminology. However, I consider that this was an imprecise use of language, and it would have been more rigorous to acknowledge this limitation at the time. Referring to these events as “extreme” in previous publications was (1) reflecting a lack of rigor and (2) have significant implications for how results are interpreted and cited.

1) Regarding the lack of rigor, my concern relates to the underlying probability distribution of the data and the associated mathematical assumptions and conclusions. Referring to precipitation exceeding the 95th percentile as “extreme” implicitly undermines the relevance of a GEV-based framework. Would one consider a Gumbel distribution to be equivalent to a Gamma or a normal distribution? The answer is no; consequently, this issue should at least be acknowledged as an imprecise use of terminology.
2) If precipitation exceeding the 95th percentile is classified as “extreme,” then subsequent studies may rely on this definition and claim that we already have a solid understanding of extremes. However, I would argue that the March 2022 event in Antarctica clearly showed that truly extreme events behave as outliers, with causes and impacts that are fundamentally different from those of events above the 95th percentile. Consequently, our current understanding of extremes in Antarctica remains very limited, which makes it all the more important to use precise terminology and to clearly distinguish between *intense* events and genuinely *extreme* ones.

Regarding the use of the term “extreme” in your study, I do not think that the reported trends reflect changes in extreme precipitation in Antarctica. To support my point, I downloaded ERA5 precipitation data for Antarctica and computed trends in annual maxima (This corresponds to a 2-year return level, i.e., the 99.86th percentile, meaning one event every 2 years, which is well beyond the “extremes” considered by the authors) and in the 10-year return level of daily precipitation using a non-stationary GEV framework (*Figure R1*). I applied a 5% significance level, which differs from the 10% significance level used in your Figure 2. Note that the scale is expressed in mm decade^{-1}, as I analyze the trend of individual events rather than the total precipitation from events exceeding the 95th percentile (i.e., approximately 18 events per year). Hereafter, I will refer to precipitation above the 95th percentile as EPEs, to clearly indicate that I am referring to the definition used in your paper.

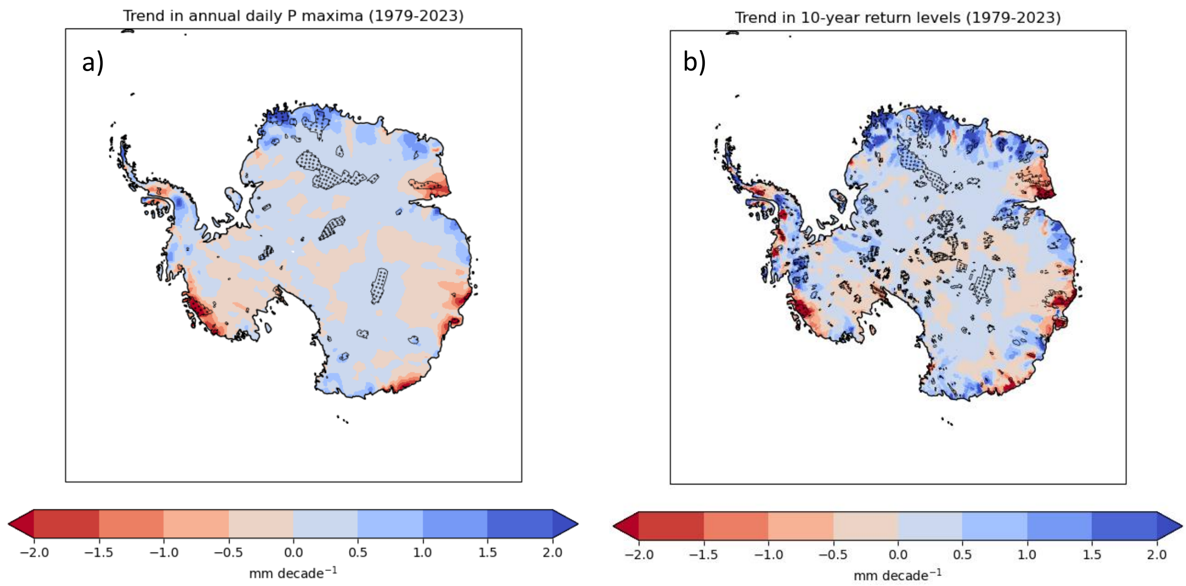

Figure R1: a) trends in ERA5 annual precipitation maxima over 1979-2023, b) same as a) but for 10-year return level precipitation. Stippling indicates where trends are statistically significant at the 95% confidence level.

In spite of an overall similar pattern, there are key differences. In your Figure 2a, regions showing strongly negative trends in mean precipitation, such as Wilkes Land, were progressively shifting toward positive trends when considering EPEs (Figure 2f). When examining annual maxima, and even more so the 10-year return levels, this shift becomes clearer. Conversely, regions such as Enderby Land or the Antarctic Peninsula shift from increasing trends to decreasing trends.

This is particularly concerning because these changes occur in the I–Ipp, J–Jpp, and especially Ap–B catchments, suggesting that trends derived from EPEs in key basins of your study are unlikely to be representative of true extremes. These shifts are actually indicating that the entire probability distribution is evolving, with notably different behaviors in the upper tail. Moreover, the F–G basin (along the coast of Marie Byrd Land) shows widespread regions with significantly negative trends, suggesting that the basin-scale trend may be statistically significant.

In my first review, I was only expecting a clarification in the text and an adjustment in wording, not a complete revision of the study. Again, I am not requesting a major change to the study. I am not suggesting a completely new analysis based on annual maxima or return-period events (e.g., 10-year events). I am simply asking that the manuscript clearly states that precipitation above the 95th percentile should be described as intense rather than extreme.

A sentence such as the following would be sufficient:

“In this paper, we consider intense events as those exceeding the 95th percentile, which have generally been referred to as extreme precipitation events in the previous literature. For simplicity, we will refer to them as EPEs, although, for the sake of scientific rigor, we note that these events should be considered intense rather than extreme.”

Then, the definition of EPEs needs to be clarified in the very first sentence of the introduction. In the current manuscript, it is only explicitly defined in the Methods section, which makes several statements in the introduction unclear or misleading. For instance, the definition given in the first sentence of the introduction as: “Extreme precipitation events (EPEs, i.e., where the amount of precipitation at a location is substantially greater than normal)" is ambiguous. What exactly is meant by "substantially"?

Similarly, the statement "EPEs account for over 40% of the continent's total precipitation and almost 70% of the variance in annual precipitation" relies on the definition from Turner et al. (2019), but this is not clearly stated upfront.

I also suggested revising sentences such as: "EPEs over recent decades in Antarctica is hampered by such events being rare and difficult to separate from the substantial interannual precipitation variability." Here again, the term "rare" is not appropriate for events defined as exceeding the 95th percentile.

If EPEs were clearly defined in the first sentence of the introduction, it would immediately establish that you are using something similar to Turner's definition, and the subsequent statements would become coherent and meaningful, for example:

"Positive trends in the number of EPEs occurring over recent decades have been identified over parts of East Antarctica" or "Antarctic precipitation and EPEs are the major component of surface mass balance." Without an explicit definition at the beginning, these statements lack precision and may even be misleading or incorrect.

2. Overall, the manuscript is primarily a D&A study and already contains substantial and valuable material for discussion, including uncertainties, and discrepancies between CESM1 and CESM2, as well as comparisons with previous studies. In my view, the abstract would be better concluded by referring to the differences between CESM1 and CESM2, which are more robustly supported by the analysis presented. I believe the discussion should also be slightly extended and focus more on these aspects, in particular on the robustness of the results.

Indeed, current final sentence of the abstract (referring to ARs) is based on conclusions that are not sufficiently robust and should only be treated as a discussion point. ARs currently represent only a secondary and still debatable element of discussion and key questions remain unresolved on AR impact: do AR precipitation intensities increase more than non-AR precipitation? Or could the apparent change simply result from the thresholds used in Wille's ARDT methodology? Concluding the abstract on this point would require a more solid basis.

Indeed, in my first review I suggested computing trends in AR intensity during AR days and non-AR days, and then comparing both, because the current figure does not allow us to draw robust conclusions about the actual role of atmospheric rivers in EPE trends. Here, you assume that $\sum(EPE) = \sum(EPE_{ar}) + \sum(EPE_{nonar})$. Then if the sum of EPE increased by a rate of α in mm y^{-1}, and if $d(\sum(EPE_{ar}))/dt = \alpha$, then $d(\sum(EPE_{nonar}))/dt = 0$ implying that 100% of the change is attributed to ARs. While mathematically correct under this decomposition, this does not allow one to conclude on the actual contribution of ARs, because the number of EPE_{ar} and EPE_{non-ar} events is itself evolving over time.

Indeed, let us consider that the intensity of both EPE_{ar} and EPE_{non-ar} changes at the same rate. In that case, all events would be intensifying, not only those associated with atmospheric rivers. In your calculation, both $\sum(EPE_{ar})$ and $\sum(EPE_{nonar})$ would increase. However, if you also suppose that the number of AR events increased over time. In this situation, the changes in $\sum EPE_{non-ar}$ could decrease or remain constant, while $\sum EPE_{ar}$ would increase even more strongly, even though the intensity of both types of events evolves similarly. One could then conclude that the observed change in EPEs is driven by an increase in the number of AR events rather than by changes in their intensity. However, this interpretation would still be misleading, because non-AR events would be experiencing the same underlying changes in intensity. The actual driver of the signal would therefore not be atmospheric rivers per se, but rather a general increase in atmospheric moisture transport affecting all events. In the present example, since both AR and non-AR event intensities would increase at the same rate, the difference in $\sum EPE_{ar}$ and $\sum EPE_{non\text{-}ar}$ trends would primarily reflect how events are classified by the ARDT method rather than a true change in the physical properties of the two categories. Actually, this is possible because Wille's ARDT algorithm mechanically classifies a greater number of moist atmospheric events as "AR" under conditions of overall increasing atmospheric moisture. This can generate an apparent AR-related signal even if the underlying change is not specifically driven by atmospheric rivers themselves.

In my first review, I already noted that the authors did not assess whether the observed changes in trends result from changes in frequency, intensity, or a combination of both. The authors' write "Following the reviewer's suggestion, we also tested an explicit decomposition of ERA5 precipitation into AR-associated and non-AR-associated components. In this approach, ERA5-based total precipitation is expressed as the sum of (i) AR-associated total precipitation and (ii) non-AR-associated total precipitation, where non-AR-associated precipitation is defined as total precipitation minus AR-associated precipitation.". My comment was unclear, because additional calculations were not what I was expecting and do not resolve this issue.

At this stage, I suggest analyzing whether trends in precipitation intensity differ between the EPE_{ar} and EPE_{non-ar} subsamples. For this task, the authors have to compute the trend considering the mean intensity of EPE event and not the sum of EPE events. Atmospheric rivers could be considered a driver of the observed change only if non-AR events present (significantly) weaker trends. It would thus be necessary to test the significance of EPE intensity trends during AR days and non-AR days using a bootstrap approach.

3. Figures 2 and S1 show substantial differences between CESM2-AAER and CESM1-AER, and this point deserves much more discussion in the manuscript. In particular, while I understand that the sum of trends from GHG, AER, and O_3 forcings is not expected to reproduce the trends from the ALL simulation, the spatial patterns in CESM1 appear remarkably similar across all forcings. What could explain this? This is especially puzzling given that trends associated with AER in coastal regions of Antarctica are clearly distinct from those retrieved over the Southern Ocean in Figure 4 of Deser et al. (2020), which describes the CESM1 Large Ensemble. Exploring, and ideally explaining, the origin of these differences and those with CESM2 would be highly valuable for the community, as it could provide important insights into the limitations of detection and attribution analyses in Antarctica.

Reference : Deser, C., Phillips, A. S., Simpson, I. R., Rosenbloom, N., Coleman, D., Lehner, F., ... & Stevenson, S. (2020). Isolating the evolving contributions of anthropogenic aerosols and greenhouse gases: A new CESM1 large ensemble community resource. Journal of climate, 33(18), 7835-7858. https://doi.org/10.1175/JCLI-D-20-0123.1

4. After reading the initial version of the manuscript, I did not understand that the EPEs were computed based on "precipitation days," defined as days with precipitation exceeding 0.02 mm day^{-1}. I thank the reviewers for their response, as they clarified a point that I had not noticed during my first reading, even though it was already stated in the paper. I apologize for having missed this detail.

This results in the use of a "location-dependent" percentile used to define extremes since the 95th percentile is computed over different number of days depending on the location. I find this approach debatable, as this threshold is excluding more days inland than at the coast. As a result, basin-scale precipitation also becomes "disproportionately" influenced by coastal events, rather than reflecting the inland signal. It also implies that the number of days considered at each grid point, varies with the forcing and across simulations, which complicates the interpretation of trends.

A second issue concerns the methodology used to define 95th percentile precipitation at the basin scale. You identify precipitation days exceeding the 95th percentile at each grid point, and then aggregate these values annually to compute a basin-averaged "extreme" precipitation. This is not consistent with standard hydrological practice, where extreme precipitation is typically defined at the basin scale to allow meaningful comparisons with extreme discharge. From a hydrological perspective, each precipitation event should be treated independently rather than merged with others. I would suggest to first compute daily basin-averaged precipitation, then identify days when this basin-averaged precipitation exceeds its 95th percentile, and finally sum these values over the year. Even if this should not fundamentally change the conclusions, this would provide a more physically consistent probability distribution of precipitation at the basin scale.

Minor comments

Line 116 vs. line 395: in the document and in the Supplementary material the term AAER is still used for CESM2.

Line 221 : "Over the Antarctic Peninsula, the ERA5 trends in total and extreme precipitation differ in direction, with total precipitation increasing and extreme precipitation decreasing, although only the trends over the northern Antarctic Peninsula are significant."=> if I am not wrong, extreme precipitation is increasing in the I-Ipp region.

Line 387 : "indicating that greenhouse gas forcing is the primary driver of the positive trends in the CESM2 ALL ensemble, consistent with CESM1." => In lines 344–363, it is stated that the scaling factors are generally associated with relatively small confidence intervals for all forcings, including GHG. Is this in contradiction with the present sentence?