The reviewers have answered all my comments and clarified parts of the methods section which makes it more readable and understandable. The additional material in the supplementary information also helps the interpretation of the analysis. Overall, the manuscript is in a good shape. While reading the manuscript again (with a better understanding of the method) I got the following question (apologies for not having asked this in the previous round):

Do we expect the scores of the SST component 1 (ENSO) pattern from ERA5 to be correlated with free running climate simulations? For component 2, the authors argue that it represents mainly a forced warming signal and therefore should be similar in ERA5 and model runs (l149). I see how for component 2 this could be meaningful. For component 1, the authors write "This suggests that in the first half of the timeseries, El Niño conditions were dominant, while in the second half of the time series, La Niña conditions were." (l144). I would argue that this reflects natural variability. Therefore, I'm wondering why the correlation of the scores of component 1 between ERA5 should be correlated to the corresponding scores of a freely running climate model simulation. Does the correlation of scores really inform about the skill of that climate model to represent the dynamics. Wouldn't it be more meaningful to calculate the NRMSE score only for the loadings (ignoring the scores)?
Thanks for the insightful comment. Yes, we expect the scores of the simulations to match ERA5 because the simulations are not freely running; the simulations are from AMIP and therefore forced with the observed SST. Thus, it is a good sign if the model timeseries and reanalyses timeseries are in agreement, even for natural variability modes; this suggests that the atmospheric model is skillful in translating these SST patterns into precipitation.

Further suggestion: The authors made a number of methodological choices (0.6 NRMSE threshold, threshold must be met by at least 2 components, considering the first 4 components, ...?). While I see the necessity of such methodological choices, I would appreciate a little sensitivity analysis where some of these choices are altered and the results from figure 6 are shown again. Alternatively, the choices could be listed and shortly discussed in the discussion section again.
Thank you for the suggestion; we agree it would be ideal to test the sensitivity of these choices. Unfortunately, this falls outside the scope of the current project, which was funded between 2022 and 2023. However, in the revised version, we did provide a brief discussion of these choices in lines 108-111 and again in the conclusion section (l. 288-290).

L57: "latents variables"
Thank you for pointing this out. We have corrected the revised version.

L73: "For instance, if the SST signal is negative at a particular location while the score is positive, this indicates a negative association with the corresponding precipitation loading pattern." Thanks for adding this explanation, i think it's useful. I think this example should be even more specific. What has a "negative association with the corresponding signal"? And how does the time dimension come in? This seems important to me as the scores change sign over time in your example. Would the following be correct? "For instance, if the SST loading is positive in the ElNino region while the score is negative in a given year, this indicates that in that year, warm SSTs in the El Nino region are negatively associated with the corresponding precipitation loading pattern."-

We thank the reviewer for the suggestion and improved the explanation between L.73 and L75.

L101: "were separately calculated and after that, used a normalisation." -> as normalization?
In the last revision we modified this section and believe it is now written in a clearer and more detailed way. With this, we believe that the use of the term 'normalization' is now properly contextualized and aligned with the intended meaning we want to convey.