Second review of Fischer et al. (egusphere-2024-1253)

Despite answering in detail to most of my comments, the authors did not fully understand the main issues I reported. I respect the authors' original methodological choices, but I still think some of the requests made in the first round of revision are legitimate and can be tackled without modifying the general approach. This does not change my overall opinion of the paper, which I consider an excellent piece of work and worth of publication. Nevertheless, I would like these two points to be more carefully addressed in the final version.

- Significance. I do not agree that performing significance tests with ensemble data is not meaningful since "there is no limit in generating more ensemble members", nor that "even tiny differences would appear as significant": in general, we do not know which changes are significant and which are not before performing a statistical test. It could well be that most of the changes are indeed significant, but this can't be taken for granted without a proper calculation.
The test I proposed is quite simple: take the ensemble standard deviation of 10-year averages in the historical period and, if the future-hist difference at a specific point is larger than that, that is an indication that the change is significant. I would apply this to regime composites of the change in Figs. 4,5,6,7: is the future-hist change larger (in absolute value) than the historical standard deviation of 10-year averages? Of course, this is only one possible measure of significance, but alternative ways are also possible. I don't think that this repeats Yettella and Kay (2017), since they did not investigate regime-specific changes.

- Interpretation. Your claim is quite strong, namely that regime frequency changes do not matter for the understanding of the impacts of future circulation changes on seasonal precipitation. However, I think the results you show are tightly linked with some of your methodological choices, and this should be discussed more in depth. In particular, these two choices impact the value of gamma:

- since you separately removed the climatology on the historical and future ensembles (also, the calculation of the normalization factor is done separately), part of the climate change signal on the circulation is removed, thus reducing the amplitude of regime frequency changes;
- on the other side, you do not remove the mean climatological change in the precipitation fields, so that regime intensity changes contain that signal (most of which is independent from the regimes, as you observe).

As far as I understand, the first choice reduces the "frequency effect", while the second enhances the "intensity effect". This is not to question your methodological choices - which are legitimate and are always to some extent "subjective" -, but to better inform the community that it is actually quite complicated to disentangle the two effects, and the results you get depend - at least to some extent - on some of your assumptions.