We are grateful for your thoughtful and constructive comments, and we hope that the responses provided below address the remaining queries with sufficient clarity.

SD Comments v2:

The authors have adequately addressed the comments made during the first round of peer review, and the manuscript has been substantially improved, so I would recommend that it be accepted after some minor corrections, mostly related to the ensemble mean and the use of the EURO-CORDEX hindcast simulations, which I think is important to clarify for other studies.

Minor comments:

The author claimed that the ensemble was constructed only from simulations that provided the parameters necessary to construct the indices described in Table 2. However, to my knowledge, the EURO-CORDEX data are consistent in the variables they provide and their spatial resolution (0.11°), which means that all simulations participating in the hindcast experiment provide all the variables listed here: http://is-enes-data.github.io/CORDEX_variables_requirement_table.pdf from which the authors could calculate the indices.

We appreciate the reviewer's observation regarding the expected consistency of EURO-CORDEX simulations in terms of variable availability and resolution, in line with established CORDEX protocols. However, despite these formal requirements, there were practical limitations at the time we initiated the analysis. Specifically, several ESGF nodes were offline, making it impossible to access some simulations (for instance, the current issue with RegCM4-6 outputs, as discussed below).

Given these accessibility constraints, we proceeded with the subset of simulations that were readily available at that time. As the issue persisted for a while and considering that the current study is intended as a proof-of-concept, the primary objective was to demonstrate the effectiveness of the methodology rather than to exhaustively include all possible model runs. Expanding the ensemble would likely not have significantly altered the results, nor would it have influenced the conclusions of the study.

For these reasons, we respectfully maintain that it is not necessary to expand the ensemble further at this stage.

Also with a quick search on the Earth System Grid Federation https://esgf-metagrid.cloud.dkrz.de/search I could not find the RegCM4-6 data.

The RegCM4-6 data is listed via the ESGF LiU server hosted at https://esg-dn1.nsc.liu.se/search/cordex/, although it may currently be experiencing accessibility issues. It is likely that the RegCM4-6 data is temporarily unavailable due to disruptions at the ESGF-ICTP node (esgf-ictp.hpc.cineca.it), which was hosted on the Marconi system at the CINECA supercomputing centre. This system sustained significant damage during the severe flooding in Bologna on October 20th, 2024, resulting in prolonged outages for both computing and data access services. We were informed that this is a temporary issue and expect data access to resume but have no information regarding the timeline.

I think it's very important to clarify what criterion was used for a model to be included as a member in the ensemble.

In light of the comments above, we have clarified the sentence in the manuscript describing the ensemble selection by explicitly stating that the included models were also those "available through ESGF nodes at the time data collection began."

Finally, I don't understand why there were two models, the CNRM-ALADIN63 and the ICTP-RegCM4-6, that needed interpolation, since all EUROCORDEX hindcast data are provided at the same spatial resolution as mentioned above.

We acknowledge the reviewer's observation and agree that all EURO-CORDEX simulations conform to a nominal resolution of 0.11°. However, modelling institutions retain flexibility in choosing the projection and some domain configuration best suited to their models. As a result, while all simulations cover the common EURO-CORDEX domain, some differences (e.g., projection and larger grid sizes) may be sometimes present. In these cases, interpolation was required to ensure consistency across grids for subsequent analysis. To minimize errors, the indices listed in Table 2 were calculated individually before any interpolation, as stated in the manuscript.

It seems that the later models are not driven by ERA-interim, but by general circulation models, and if this is the case, then the ensemble model should be rebuilt using only hindcast simulations.

Indeed, there is an important distinction between simulations driven by reanalysis datasets such as ERA-Interim (referred to as evaluation runs) and those driven by general circulation models (GCMs), which include historical and future climate projections.

For the purposes of this study, we specifically selected ERA-Interim-driven simulations to ensure consistency across the different data-sets and to maintain alignment with observation data. Reanalysis-driven simulations are generally preferred for evaluation studies, as they are constrained by observations and therefore offer a more realistic representation of the historical climate compared to GCM-driven simulations, which reflect internally consistent but divergent "mirror world" realizations.

Given the aims of this work and the nature of the analysis, rebuilding the ensemble to include only GCM-driven simulations would not be appropriate and would reduce the comparability and reliability of the results in relation to observed climate variability.

#1: line 159: "The observations are the 30-year (1980-2010)" Please change to 31-year everywhere in the text.

Changes made as suggested.