

This study is a first attempt to mobilize long transient regional climate simulation to understand past climate changes. The study focuses on changes in precipitation and mean 2 m temperature within the Eastern Mediterranean (EM) and Nile River (NR) basin regions from 500 BCE to 1850 CE. The authors evaluate the RCM's capability to simulate homogeneous regions during winter (DJF) and summer (JJA) under current climate conditions (1980-2018), then compare precipitation and 2 m temperature between the Roman period (ERP: 400-362 BCE) and the pre-industrial period (PI: 1800-1850 CE), as well as changes in large scale teleconnection patterns.

I am convinced that dynamical downscaling is a valuable tool to refine coarse GCM outputs to improve our understanding of past climate changes, especially in regions with diverse topography. While the scientific significance is excellent, the methodology used to identify homogeneous regions is not straightforward and the presentation quality deserves to be improved by selecting a subset of results and further highlighting the main take-home messages. Below are a couple of major and minor comments that require consideration for the paper's acceptance for publication in *Climate of the Past*.

Dear reviewer, thank you very much for the comments and suggestions. We have revised the manuscript accordingly, and please see the following responses for details. In the evaluation part, we have changed the comparison to the datasets ERA-Interim, CRU and CCLM. ERA-Interim is the driving data of the CCLM because the MPI-ESM – LR simulation is only available to the year 2000.

### Major Comments:

- While the study focuses on winter and summer, the most contrasting seasons, the primary changes between ERP and PI predominantly occur in spring and late summer-fall, especially concerning the mean state (Fig. 7). Including variability (standard deviation for each month for each period and inter-period difference) would provide a more comprehensive understanding of the changes. The authors are encouraged to consider analyzing or discussing the main modes of variability and associated large-scale teleconnections for these seasons as well.

For showing the variability for each month, each period, and each region, as well as comparing the CCLM output to the forcing MPI-ESM, the Box plots are shown on top of the mean annual cycles in Figure R1. Additionally, the associated large-scale teleconnections for four seasons are calculated and plotted and shown in Figure R2 to R5. As illustrated in Figure R1, showing the annual cycle of precipitation and temperature across regions EM, NR, and EMNR, a distinct contrast between the Early Roman Period (ERP) and the pre-industrial (PI) periods is noticeable, particularly from August to January in the case of precipitation. Interestingly, the spatial mean of temperature exhibits minimal differences between the two periods, except during the summer months (July, August, and September).

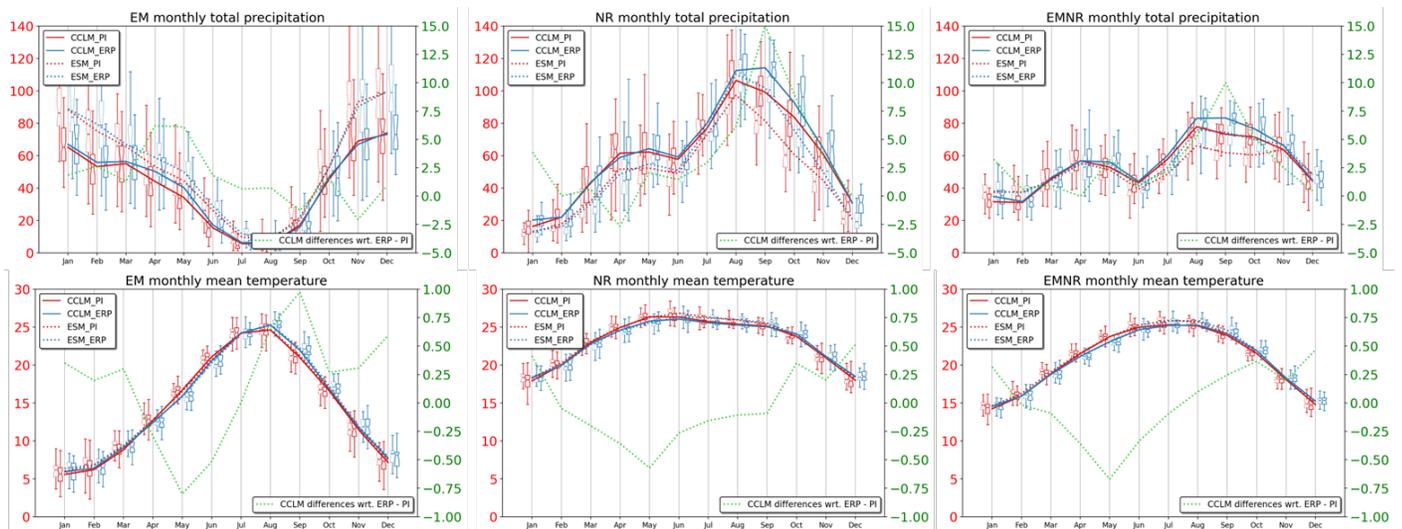


Figure R1. Precipitation and temperature annual cycle of the PI and ERP periods for EM, NR and EMNR. Monthly mean total precipitation and monthly mean temperature are presented on the left y-axis. Differences between ERP and PI are shown on the right y-axis.

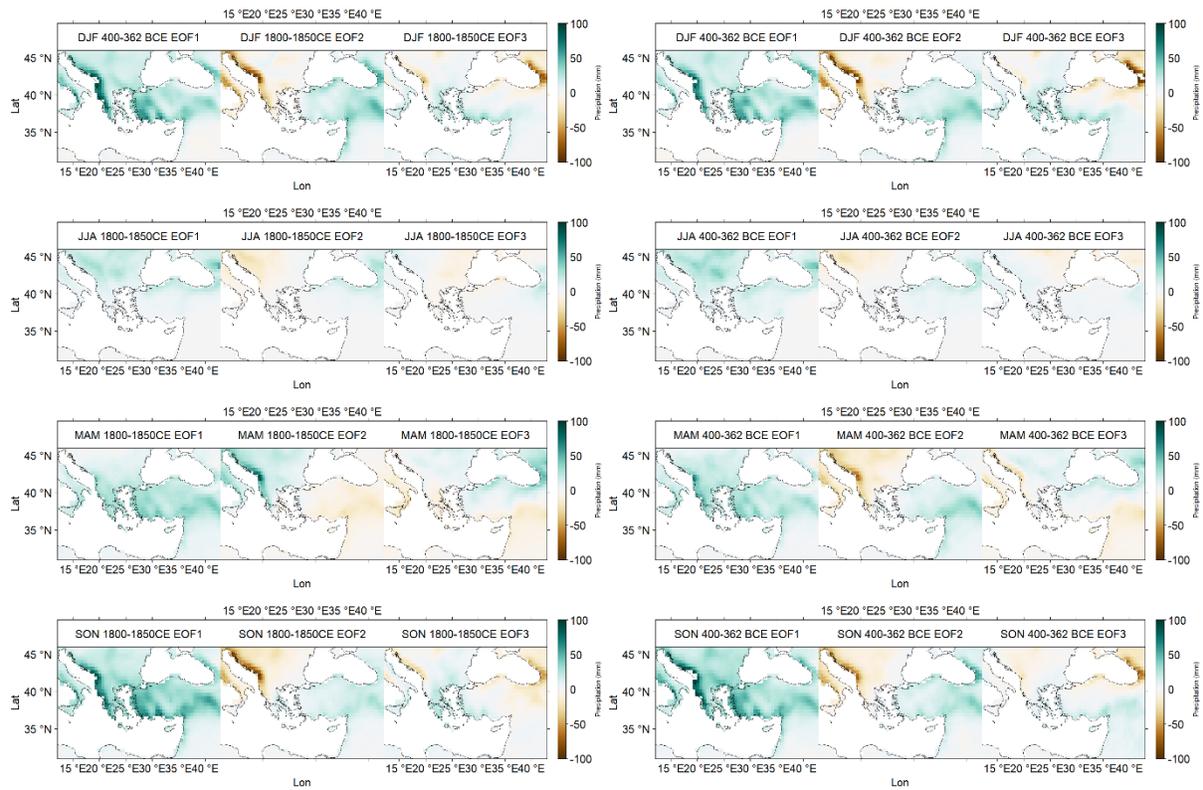


Figure R2. Patterns of the first three non-rotated EOFs of four seasons (DJF, JJA, MAM, SON) precipitation for PI (1800-1850 CE) and ERP (400-362 BCE).

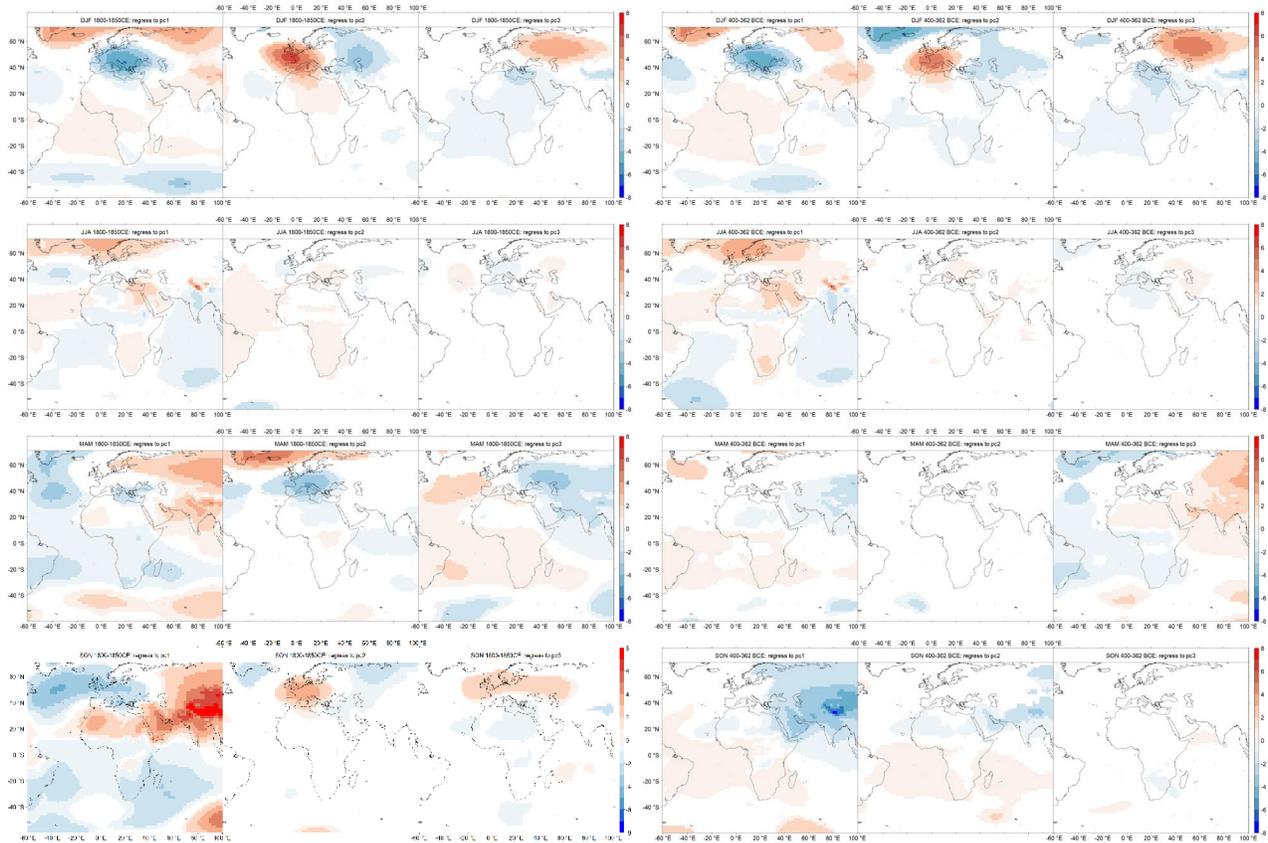


Figure R3. Regression maps between the PCs of the first three non-rotated EOFs for precipitation (four seasons: DJF, JJA, MAM, SON) and SLP from the MPI-ESM simulation for the PI and ERP periods. Only the grid box with statistical significance at the 95% confidence level is plotted.

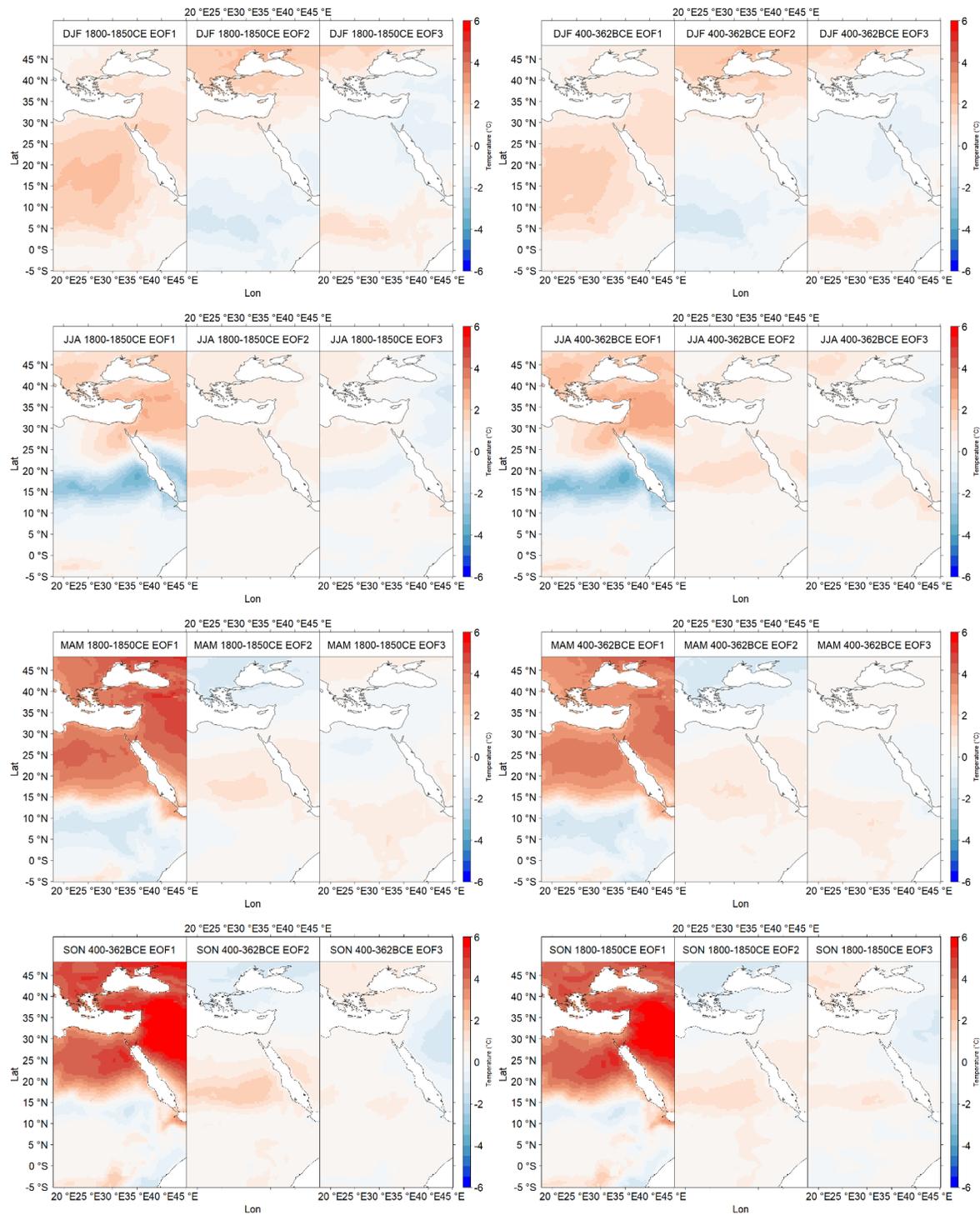


Figure R4. Patterns of the first three non-rotated EOFs of four seasons (DJF, JJA, MAM, SON) temperature for PI (1800-1850 CE) and ERP (400-362 BCE).

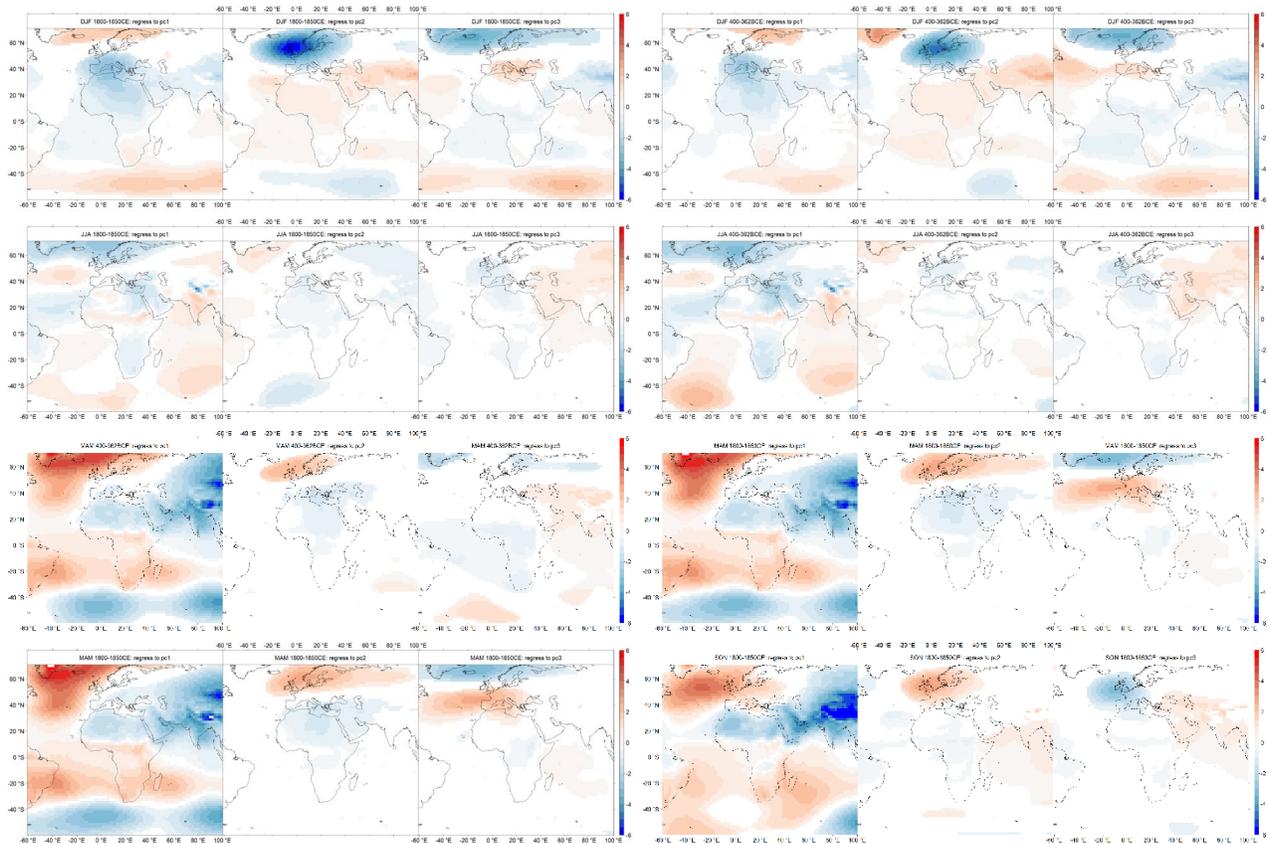


Figure R5. Regression maps between the PCs of the first three non-rotated EOFs for temperature (four seasons: DJF, JJA, MAM, SON) and SLP from the MPI-ESM simulation for the PI and ERP periods. Only the grid box with statistical significance at the 95 % confidence level is plotted.

- The paper lacks a discussion on the added value of dynamical downscaling compared to the forcing GCM. A comparison of the two solutions in terms of spatial and temporal characteristics of homogeneous regions (e.g., Figs. 2-3) and annual cycles (Fig. 7) would help illustrate the usefulness of long transient RCM simulations for paleoclimate.

For providing the added value of the dynamical downscaling compared to the forcing data, we have changed the evaluation part by comparing CRU (observational data), Era-Interim (forcing global data) and CCLM (regional climate modelling output) to further address the added value of dynamical downscaling of global climate modelling date see Figure R6 to R9 for both EM and NR region. In addition to this we have also compared the annual cycle of the three data set for both precipitation and temperature, see Figure R10.

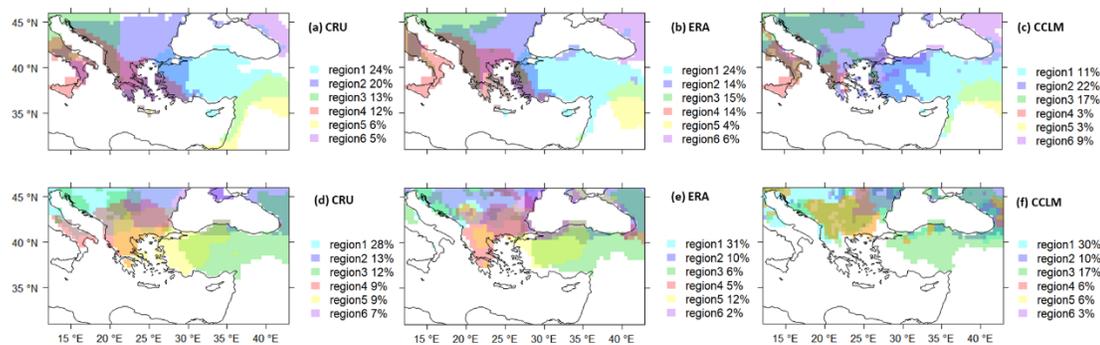


Figure R6. Regionalization of precipitation over the EM in winter (DJF, upper panel) and summer (JJA, lower panel) for CRU (a, d), ERA (b, e) and CCLM (c, f). The total explained variance of the corresponding REOF is shown in the legend.

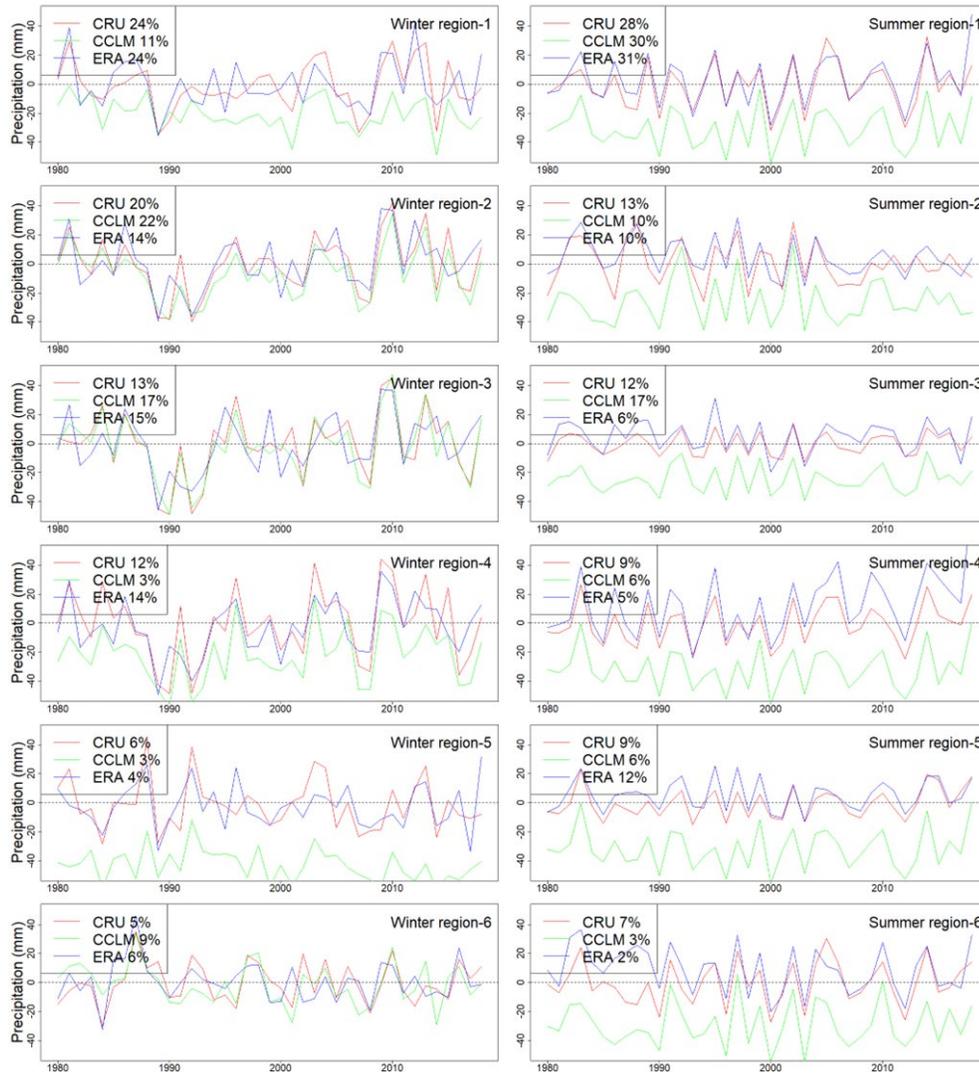


Figure R7. Seasonal mean total precipitation differences of EM regions with respect to the 1980-2018 mean total CRU winter (DJF) and summer (JJA) precipitation.

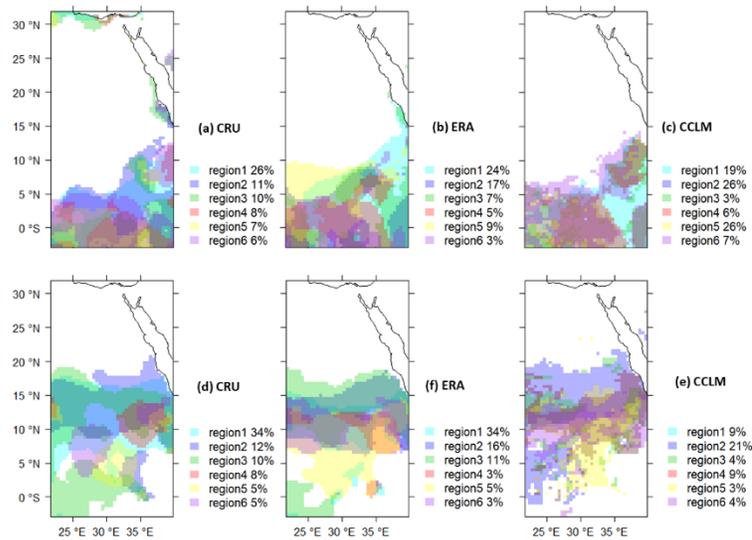


Figure R8. Regionalization of precipitation over the NR in winter (DJF, upper panel) and summer (JJA, lower panel) for CRU (a, d), ERA (b, e) and CCLM (c, f). The total explained variance of the corresponding REOF is shown in the legend.

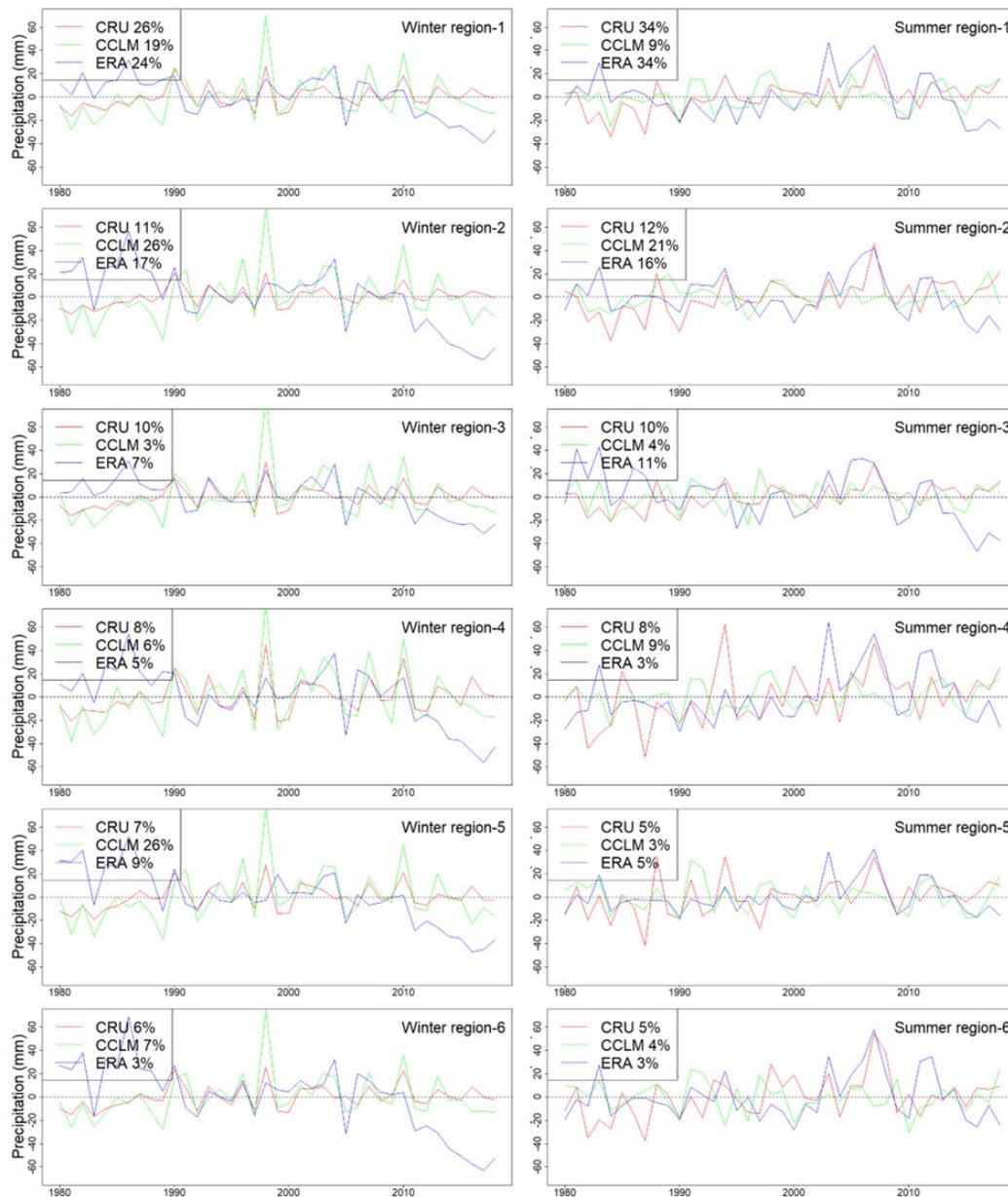


Figure R9. Seasonal mean total precipitation differences of NR regions with respect to the 1980-2018 mean total CRU winter (DJF) and summer (JJA) precipitation.

- Changes in annual cycle and teleconnection patterns presented in Section 3.2 could gain in robustness by evaluating the RCM's capability to simulate these characteristics under current climate conditions. Fig. 7 should be duplicated for the 1980-2018 period (Obs + RCM + GCM) and included at the beginning of the evaluation section.

Figure 7 has been duplicated for the 1980-2018 period as well, see Figure R10, where the CRU data is the observational data, ERA is the global forcing data of regional climate model CCLM.

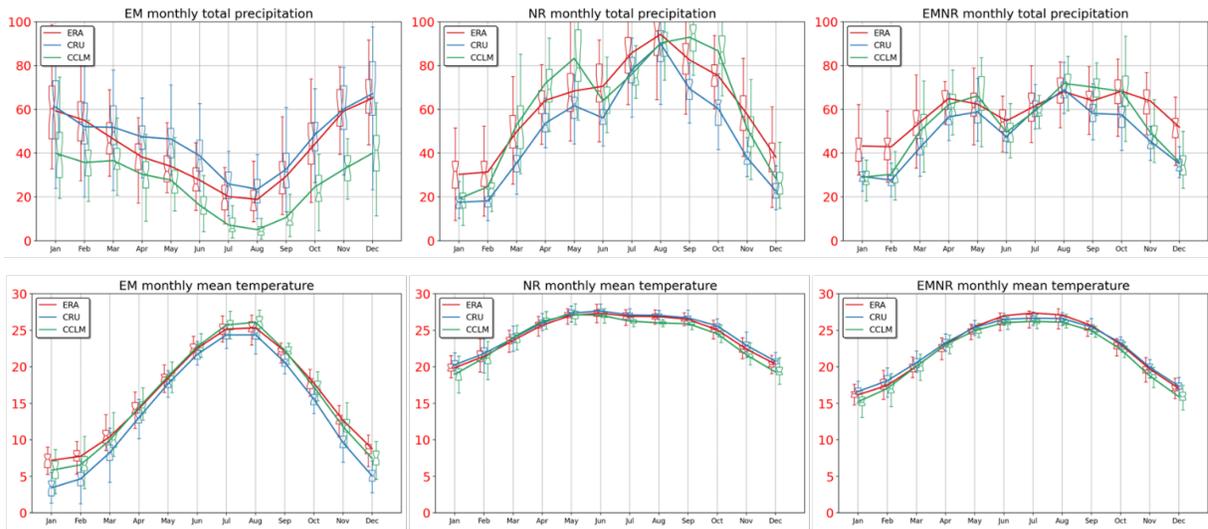


Figure R10. Precipitation and temperature annual cycle of the present period for EM, NR and EMNR of CRU, ERA and CCLM.

- Teleconnections should also be evaluated for the 1980-2018 period in the Supplementary Material and briefly discussed in the main text.

We also present the results of the teleconnections of the 1980-2018 period in Figure R11- R14. This calculation was done when we were preparing the manuscript, it will be provided in the supplementary if necessary. As discussed in the manuscript, the relationship between the sea level pressure and summer precipitation over the EM region does not show a clear pattern as well. Furthermore, only the first EOF of the winter precipitation over the EM region shows a relation with the low pressure center over the EM region and strong positive sea level pressure over the northern Atlantic ocean.

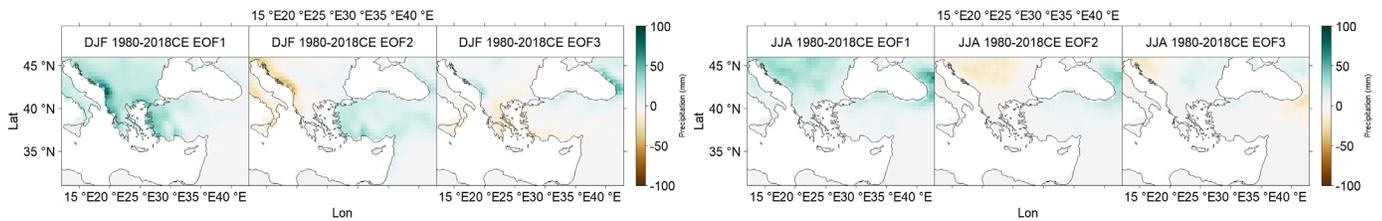


Figure R11. Patterns of the first three non-rotated EOFs of DJF and JJA precipitation for 1980-2018 CE.

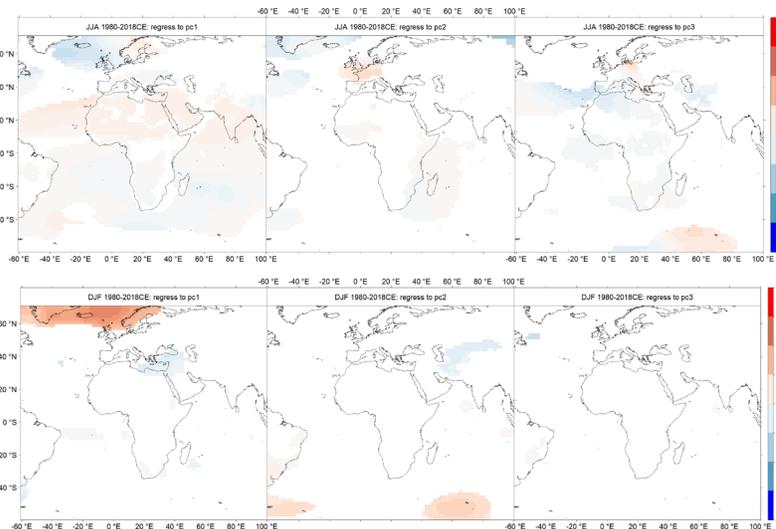


Figure R12. Regression maps between the PCs of the first three non-rotated EOFs for precipitation (seasons: DJF, JJA) and SLP from the ERA-Interim of 1980-2018 period. Only the grid box with statistical significance at the 95 % confidence level is plotted.

- Regarding the EOF approach, there are a couple of questions:

- The rationale for retaining only grid points with positive values for identifying homogeneous region needs clarification, considering that EOF loadings can include strong negative and positive values simultaneously e.g. in the case of a dipole pattern. The possibility of one specific grid point being classified in more than one region should be addressed, along with the sensitivity to the percentile threshold.

The identification of the homogeneous region is done based on the EOF pairs (Table 1 and 2 in the manuscript) which we have calculated with reference to CRU data. Further, the homogeneous regions between the three data set were identified by setting a threshold. Where the value of one specific grid is equal or greater than the threshold value is part of the subregion. We have chosen the threshold value as 80<sup>th</sup> (75<sup>th</sup>) percentile of the precipitation (temperature) loading. As the threshold is set as the percentile of the precipitation (temperature), thus strong negative and positive values when considering the EOF loadings with a dipole pattern are all considered as well.

Here we are presenting the results of using different percentiles (EM region as an example). In Figure R13, we can see that by setting different thresholds, the subregion experiences more overlap when the threshold is too low and covers only small parts of the domain when the threshold is too high. In addition, in Table R1 and R2, we are showing the standard deviation and the percentage of the grid points which have been retained as a subregion. This justifies the selected thresholds in our manuscript.

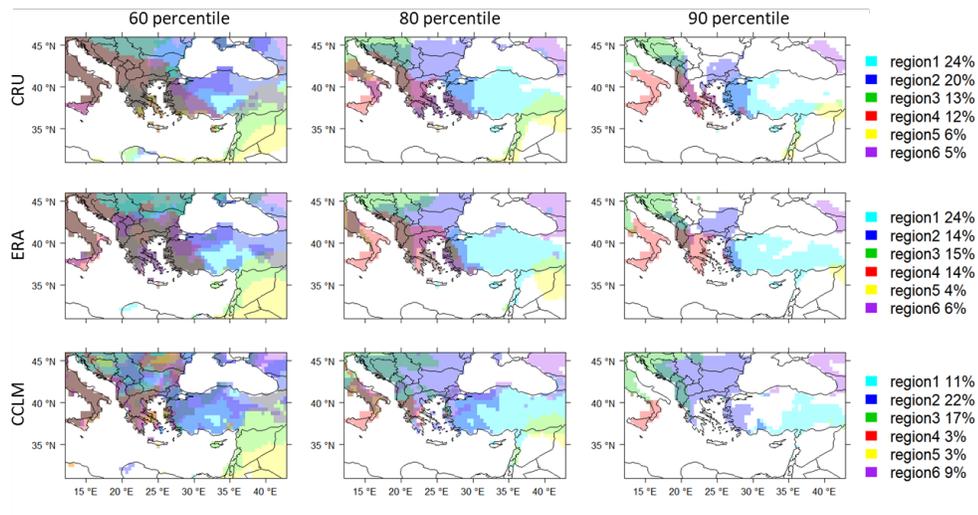


Figure R13. Subregions when using different thresholds for the EM region.

Table R1. Standard deviation of the subregions when using different thresholds.

	Region-1	Region-2	Region-3	Region-4	Region-5	Region-6
CRU-60	11.27	8.30	11.16	9.53	6.88	8.73
CRU-80	10.06	6.68	10.58	7.08	4.98	8.30
CRU-90	9.09	5.18	9.21	4.92	2.95	6.86
ERA-60	10.30	6.80	10.29	9.31	4.58	8.79
ERA-80	8,24	4.74	9.82	7.05	3.69	10.33
ERA-90	7.37	2.97	8.47	5.05	1.94	9.40
CCLM-60	6.14	8.47	12.10	6.07	2.66	10.29
CCLM-80	4.74	7.13	12.69	6.56	1.97	12.88
CCLM-90	3.81	5.60	12.23	5.54	0.32	14.03

Table R2. Covered grid box percentage of the whole region of the subregions when using different thresholds.

	Region-1	Region-2	Region-3	Region-4	Region-5	Region-6	sum
CRU-60	0.32	0.33	0.20	0.18	0.16	0.13	1.32
CRU-80	0.19	0.21	0.08	0.10	0.07	0.03	0.68
CRU-90	0.09	0.09	0.05	0.06	0.02	0.02	0.33
ERA-60	0.34	0.30	0.23	0.19	0.12	0.14	1.32
ERA-80	0.19	0.17	0.10	0.12	0.04	0.04	0.66
ERA-90	0.12	0.07	0.05	0.07	0.009	0.02	0.331
CCLM-60	0.29	0.35	0.22	0.13	0.17	0.18	1.34
CCLM-80	0.17	0.23	0.11	0.05	0.04	0.07	0.67
CCLM-90	0.06	0.15	0.07	0.02	0.001	0.03	0.331

2. Considering the lengthy model evaluation section, it might be advantageous to perform one EOF analysis (or a more straightforward method: see my next comment) for the entire region (all grid points within the EM-NR region) and the entire annual cycle (monthly anomalies) simultaneously. Alternatively, some results could be moved to the Supplementary Material to shorten the section and to let room for a discussion on RCM added values compared to its forcing GCM.

During the preparation of this manuscript, we have tried first to use the entire region to perform the EOF analysis. Due to very different precipitation characteristics of the EM and NR region, the EOF analysis cannot identify the variabilities when using the entire region. Thus, we have decided to separate the region.

3. The EOF approach is not straightforward to identify homogeneous regions in terms of temporal variability. As illustrated in Figures 3 and 5, it becomes apparent that certain regions exhibit highly similar temporal variability. I believe the paper could gain valuable insights by either transitioning to or, at least, discussing clustering analyses (e.g., k-means, hierarchical clustering) since these techniques have proved to efficiently identify regions with shared temporal variability.

To better demonstrate the reason why we are using EOF analysis, we will add a paragraph discussing about those different methods. Compared to machine learning based clustering, regionalization based on EOF remains the physical meaning and the principal components retained from the EOF can further help to demonstrate the temporal variation and link to large scale circulation in climate research.

4. The rationale behind performing rotated EOFs for the present-day period while classical EOFs for the PI and ERP periods should be clarified.

As we mentioned in the manuscript. The classic patterns, EOFs, are orthogonal, hence, the eigenvectors are uncorrelated. For some applications, this is a useful characteristic (i.e., setting up multiple regression models with predictors that are not collinear). In meteorological applications, however, the orthogonality constraint may be disadvantageous, because most processes in the real world are not orthogonal (Storch and Zwiers, 1984). We thus applied the VARIMAX rotation to obtain rotated EOFs (REOFs) that are physically more consistent than the non-rotated patterns. REOFs are thus used for the validation of the model set-up in the present period (1980 - 2018 CE). While non-rotated EOFs are implemented for the comparison of the mean climate conditions in PI and ERP times, ensuring the preservation of meteorological characteristics in temperature and precipitation. The application of non-rotated EOFs guarantees the accurate interpretation of the regression involving the non-rotated principal components of precipitation and temperature with respect to large-scale circulation patterns.

5. Important details about the RCM simulation are missing in Section 2.1, such as the horizontal resolution and more information on changes in forcings, particularly land cover changes over the 2500-year period. Additionally, the high correlation of simulated precipitation's temporal variability with observations (Figs. 3 and 5) requires clarification, given the RCM simulation is forced with outputs from an ocean-atmosphere coupled GCM.

We are currently preparing a paper which is mainly discussing the set up and the changes in the different external forcing in detail (Hartmann, et al. 2024). We can share the manuscript confidentially to further illustrate the above-mentioned information. The

external forcings are based on the recommendation for the PMIP4 past1000 contribution to CMIP6 (Jungclauss et al., 2017). We will add some more information in Section 2.1. Here is some more detailed information:

The orbital forcing is represented by the eccentricity, the obliquity and the longitude of perihelion.

The total solar irradiance represents the solar forcing.

The volcanic by the aerosol optical depth.

The changes of greenhouse gas concentrations are given in equivalent CO<sub>2</sub> concentrations which consists of CO<sub>2</sub>, CH<sub>4</sub> and N<sub>2</sub>O.

The land use changes are introduced as LAI, maximum and minimum plant coverage derived from the MPI-ESM transient simulation.

### **Minor comments**

All the minor comments will be addressed as suggested.