Clim. Past Discuss.,
https://doi.org/10.5194/cp-2020-8-RC2, 2020

Climate of the Past Discussions

Interactive comment on "Dynamical and hydrological changes in climate simulations of the last millennium" *by* Pedro José Roldán-Gómez et al.

Anonymous Referee #2

Received and published: 17 April 2020

The paper by Roldan-Gomez and co-authors aims at evaluating the relative influence of external forcings on large-scale changes in PMIP2/CMIP5 last millennium climate model simulations including the historical period. To address this issue they relied on various statistical method and mainly EOFs analyzes and evolutions of their related PCs. Even though the paper is generally well written with potentially interesting results I have several concerns regarding the method and interpretations. The authors need to significantly improve the paper, as there are many important points to clarify or to be corrected before publication. I have listed bellow my main comments and criticism to be addressed:

Models and methods:

1. First of all they show time series covering the last millennium and the historical period as continuous model experiments. As far as I know this might not be the case for most of the model experiments used in this paper as the historical experiments in CMIP5 are branched off the pre-industrial control runs and are not a continuation of the LM simulations. The authors need to explain how they build the time series anomalies to make them look like seamless long climate model integrations. This is very important since this study discuss long-term trends and secular changes which depend on long term integration of external forcing histories. Historical runs branched of piControl runs might therefore include different initial mean background climate condition and trends. This should be clearly evaluated and the method used to take that into account when comparing to LM runs. How were the anomalies computed for each experiments used (piControl, LM, Historical) ?

2. The authors states that the model simulations were concatenated and time series low-pass filtered with a centered 31 years moving average. Which frequency cut-off was used to filter-out? The 31 years moving window was used to compute the anomalies? This should be clarified.

3. The method used to estimate the Total External Forcing (TEF) obtained by composing the contributions of several forcing factors should be explained in the method section.

4. This section does not give enough specific and explanations as how the EOFs analyzes is developed across PMIP3 models used. How the PC selection linked to the forcing is done? Which statistical method did you consider to evaluate the spurious results related to the different forcing data-sets and implementation strategies?

5. In the PMIP3 ensemble simulation, some model multiple realizations are included in the analyzes. From my understanding, each model experiments are given the same weight when performing the EOF analyzes or ensemble averaging. This will tend to

give mode weight to a few models. The authors state that the results are not affected by this sampling bias but they don't show and provide statistical measures in the subsequent analyzes to prove it. I suggest that a weighting is applied considering the number of experiments for each model to correct the sampling bias and make sure the results are unchanged.

6. The author state on page 8: "Some long-term changes in the external forcing, like the one during the transition from MCA to LIA, are significant enough to be obtained not only by performing PC analyzes but also by directly looking at the evolution of the variables during these two periods." I don't understand this sentence? Does that mean the authors assume that the leading PCs across LM ensemble for the considered variable and the actual evolution of the considered variable during the transition from MCA to LIA are the same? The authors should clarify this statement and prove it. Which long term external forcing changes during MCA/LIA are the authors referring to? This statement needs to be accompanied with quantified analyzes with statistical significance estimates.

Over the method section needs significant rewriting with a more systematic explanation of which methods is used to evaluated the statistical significance and relevance of the analyzes displayed in the results section. The authors should also clearly make a choice regarding the frequency window they want to investigate. Many mixed statements are presented in the results sections, regarding mean climate anomalies during the MCA relatively to LIA, secular trends and climate modes of variability occurring at various timescales. As it stands we cannot really makes sense and relate some assertion regarding climate modes of variability relying on displayed analyzes.

Results sections:

7. The authors make the following statement on page 8 in the 3.1 results section: "The peaks in volcanic forcing after the main eruptions are related to periods with lower global temperatures, while the multidecadal variability and long-term trends associated

with solar and anthropogenic forcings correspond with the long-term changes in temperatures that define periods of the MCA, LIA, and industrial era." Which analyzes attribute the multidecadal variability and long-term trends with solar and anthropogenic forcings? This is merely assertion not proven by presented results especially with latest forcing datasets used in PMIP3 which have shown a very weak or no fingerprint of solar irradiance forcing during the LM. The authors need to provide analyzes for the multidecadal variability and trends proving otherwise.

8. Page 9: "For the 20th century, all the analyzed simulations consistently show a warming, but trends strongly differ among simulations due to the different climate sensitivities of each model and the considered forcings". To which forcing this stronger sensitivity refers too? References should be cited to consolidate this assertion.

9. Page 9: "In a related and most relevant note, changes in the ensemble associated with external forcing are in general more relevant than those of internal variability." To which timescale this statement refers too? Is it for decadal or secular trends? This should be quantified and specifically quantified related to the frequency domain the authors want to discuss.

10. Page 9: "Note that most of the analyzed simulations show correlations larger than 0.5 and for simulations of the same model the correlations reach values around 0.9, both when analysing the whole period and when considering only the pre-industrial era. This indicates that even if the EOF has been obtained with a combined analysis, it is also representative of the individual simulations. Additionally, the use of large sets of simulations for some of the models, and 20 in particular the use of the 13 CESM-LME simulations, does not significantly bias the results, because the correlation ranges for models with individual simulations are as large as for the others." Since piControl runs are a measure of internal variability for each model, I don't understand why the authors get high correlation for both LM and piControl runs ? The method used should be clarified since the above results suggest either a flawed method or that LM changes and high correlations among model members including piControl are only due

to internal variability (the leading modes of internal variability present by construction in the piControl run?).

11. The authors also discuss changes in the leading EOF for SLP (and other hydroclimate variables) which probably reflects the first order thermodynamical response to global temperature changes due to external forcings. Yet the authors attribute it to changes in phases of the NAO, NAM and SAM or even ENSO/IPO in response to external forcings. They don't provide any analyzes that prove it. The authors states for example that there is "a tendency toward more positive phases of the NAO, NAM and SAM is observed during the MCA and industrial periods." However no relevant analyzes are shown to sustain these statements showing for example a quantified and causal link between the leading EOF for SLP and the actual changes in (internal) variability modes. The authors rather present long-term mean anomalies between MCA and LIA or time-series of leading PCs for global scale variables. Yet by definition internal modes of variability are characterized by leading pattern and frequencies prevalence that are not analyzes in the present paper. This comment applies almost to all the points discussed in the results section where many descriptive and speculative assertion.

12. For example, the presented and discussed results for SLP changes are confusing and somewhat contradictory. For instance, the authors state "simulations of GISS show an increase of pressure after volcanic events, while simulations of CESM-LME consistently show a decrease. This difference in the global average of pressure is not related to an opposite response in different models, but to the distribution of areas with positive and negative loadings in the mode of variability associated with the forcing. As shown in Fig. 5, simulations of CESM show a larger amount of areas with negative anomalies during periods with volcanic events, while simulations of GISS tend to show more areas with positive anomalies." An other example for the wind changes: "In spite of the differences in the global balance of regional positive and negative anomalies among models, all of them produce a global weakening in zonal circulation during

volcanic eruptions. " or "In general, this global analysis shows that regional modes of variability might be indirectly influenced by external forcing".

These are descriptive assertions, which need to be quantified and evaluated in terms of significance. Based on these few examples and the overall presentation of results sections, one can conclude that the simulation changes (leading EOF and volcanoes composites) are not really significant and alternatively interpreted as mean changes, decadal and secular trends or internal variability modes acting at inter annual (such as NAO) to decadal timescales (such as SAM) depending on the authors choice. Changes in variability modes are mixed with long-term trends and mean changes. However no results are presented and assessing these various questions separately depending on the timescale.

To sum-up I suggest major revisions. The authors need to exclude statements that are not sustained by actual relevant analyzes and focus only of long-term trends and mean MCA/LIA changes. In the actual form the paper will mislead the readers regarding the responses of the variability modes and the roles of external forcings based on speculative comments. The results presentations need to be improved focusing on specific timescale based on statistically significant signals analyzed with the appropriate method.

Interactive comment on Clim. Past Discuss., https://doi.org/10.5194/cp-2020-8, 2020.