

Response to Reviewers

Reviewer 1

Main points

A first point concerns the considered ensemble of simulations: I wonder why MPI-ESMP, for instance, is not included in the list (a millennial control run is also available in the CNIP reportistory for this model). Also, there are at least two full-forcing past1000 simulations with GISS-E2-R available, which one is used here? And why not use both?

- The MPI-ESM-P model was processed but excluded from the multi-model ensemble by error, which has now been corrected. In addition we now explain which GISS-E2-R realization was used and why we selected only one (“For GISS-E-2-R, we include only one contributing realisation (r1i1p121) to constitute a multi-model ensemble of one ensemble member from each model.”)

A focus on statistical significance is of course important to substantiate all the results in section 4.1 and 4.2, and especially those concerning differences in ENSO statistics/metrics through the last millennium as these are part of the main conclusions of the study. The methods are well described in section 2.3, but I failed to see the significances for instance for figures 7 and 8.

- Figure 7 has now been removed on recommendation of Reviewer 2. The significance of correlations in Figure 8 is now included.

Also, I wonder why the authors decided to use the 20CR when only the 1976-2005 period of historical simulations is considered (Table 1). Why not use the whole period covered by the 20CR, or use instead more or different available reanalysis products? As the authors also report in the introduction, the instrumental record provides limited guidance for understanding the range of ENSO behaviors. Still, observations indicate that ENSO properties have changed over the last several decades, in particular with increased frequency of so-called Central Pacific events in most recent decades - the ones considered in this assessment (see, for instance: Pascolini-Campbell et al., 2014). It has been emphasized that different “types” of El Niño exist during the observational period that have substantially different characteristics (including different teleconnections, as shown for instance by Graf and Zanchettin, 2012). This observed behavior should be considered when discussing the simulations-reanalyses comparison over such a short and peculiar period of time.

- We selected a short period of time for comparison of the historical experiment and the 20CR reanalysis data so that we could focus on ENSO characteristics, rather than the anthropogenic signal evident over the extended period. We have now made clear there is greater variability in ENSO over the extended period, and that the recent period is unusual, and hence that we are only evaluating part of the ENSO system in comparing datasets over a short period (“Models were compared to twentieth century reanalysis data (20CR) (Compo and Whitaker, 2011) as a widely proxy on observed climate (King et al., 2014; Klingaman and Woolnough, 2013). We compare datasets for the period of 1976-2005, rather than an extended period, due to greenhouse forced non-stationarities over the post-industrial era. It should be noted that ENSO properties have changed over the last several decades, in particular with increased frequency of Central Pacific centred events in recent decades, which have substantially different characteristics (Pascolini-Campbell et al., 2014). Hence model skill in recent decades does not ensure that all variations of ENSO are equally well captured.”)

A similar question concerns the limited temporal domain used for the Mid-Holocene simulations: is 100-year a long enough period to guarantee robust estimates about ENSO behavior, given the variability that is reported about the last millennium? The authors should consider expanding the “model evaluation” section and related discussion: in fact, they mention six metrics used to evaluate ENSO, but in the following text there is very limited discussion on this.

- The model evaluation section has now been clarified to make clear that we are referring to the Bellenger et al study on model evaluation of ENSO. The 100 year period for the mid Holocene is used to provide an estimate of the mean climate state and represents the minimum period provided by all models. We do not suggest that 100-year period provides an exhaustive representation of possible ENSO behavior, but rather that it useful for examining possible boundary condition influences.

A deeper analysis could substantiate interpretation of some results which appears at occasions to be not conclusive. For instance, concerning the difference between historical and last millennium simulations in Figure 1 (section 4.2), the authors provide an only vague interpretation (1592/21-25), while I felt it was exactly the aim of this study to provide an answer to this regard. I also did not find conclusive the analysis of internal versus externally forced ENSO variability in section 4.1. The authors themselves agree that this is the case (1598/26-29), so I wonder what the aim of this section is: overall, I suggest the authors to either deepen the analysis or largely restructure/reduce this section. Some specific concerns/suggestions I have on this are: when external forcing is considered, such as variable solar irradiance, why not substantiating the results with a wavelet coherence analysis (1590/15-24)? Also, the assessment of the role of volcanic forcing is too vague: no result is shown (e.g., from a superposed epoch analysis as typically done in these cases), only three major eruptions are reported in Figure 3 (but not the 1815 Tambora, why?), and only one eruption is discussed in the text. Later on, volcanic forcing (1591/23) as well as combined volcanic and solar forcing (1591/27) are reported again as a possible important factor for ENSO evolution. The summarizing paragraph (1591/21-1592/2) appears again to be too vague (“may be... may reveal...”).

- We have now shortened this section as recommended and we also provide a more specific discussion of volcanic forcings is now provided in section 6, with greater reference to previous model-based studies.

Minor points

1581/7: typo (“is a is a”)

- This has been corrected.

1581/8: maybe Zou et al. (2014) is a worthy addition here

- The Zou et al (2014) study is primarily focused on pre-industrial CMIP5 simulations, not future projections, so is not added here.

1581/17: I guess it is “does NOT capture”

- This has been corrected.

1585/14: remove “in”

- This has been removed.

1585/18: isn't it Fig. 1 (and not Fig. S1)?

- Fig. 1 shows composited anomalies and Fig S1 the EOF analysis.

1587/13: please check that acronym SD is defined

- This has been corrected.

1587/21: is MIROC5 the same as MIROC-ESM?

- This has been corrected to MIROC-ESM

1588/5: I am not sure what “physically plausible” means in this instance, maybe expand a bit?

- This description was vague and has now been deleted during the shortening of this section.

1590/15-21: Can you be more specific here about the role of solar activity? Are the prevailing La Nina like conditions induced by increased solar activity a result of this study or from previous ones? Actually Figure 3 does not seem to show this as the 1258 seems rather associated to cold anomalies.

- This discussion on solar activity has now been deleted as it was not contributing to the discussion of teleconnection stability.

1591/27: combination of

- This has been corrected

1592/17: resemble

- This has been corrected

1593/4-5: I wonder whether the linear relationship is really different for the two experiments, or, rather, the regression is for both not significant (and then differences do not really matter).

1593/6-7: I was not able to see where significance is reported? I think it is important to report it since by eye I wouldn't say that for some regions/variables the changes are so dramatic...

1593/15: same as above: where is significance reported?

- The statistical significance of correlations is now shown in Figure 7 and 8, and in the discussion.

1594/5: sites in the tropical ...

- This has been corrected

1596/23: we find that ENSO...

- This has been added.

1597/3-5: I think the use of parentheses here is confusing

- This has been expanded and written out explicitly ('Models suggest it may be inherently difficult to deconvolve variability in the NINO3.4 region and local-scale, teleconnected climatic change in the West and East Pacific regions. The West Pacific Warm Pool is likely sensitive to subtle shifts in the western extent of the warm tongue characterising positive (El Niño) episodes, and conversely to the cool anomalies charactering La Niño episodes.'

1597/25: “the stability ...is ...variable” sounds strange, so maybe rephrase?

- We have removed “stability” .

1598/6: why necessarily?

- Necessarily has been removed

1598/8: volcanic)

- It is unclear what should be changed here.

Fig. 1 caption: check space in “La Niña”

- This plot has been replaced by a multi-panel plot showing each model.

Fig. 2: there is a strong peak at 6-year period in the historical IPSL-CM5A-LR simulation, any thoughts on this?

- This peak was likely an artifact of processing the short historical period (1976-2005). When models are processed over 1906-2005, most reveal a strong peak around this period, which is now discussed in section 4.1.

Fig. 3: the anomalies for bcc-csm1-1 are noticeably mostly negative, so I wonder how anomalies are exactly calculated (not from full-period average?)

- This was an artifact of processing error and has now been corrected. We thank the Reviewer for identifying this, which has been corrected and now included in Supplementary Figure 4.

Fig. 4 caption: check panel for 20CR precip

- It is unclear what should be corrected here.

Fig. 5: maybe it could be useful to add a Box-Whisker plot for the past1000 simulations, to see how they compare with the piControl.

- We do not directly compare the piControl simulation to the Last Millennium because they have a different number of contributing models.

Fig. 8: To me it seems that the only changes in the West Pacific for temperature are associated to volcanic eruptions (1258, Kuwae). Does this support the hypothesis of a volcanic influence? The question is also how much short-term effects could affect the long-term (100 year in this case) statistics. Was any smoothing applied to the series? How would the statistics change if the data around the years of major eruptions are removed from the analysis?

- We now discuss details of the timing of volcanic eruptions and changes in the remote-local relationships in section 6, including the persistence of such influences. We did not apply smoothing to this aspect of the analysis.

Supp. Fig. 2: should one of the “showing” be removed?

- This figure has now been deleted.

Supp. Fig. 6: what does the blue shading indicate in panel a?

- This figure has now been deleted.