Answers to: RC1

https://doi.org/10.5194/egusphere-2023-757-RC1

Dear anonymous reviewer 1,

First of all, many thanks for considering our manuscript for publication and taking the time to review. We value that you think our work is worthy of publication after revision and responding to your feedback. We appreciate your positive words, and we acknowledge your critiques and constructive feedback. In this document we will aim to answer your questions and respond to your feedback, in a way that we hope to be satisfying. Ultimately, we think the manuscript will be improved and will be accepted for publication after intended revisions.

General comments.

This paper explores mean state and jet variability changes in response to increased CO_2 and mid-Pliocene boundary conditions through previously published general circulation model experiments using the CESM. The experiments are well-designed, and some interesting, although perhaps not too surprising, results are found, that changes in climate variability can be very different in response to increased CO_2 relative to other boundary condition forcings that also create warmer climates. Overall, the structure of the manuscript is logical, and figures are clear. However, I find much of the results section to be very descriptive of the figures, with little interpretation, leaving the reader to try to find their own interpretation - more interpretation of the results (and less description) would greatly improve this manuscript.

We agree the Results section is quite descriptive and could use more interpretation. This was a choice we made, but we understand it doesn't read well and leaves interpretation to the reader, which is not the intent. We propose the following:

- We will reduce the level of description in the Results, in those places where the description is not really necessary for answering the research questions, or where description is redundant because it is obvious from the presented Figures. Example 1: L263 – 265 ("Temperatures increase … higher latitudes.") these two sentences will be combined and thus shortened, as it is merely description. Example 2: L289-290 ("The dynamical … Eurasian continent.") this sentence will be removed as it is again merely a description of the figure with no interpretation.
- We will increase interpretation in the Results in two ways.
    - First is to move some of the interpretation that is currently presented in the Discussion, specifically Section 4.2 Physical and dynamical interpretation, to the parts in the Results where it would fit. For example L438-445, a paragraph offering interpretation of the CO_2 doubling results and making a link between the mean changes and variability changes, can almost entirely

    be moved to the Results section 3.2.1 (Sea-level pressure variability / CO_2 doubling).

  - Secondly we intend to include subquestions to the main research question in the Introduction (including: "How do changes in the mean winter state relate to changes in winter atmospheric variability?") as well as some hypothesis based on literature that is now newly presented in the Discussion (mainly section 4.1). In the Results section, when presenting our results we can then refer back to the questions posed in the introduction, as well as to our initial hypothesis. For example, linking mean state changes to changing in variability is what we are doing in the previously mentioned paragraph L438-445. Moving that to the Results, and referring to the question asked, will hopefully increase the level of readability of our Results section. (more on rewriting/reordering of the Introduction below)

The discussion section leans heavily on prior literature and could do a much better job of putting the new results in the context of literature and highlighting the novel results in this study and their relevance.

We agree that our Discussion is currently introducing some literature that was not treated earlier in the paper (for example, in the Introduction). Next to that, how the Discussion is currently written can indeed be confusing as to when we refer to our own results, and when to previously published results. We propose the following:

- We intend to reorder the Introduction (more on that below), where we will introduce a paragraph on the effects of the different boundary conditions of the Pliocene on the climate. This would include parts of section 4.1 Sensitivity to mid-Pliocene boundary conditions. For example paragraph L417-424 is treating SAT and SIE response to different mid-Pliocene boundary conditions in other climate models, which would actually be relevant to mention in the Introduction.

Including uncertainty estimates on some of the quantities reported would also improve the paper.

Thank you, we will take that into consideration. Your suggestion to calculate uncertainties of the changes in variance explained is one thing we will do. We furthermore plan to include statistical significance information on the correlation coefficients calculated, for example in Figure 7 (where some correlations might not be statistically significant), and Figure 8.

The research is motivated by the question of whether "the mid-Pliocene climate be used to assess the response of present-day Northern Hemisphere winter atmospheric variability, such as the NAO, NAM and PNA, to increased CO_2"; however, I remain a little unclear as to why this is a valuable question to ask? If we trust the model in reproducing climate variability of the mid-Pliocene, why don't we just trust the models for the future? One argument for using paleo data is that we have proxy observations and so don't need to rely only on models, but here you are just using a model, so any model biases remain. Exploring and understanding climate variability in different climates is interesting and useful for understanding the underlying climate system and behaviour of internal variability, and your result that variability is very different in two different warm climates is interesting. These insights could be helpful, in a more indirect way, in helping understand projected future changes. I think the research focus is interesting and worthy of publication, but the motivation does not convince me as it is currently presented in the introduction, and I found the structure and order of the introduction rather confusing.

For example, the paragraph starting on line 42: “An issue with investigating the response of climate variability to increasing CO2 in the near-future is that the present-day climate system is not in equilibrium with the mean forcing….” – are you saying that the present day isn’t an analog for the future because we’re not in equilibrium? Or that, for model simulations in which the climate is changing, it is difficult to assess what is trend and what is a change in variability? I would argue that, at least in the near-future, the climate won’t be in equilibrium, so the modelled simulations of future climate seem like a better analog than a past climate that is in equilibrium? Even if you argue that you can only study this in an equilibrium climate, why not use simulations of future changes that have reached equilibrium, e.g. 4xCO2?

We acknowledge your critical feedback on the formulation of the motivation as well as the structure of the introduction. We believe our motivation is valid, but can understand that how it is currently phrased and presented (which includes the research question and structure of the introduction) can be confusing, or not convincing. Instead of trying to answer all the question you asked above, we propose a different structuring of the Introduction and a more clear formulation of the motivation and research questions. We believe that will cover most and hopefully all criticisms you pose.

Proposal of paragraphs of the Introduction:

1. Introduce the midPliocene as most recent period with atmospheric CO2 similar to present-day. Explain why midPliocene is considered a possible ‘best analog’ for near-future climate (when considering mean temperature and precipitation and compared to RCP4.5 projections, see Burke et al 2018) – would include parts of current L42-51
2. Introduce midPliocene modelling efforts – would include current L52-60
3. Explain midPliocene can also be relevant for long term climate projections, in terms of CO2 (400ppm similar to optimistic SSP1-2.6 scenario in 2300), or when considering reduced Greenland ice sheet and West Antarctic ice sheet. – would include some parts and literature that is now presented in Discussion 4.4 The mid-Pliocene as future analog?
4. Explain next to CO2 and reduced GIS, there were also closed gateways, which previous studies have shown to have effect on mean climate – move some literature and sentences from Discussion section 4.1 Sensitivity to mid-Pliocene boundary conditions here
5. Apart from mean climate, we can also study climate variability in the past, in order to further understand climate dynamics & variability response in warm climates, as well as how mean climate changes relate to changes in variability - Cite work on midPliocene ENSO, a.o.

6. Also very relevant to study winter variability, such as NAO. Large impacts, but future projections not in agreement. What happens with winter variability in warmer climates? – This would largely be L19-41
7. Brings us to the rephrased research question: "Can the mid-Pliocene be used to investigate the response of NH winter variability to a warmer climate?" with subquestions: 1. Is there a difference in the response to elevated CO2, and to mid-Pliocene boundary conditions other than CO2, including closed Arctic gateways and reduced ice sheets? And 2. How do changes in mean winter climate relate to changes in atmospheric variability?
8. Introduce the model used and the specific simulations used to answer the questions – corresponding to L87-L95
9. Briefly hypothesize what we expect the answers to be based on previous studies – this would consist of parts from L61-L76, section 4.1, and section 4.3
10. Paper outline – current L96-100

In the current Introduction, the focus might be too much on using the midPliocene to 'fix' the fact that current climate projections on winter variability are uncertain or not consistent. From that train of thought, we can understand your comment "why don't we just trust the models for the future". Hopefully, with the newly proposed structure of the Introduction, it is clearer that we study the midPliocene because it can help us to understand warm, high CO2 climates, and we investigate winter variability in that climate to get a better grasp of winter variability in warm climates.

Lastly, the title of the paper suggests this is a response that is consistent across many past warm climate conditions, rather than just the mid-Pliocene – given the suggested dependence on orography, and possible dependence on SSTs, this may not be true; I suggest to be more specific in the title so it's not potentially misleading.

Thank you, we agree the title can be potentially misleading. The other reviewer also has a comment on the title, so to answer to both, we propose the following title:

"**Mid-Pliocene not analoguous to high CO_2 climate when considering Northern Hemisphere winter variability**"

Specific comments

'the geological past climate was in equilibrium with forcing' – if this was always true, the climate couldn't have changed in the past? You could argue that it was, most of the time, more in equilibrium than we are today.

L44. Indeed, depending on the timescale considered, the climate was not always in equilibrium, as it has adapted to changing ice sheets, atmospheric CO2 levels, etc. To clarify, we will add a timescale, for example ".. on timescales relevant for climate variability..".

You mention variability on decadal timescales, but not the PDO (Pacific Decadal Oscillation), which surprised me – is there a reason to not discuss the PDO in modes of variability of the

North Pacific? Particularly given, on line 175, you say you are mainly interested in interannual and decadal variability.

The reason we did not mention the PDO in the Introduction is that consider atmospheric variability, and the PDO is classically defined as an oceanic mode of variability. However, it is true that the PDO has a clear connection to the atmosphere, with the PNA as the most obvious teleconnection (e.g. Chen et al 2018). In that light, it is relevant to mention the PDO earlier in the manuscript. We will include a sentence in the Introduction that mentions the PDO as well as ENSO and their deterministic links to NH atmospheric variability.

Line 173. What re-analysis data do you use? Do you have enough years to use a 50-year window?

We compare with NOAA's CR20 reanalysis, as is stated in section 2.3 Validation of E280 with reanalysis data. We propose to remove the mention of reanalysis in L173 and move the explanation that we apply Lowess filtering to remove climate change trends to section 2.3

Line 175 'A window size of 50 years was chosen since we are mainly interested in interannual and decadal winter variability' what degree did you use for your Lowess smoothing, and aren't you removing all of the interannual variability and most of the decadal variability by using a 50 year window?

- *Check Lowess degree*
- We understand the confusing regarding the Lowess filtering; we **filter** the data by **removing** a Lowess smoothing with a 50y window. So we are left with anomaly data where all the variability with periods larger than 50years are removed, and the periods below 50y are still present. We propose to rephrase: "Before analysis, a Lowess smoothing using a 50 year moving window is removed from anomalies at each spatial grid point."

Line 185: how do you determine level of zonality? Just subjectively by looking at the EOFs or some objective method?

We determine level of zonality by subjectively looking. We have employed a method earlier that used the same definitions used for changing sign of the EOF (L188-190), but this still required a lot of (subjective) tuning so that it worked well. In the end, we are just dealing with two times four EOFs per basin (CR20, E280, Eoi280, E560), so we chose to determine the level of zonality qualitatively.

Furthermore, we propose to stick to the known nomenclature (NAO, PNA, etc) instead of NAtl-z, NPac-a, in order to answer comments from Reviewer nr2. We think that the current naming convention, including the definition of 'zonality', might confuse the reader. Sticking to known nomenclature will help, with the side note that the modes in the midPliocene might not be exactly what we expect from the present-day. This is for example also what has been done in previous studies regarding NAO in past climate (e.g. the LGM, see https://journals.ametsoc.org/view/journals/clim/23/11/2010jcli3372.1.xml)

Figure 1 caption. I think d) should be NPac-z not NPac-a?

Correct. We will change this in NPO.

Figure 1. It's a little confusing whether the contours are the absolute CR20 values, or the differences between E280 and CR20. I think perhaps it is differences for a. and b. and then absolute values for the others, but the caption does not make this clear.

It does specify it currently in the caption: "For (a) and (b), the contours represent the difference between the E_{280} and CR20, while (c)-(g) show the CR20 EOFs in contours." To clarify, we can add ".. show the **absolute values of the** CR20 EOFs …"

Line 198. MAX and MIN would seem more related than PLUS and MIN

It was chosen to stick with nomenclature that is used in literature, at least commonly with the NAO phases (NAO+ and NAO-). Since we plan to adopt NPO instead of NPac-z, we will also change this to call the phases NPO+ and NPO- instead of NPac-z PLUS and NPac-z MIN. We hope that will be clearer to the reader.

Section 2.3. The description of the re-analysis dataset would be better earlier, before you mention that you've used re-analysis data in line 173.

We decide to remove the mention of the reanalysis in L173 (see earlier answer).

The way you have described the re-analysis data is confusing as 'we use assimilated sea-level pressure data from the NOAA 20C reanalysis….' – this almost sounds like you've done some post-processing to the re-analysis data, or that you're using the data that was assimilated into the re-analysis. 'The data runs from 1836 to 2015 and is assimilated using surface pressure observations on a 1.0◦ latitude x 1.0◦ longitude grid' This sounds like the surface pressure observations are on the 1x1 degree grid, and assimilated the re-analysis, rather than the re-analysis data assimilating the surface pressure observations.

We understand the confusion regarding the current phrasing. We propose to rephrase it as such: "We use **monthly mean sea-level pressure data** from the NOAA … as CR20). The **reanalysis covers the period from** from 1836 to 2015 and **assimilates surface pressure observations in combination with a forecasting model to estimate a set of atmospheric variables.** An evaluation …. Slivinski et al. **The CR20 data is presented on a 1deg latitude x 1deg longitude grid, and** we interpolate the …. Between the E280 and CR20."

I think it is worth mentioning that a benefit of the CR20 for these longer time periods is that there is more consistency in the data that is assimilated than for, for example, ERA5, which includes satellite data in more recent periods. That said, the amount of data being assimilated certainly does change with tie in CR20.

Agreed, and the length of the dataset, plus the fact that it covers more of what can be considered 'pre-industrial', was the key reason for us to choose this reanalysis over ERA5 and ERA20C. We can add a sentence in paragraph L201-207 to clarify: "The length of the CR20 dataset and the fact that it covers a period closer to what can be considered pre-industrial motivated us to use this reanalysis data over more recent reanalyses such as ERA5."

Section 2.3. Why do you use all of the data in the re-analysis, through to 2015, when your simulation is just pre-industrial, as you mention? Do you de-trend the re-analysis data to try to take out any climate change signal in MSLP? Would choosing a shorter time period that is closer to your pre-industrial modelled dataset not be better?

See also the previous answer. Yes, we detrend the data, as is mentioned in 173-174. We will remove that part and move that mention of the detrending here.

Line 209. Do you mean spatial mean MSLP? Is the model bias in global mean sea level pressure particularly meaningful? Seems reasonable this is relatively small since sea level pressure is essentially a measure of the amount of atmosphere above a point, so the global mean MSLP is broadly a measure of how much atmosphere there is in your model?

Yes, we mean spatial mean MSLP. Indeed, it might not be particularly meaningful, although it is not global mean MSLP but NH winter mean MSLP. There could be a discrepancy in atmospheric mass distribution over the hemispheres in a particular season. However, your critique is valid, and we propose to remove this sentence altogether, and instead compute RMSE (computed per grid point, then averaged).

Figure 2. It might be interesting to show Eoi400-Eoi280 as an addition panel in fig 2. I realize this isn't a doubling, and so isn't equivalent to panel b, but it would be interesting to see if the pattern response to increased CO2 is similar in the Pliocene (or you could scale the responses, e.g. to 1 degree of global warming)? This would also help with seeing your result described on lines 253-254, as it seems this result (that the MSLP difference are predominantly caused by the different surface boundary conditions, not the different CO2 levels) is quite a key one, given that one of your main conclusions is about the change in variability in response to these two forcings. However, from the panels presented in Fig. 2 and Fig. S2, it is not easy to understand what comparison to make to come to this conclusion. I think adding Eoi400-Eoi280 to Fig 2 would help.

Thank you for this suggestion. We have actually performed a 'CO2 doubling' experiment in the mid-Pliocene, the Eoi560 (see Baatsen et al 2022). The results of this simulation can be used to compare to the CO2 doubling with pre-industrial BCs (i.e. Eoi560-Eoi280 versus E560-E280). In addition, it can be used to compare to the mid-Pliocene BC effects at high CO2 (i.e. Eoi560-E560 versus Eoi280-E280). We chose not to include the results of this simulation, as we figured it would not be necessary to answer the research questions. However, considering your point here, as well as your later comment on assessing nonlinearity, we propose to include some Eoi560 results in the revised manuscript.

We are not convinced to include the Eoi400 – E280 results in Figure 2. We think it might distract from the results discussed: MSLP response to the different boundary conditions/CO2. We propose to discuss the Eoi400 vs E280 results in more detail in the Discussion, in the section where we currently discuss nonlinearity (L499-506). Specifically, we propose to show a set of figures in the Supplement that we will then refer to, namely the MSLP difference and SLP SD difference for the following five combination (i.e. 5 panels per variable):

1. E560 – E280 (effect of CO2 doubling)
2. Eoi560 – Eoi280 (effect of CO2 doubling with mid-Pliocene BCs)
3. Eoi280 – E280 (effect of mid-Pliocene BCs)
4. Eoi560 – E560 (effect of mid-Pliocene BCs at high CO2)
5. Eoi400 – E280 ('true' mid-Pliocene conditions vs pre-industrial reference)

With this information, we believe that we can properly address nonlinearity, as well as the combined effects of CO2 and other mid-Pliocene BCs in the 'true' mid-Pliocene Eoi400 simulations.

Figure 2. Is any of the signal over Greenland in panel b likely to be because of differences in the height of the ice sheet, and how the interpolation to sea level is performed?

In panel b not, because the GIS is the same in E280 and E560. In panel c (Eoi280 – E280) it could be, since the GIS is different in both simulations. The difference in height, and as such the interpolation of pressure to sea-level, might explain the slightly 'spotty' differences over Greenland. We don't think it is necessary to mention this in the manuscript.

Fig. 3. In understanding the sea ice changes, it seems there are significant changes in land-ocean boundaries (as shown in the coastlines in panel c, which is different to that for present day) but this doesn't seem to be discussed. For example, there looks like there is large sea ice retreat over Hudson's bay, but this looks to be more related to shifting coastlines? Mentioning this would be useful.

You are correct, some of the sea-ice changes in the Eoi280 are due to coastline changes, for example over the Hudson Bay area, but also the Bering Strait. We will make a mention of this in the revised manuscript: "The Eoi280 shows sea-ice retreat over areas that are seas in the present-day geography, but were land in the mid-Pliocene, for example in the Bering strait and Canadian Arctic Archipelago."

Section 3.1. The results are useful for putting the variability differences in context of mean climate changes, but this isn't mentioned explicitly in the text, so the reader is left to add this interpretation themselves. Also, the section is very descriptive, with little attempt to provide mechanistic or physical explanations for the differences you see. For example, can you give some suggestions as to why you see such a strong response in the upper level circulation in response to Pliocene surface boundary conditions? Do you think this is related to differences in topography? Or tropical SSTs? Or something else?

As mentioned before, we will explicitly add the subquestion "How do changes in mean winter climate relate to changes in atmospheric variability?", to place the mean winter results in context of answering the research question. Next to that, as explained before, we will move some of the interpretation that is currently presented in the Discussion to the Results section. In the current Discussion, specifically section 4.2, we already aim to answer some of the questions you pose here.

Fig. 5. Having the positive contours dashed and the negative dot-dashed is a rather subtle difference (although after looking at Fig. 6 I now realize they are also coloured – stating this

in the caption would help). I would find it easier to interpret the figures is positive was solid contours and negative dashed (and zero bold if you wanted to distinguish the zero contour). At least mention whatever convention you use in the caption to make it easy to get the information.

We use negative contours dot-dashed and positive dashed throughout all the figures. We have chosen to use a solid white zero line in the difference contours (fig5,6 a) but dashed black zero line in the EOF figures (fig5,6 others) for the following reasons: solid white is well visible in the difference plot because of the colormap used. In the EOF plots however, the colormap is diverging around white, so a white line to indicate the zero contour is almost not visible. A solid black line was confusing because of the coastlines. This is why we chose to use black dashed for the EOF figures. We will mention the convection in the captions where necessary.

Changes in variance explained – I think you should calculate uncertainties on these values (e.g. using bootstrap resampling), to get a sense of how robust the differences you are reporting are. With 200 months I would hypothesise that changes of 10% are robust, but it depends on low frequency variability in the model. But are differences of 63.9 vs 66.7, or 17.5 vs 20.3 robust, or just noise, I'm unsure.

Thank you for this useful suggestion. We are not sure if bootstrap resampling is the best choice, but we will introduce a method that can help us assess how robust changes in percentage of variance explained are. For example, by computing the EOFs and perc of var explained for every possible 50year sample in the 200year dataset, and using that to construct a 5-95% confidence interval, or a standard deviation.

Section 3.2.3. I am not sure what to conclude from the section – more context is required. I also wonder whether the correlation between the NHem leading mode and the zonal or azonal modes is sensitive to which mode is picked up as the leading mode in Nhem and therefore how meaningful it is in a physical sense? Is the second leading mode of Nhem in Eoi280 more similar to the leading mode in E560 and E280, and therefore would correlated better with NPac-a?

The aim of this section is to assess how the level of (tele)connection between the different modes in the different basins changes. We believe this is relevant information in order to answer to research questions. We also wondered if the NHem mode is not just a construct of the leading modes in the N Atlantic and N Pacific. Considering the spatial patterns (Fig 5,6), this seems to be the case. However, the correlation coefficients presented in figure 7 show the level with which the modes correlate in space (i.e. how similar their principal components are). Qualitatively assessing similarity of the spatial pattern is not enough to know this. For example, in the E560 (fig7b), the NAtl-a mode shows a stronger (anti)correlation with the NHem (-0.57), than the NAtl-z (+0.43). However, from fig 5 you would not so easily see that the NAtl-a correlates more strongly with the NHem than the NAtl-z.

Furthermore, in line of your earlier feedback regarding uncertainty estimates, we will check for all the correlation coefficients in fig7 if they are statistically significant (e.g. p-value<0.05).

Line 378. Could you provide more details for why you conclude that the more northern jet state is stronger? NPac-z seems to correspond more to a meridional shift of the jet strength (as implied in line 379, and from Linkin and Nigam 2008). You do mention in line 228 that NPac-z is linked to meridional 'modulation' of the Asian-Pacific jet, but it's a little unclear what this means (meridional shifts? Changes in jet strength?) and the reader is left to work out a lot of these details themselves.

We conclude that the northbound jet is stronger (in comparison to the southbound jet in the same phase, that is) from the fact that our simple method to determine jet latitude (lat of max U200 in this zonal mean slice) picks up the northern jet more often than the southern jet. If the southern jet was the stronger of the two, the scatter plot distribution at negative NPac-z PC values would be skewed towards the 20-30N latitudes. Instead, it is skewed to the 45-60N latitudes. We will clarify in L228 by rephrasing "meridional modulation of the jet" as "meridional displacement of the jet", of "variations in the latitude of the jet".

Line 381. "the correlation between the two NPac modes" to me implies that the 2 modes are correlated to each other (which they can't be, being EOFs), rather than (on second reading) one is correlated with the jet intensity and one with jet latitude.

Indeed, your second reading is correct. However, we realise that the current phrasing is confusing, so we propose to change it as follows: ".. as well as the correlation between the PNA and jet intensity, and the NPO and jet latitude, suggest the following: …"

Line 414. I assumed 'we want to explore' implied that you were going to analyse some more simulations with different boundary conditions, but it seems more just literature review.

We understand this can be confusing. As mentioned before, we will move parts of this section to the Introduction, so that it doesn't read as a literature review anymore. We still think this is a useful section in the Discussion, and although we have not performed any additional sensitivity experiments in order to investigate the difference between closed gateways or a reduced GIS, other have. We think that we can use the results of those studies to put our own results into context. However, we will rephrase it like that at the start of the paragraph, so that it is clear for the reader that we are not introducing any new simulations with CCSM4-Utr here.

Line 415-416. What about sea surface temperature changes? Surely these could impact the mean state, and thus potentially variability as well? Tropical SSTs in particular can have a large impact on extratropical atmospheric dynamics (e.g. https://journals.ametsoc.org/view/journals/clim/14/4/1520-0442_2001_014_0565_tiotss_2.0.co_2.xml). Indeed, in Chandan and Peltier (2018) they note that their model does not agree with some paleoproxies of SSTs in some tropical locations. Does your model have the same bias? This is a limitation that should be discussed in this paper.

- Specifically regarding L415-416: SSTs are not boundary conditions but rather an effect of boundary conditions, hence we don't include SSTs in the list of topics in L415-416
- Yes, we expect SSTs to influence atmospheric variability. Not just tropical SSTs but also higher latitude SSTs (see e.g. Hurwitz et al that we included in our references, https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2012JD017819) . Thank you also for sharing the Yin & Battisti paper. We will outline the link between pacific SSTs and atmospheric variability briefly in the Introduction (see also our answer to your question regarding the PDO), and return to it in the Discussion. We chose to stick with SAT since we are considering atmospheric dynamics and because it is a good proxy for SSTs as well, especially in the mean patterns.
- The teleconnection with the tropics is relevant to mention, as we briefly do in section 4.3 A climate variability point of view, but we believe it is out of the scope of this study to dive into the details. We are presently preparing another manuscript exploring the mid-Pliocene tropical /ENSO teleconnection with the North Pacific in the full PlioMIP2 ensemble.
- Our model does a reasonable job in representing mid-Pliocene tropical proxies, see e.g. Baatsen et al 2022 Fig 9 https://cp.copernicus.org/articles/18/657/2022/ and discrepancies mostly occur in coastal regions with strong boundary currents, that might not be properly resolved in our model resolution.

In most of your discussion I find it hard to understand which is new information and understanding produced by this paper, and which is taken from the literature – it seems like a lot of discussion of the literature, and perhaps some of this belongs in the introduction?

Thank you for this feedback. We feel we have already addressed this issue earlier in our answering.

Line 499. I'm not sure you really address the non-additivity here, you just discuss the literature saying that the responses might be non-additive? It seems like perhaps you don't have all the experiments you would need to address non-additivity: you'd need E400 to compare the differences to CO_2 forcing (E280 vs E400), boundary condition forcing (E280 vs EOI280) relative to both together: E280 vs EOI400). And how are you defining nonlinearity and nonadditivity distinctly?

Thank you for this comment, indeed we are currently just discussing that literature study tells us the responses might be non-additive. We feel this is a relevant point to make. Regarding the nonlinearity, we feel we have addressed this point earlier in our answering, when treating the Eoi560 simulation.