

Interactive comment on “Midlatitude atmospheric circulation responses under 1.5 °C and 2.0 °C warming and implications for regional impacts” by Camille Li et al.

Anonymous Referee #1

Received and published: 24 December 2017

This study presents an analysis of the large scale circulation changes found in the HAPPI ensemble which is designed to assess the changes that occur with an additional 0.5C of warming beyond a 1.5C increase. I think this is a useful contribution and will be a beneficial resource for other users of the HAPPI ensemble. The paper is well written but I do think some aspects of the analysis and metrics could be explained more clearly as outlined in my specific comments below. The one main aspect I found confusing about the manuscript was the measure of consensus among models as outlined in my general comment and I think some improvement might be needed in this area. But overall, I think these amount to minor revisions that I recommend being made before publication.

C1

General comments:

My main confusion lies in the metric f^2 . This is some measure, that I recommend be explained more clearly, of the consensus among the models relative to the magnitude of the internal variability. I think this is being calculated by the ratio of the standard deviation across models to the standard deviation across members i.e., if the noise due to internal variability is bigger than the spread among ensemble means then models will be considered to agree as $f^2 < 1$. I struggle a bit to see how this is a useful metric. It seems like this could result in a situation where the model ensemble means really don't agree on even the sign of the response but the noise is sufficiently large that this metric would suggest there is consensus. I may not be fully understanding this metric as I don't think it is adequately explained. But my feeling is that models don't exhibit the degree of consensus that would be suggested by the presence of dots on the figures. As an example, if I understand correctly what's in Figure 9, this is showing the range of ensemble mean jet shifts across the models. (I'm actually not completely sure on whether it's the ensemble means or whether it's the spread across all ensemble members, it's not very clear). But, if it is the ensemble mean spread, then this shows that models can range in having jet shifts of e.g., in the Pacific, -4 to +4 in DJF. Yet, the lack of dots in the North Pacific in Fig 4a tells us that there is a consensus among the models here. If the model ensemble means don't agree on the sign of the change, then I don't think defining there to be a consensus if the spread among the models is smaller than the internal variability is particularly useful. My main concern is that in almost all of the lat-lon plots, virtually all locations are described as having a strong consensus because there are no dots, but I have a hard time believing that to be the case and I think it's because this measure f^2 might not really be a measure of consensus, but whether the noise is bigger than the disagreement among models and in that sense I think it's misleading. I apologize if I'm misunderstanding this metric, but if so, then I think it needs to be described more clearly.

Specific comments:

C2

I90: About the specified SST's . Is there interannual variability or is it the climatology of that time period that is being imposed. Recommend making that clearer.

I93: It's not very clear whether the SST anomalies imposed are coming from the RCP simulations with one model e.g., the particular HAPPI model, or whether it's a CMIP5 ensemble mean. It's made clearer in the conclusions that it's coming from a CMIP5 ensemble mean but I recommend it be made clear at this point.

I131: I don't think sigma has been defined. I assume that's standard deviation, but recommend making that clear.

I139: I think f^2 should be explained in more detail rather than just referring to the Sansom paper. It's pretty unclear how this is calculated and I expect it shouldn't be too lengthy to explain. Is it just the standard deviation across models of the ensemble mean response divided by sigma?

I220: it's stated that the additional features in 2C compared to 1.5C are small. But in the North Atlantic, they look pretty large. I guess it depends how you define small, but I'm not sure what the basis is for stating that the anomalies that appear in the North Atlantic are small. They're close to the magnitude of the original 1.5C anomalies over North America.

I230-232: another potentially relevant reference here is Harvey et al 2014 Equator-to-pole temperature differences and the extra-tropical storm track responses of the CMIP5 climate models. *Clim Dyn*, 43, 1175–1182.

I237-239: Couldn't the ensemble mean response in the CMIP5 models be pointed to for verification of this statement. It's stated that the as the world warms more the upper level temperature gradients weaken and we have a poleward shifting of the jet. But I don't think this is true in the east pacific during DJF where the CMIP5 models by the end of the century under RCP8.5 show a pretty good agreement on an equatorward shifting. Suggest modification of the wording to reflect this.

C3

I263: In this discussion of the Mediterranean changes, the moisture budget analysis of Seager et al 2014 might be useful. (Seager et al 2014, Causes of Increasing Aridification of the Mediterranean Region in Response to Rising Greenhouse Gases, *J. Clim.*, 27, 4655–4676). There the changes in P-E are decomposed into the various moisture flux contributions. Indeed the transient eddy moisture flux convergence is reduced which backs up the statements made here. But there are also substantial contributions from the altered mean flow moisture flux convergence as well.

I268: I think it would be worthwhile being more specific about where the weakening of the mean westerlies is i.e., "weakening of the mean westerlies OVER NORTH AFRICA signals" because otherwise readers might assume this is referring to weakening of westerlies over the Mediterranean which would be confusing since the Mediterranean is near the zero line of the zonal wind change. Similarly at line 279: "changes in u850 and " → "changes in u850 over North Africa and"

I279: "of wind responses" → "of wind responses in the 2C experiment" (because it's not clear which experiment is being referred to here).

I290: "as defined in the sense of the changes in the multi-model mean in Fig 12d" is unclear. I'm not sure exactly what this means. Does this mean that the strongest 5% are taken from all members from all models pooled together? The same goes for the caption of Fig 13.

I291: It would seem that a useful way to put this discussion of the change in the extreme percentiles into the context of a comparison with the present day climate would be to assess at what percentile does the magnitude of the 95th/5th percentile of the 2C climate occur in the PD climate. Then a statement of the form "Winters with this extreme dryness occur 5% of the time under 2C but only occur XX% of the time in PD" could be made. Otherwise, this discussion doesn't really provide any information about the change in these extremes from PD and so because of that, I don't see how it's really useful at this point. Another way to draw a comparison would be to ask how

C4

much of a reduction compared to the PD climate does the driest 5% of PD members represent i.e., a number equivalent to the 27% that's quoted but for the driest 5% of the PD members.

Figure 1: I'm confused as to why the dots indicating a lack of consensus occur where they do. If I understand correctly, all models specify the same SST's and sea ice anomalies. If I don't understand that correctly, then I think it needs to be made clearer exactly what's done with the SSTs and sea ice. If that is correct, then I don't understand why dots are occurring around the sea ice edge and over the middle of the Pacific. I would have thought the surface air temperature would be very strongly constrained by the imposed SSTs or sea ice anomalies. If so, then why would the models differ in this region? Is it because this metric is being influenced by the degree of spread among the members and there is very little spread among the members so the small spread in the response across members is actually bigger than the spread across members. This relates to my main comment above and again I wonder to what extent this metric is a useful measure of model consensus.

Figure 8 caption: it's stated that this is showing the "stationary waves". I think it would be best to be more explicit about what is actually shown i.e., "500hPa eddy geopotential height"

Figure 9 caption: It's stated that the multi-model mean shift in the eddy driven jet for the PD is shown in grey. Firstly I don't see any grey in the figure and secondly, how would a shift be calculated for PD? I suspect this is an error in the caption and that a shift for PD isn't shown. Sorry if I'm missing it. I also think it needs to be stated more clearly whether this is the spread across ensemble means or spread across all members of all models (see my general comment above).

Figure 10 caption: I don't think this is showing "winter North Pacific eddy-driven jet" because it's showing all months of the year, not just winter.

Figure 16: suggest showing the box that's used for the composite of v in panel c rather

C5

than panel a. I'm not sure why it makes sense to have that in panel a, but perhaps the authors have some reasoning.

Technical corrections:

l146: "show multi-model mean" → "show the multi-model mean" l217: "weakening in southwest" → "weakening in the southwest" l218: "strengthening in northeast" → "strengthening in the northeast" l340: "increases over Iceland increase" → "increases over Iceland" l371: "yield show drying" → "yield drying" l413: suggest "investigations of how" → "investigations into how" Figure 2 caption: "mean esponse" → "mean response"

Interactive comment on Earth Syst. Dynam. Discuss., <https://doi.org/10.5194/esd-2017-107>, 2017.

C6