Response to Reviewer 1 (James Done)

We'd like thank Dr. Done for his insightful review of our manuscript, including his positive comments and suggestions for improvement. Responses to his suggestions are included below.

1. The claim is made that the experiment assesses future thermodynamic environments. I understand that the frequency distribution of ENSO phases is the same in current and future climate. But MPAS will permit some large-scale circulation change in response to the future SSTs so I'm not sure circulation change is small compared to the thermodynamic change. Please clarify the contributions of thermodynamic change and circulation change permitted by your experimental setup. Perhaps checking the magnitude of circulation difference fields would help. More generally, can you provide more guidance on how the future change results should be interpreted? How should we interpret future changes based on fixed ENSO phase frequencies but variable atmospheric circulation response?

   It is true that we are assessing more than just thermodynamic changes since MPAS is a global model allowing for circulation changes to occur. We have removed "thermodynamic" from L8 on page 1, L11 on page 2, L3 and L29 on page 6, L9 and L34 on page 8, and L2 on page 13. We have also added L12–15 on page 2 to further clarify our methods.

   In regard to the circulation changes, please see Figs. 1 and 2 below. Fig. 1 shows a stronger wintertime Pacific jet and a deeper Aleutian Low in the future simulations. Changes in the Atlantic are less severe, but show a weaker jet over the northeast US and a slightly weaker Icelandic Low. The cross-sections in Fig. 2 show an overall stronger jet in the future simulations, with a poleward shift in the maximum.

   Given GCM difficulties in representing ENSO, we felt that the approach taken was the most defensible. Future research will need to determine how ENSO phase frequencies will change with warming, but we agree that this is another assumption in our methods and have added a sentence acknowledging this on page 7, L27.

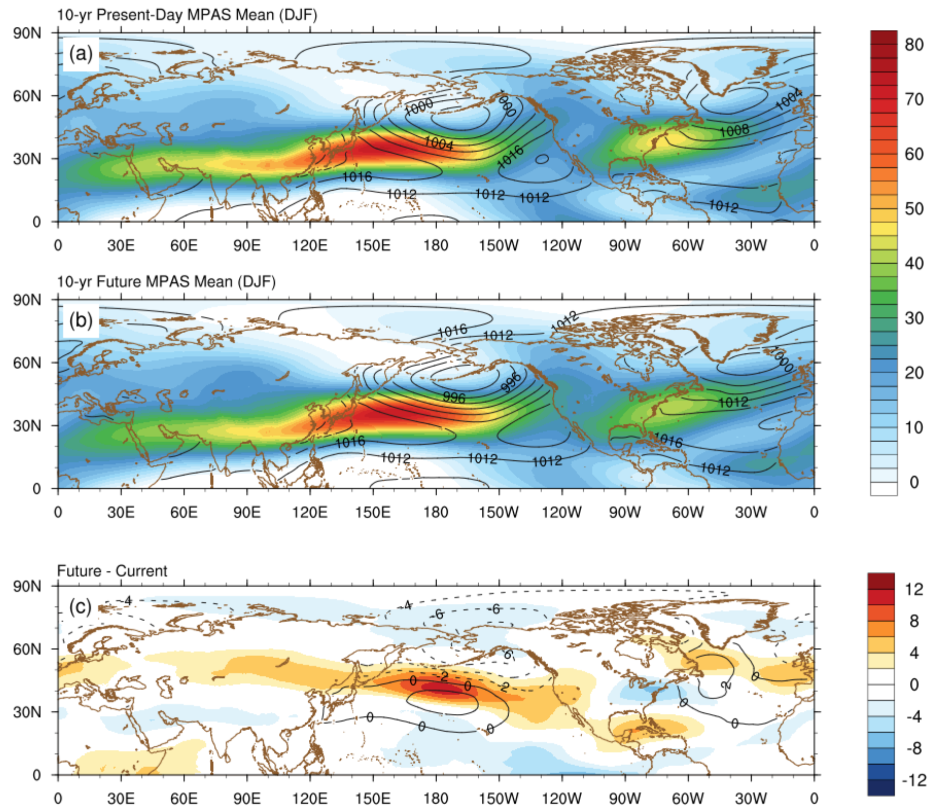

Figure 1. Average wintertime (DJF) 250-hPa zonal wind speed (ms^{-1}; shaded) and SLP (hPa; contoured) over the ten MPAS simulation years for (a) present-day, (b) future, and (c) future minus present-day 250-hPa wind speed (shaded) and SLP (contours). Wind contours are shaded every 2.5 ms^{-1} in (a), (b) and every 2 ms^{-1} in (c). SLP is contoured every 4 hPa in (a), (b) and every 2 hPa in (c).

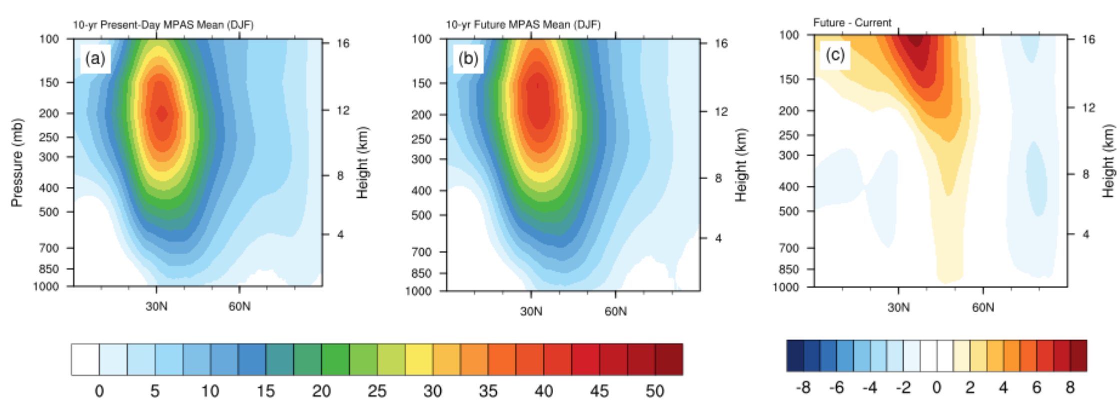

Figure 2. Average wintertime (DJF) cross-section of zonally averaged zonal wind (ms^{-1}; shaded) over the ten MPAS simulation years for (a) present-day, (b) future, and (c) future minus present-day. Contours are shaded every 2.5 ms^{-1} in (a), (b) and every 1 ms^{-1} in (c).

2. A related issue is that the experimental approach assumes that SSTs in all phases of ENSO change by the same magnitude and spatial pattern. Can you comment on how realistic this assumption is and any implications for the results?

   We agree that patterns of SST change may differ as a function of ENSO phase, and applying a single "delta" in our experiments is a limitation. However, we felt that this was the most straightforward starting point, given limited computational resources. We have added mention of this limitation, along with justification for our approach to L24–26 on page 7.

   We are aware of the "remote SST effect" on tropical cyclones, and that TC activity can exhibit considerable sensitivity to the patterns of SST change (e.g., Vecchi and Soden, 2007; Xie et al., 2010). Mizuta et al. (2014) have identified three distinct patterns of SST change in the CMIP5 CGCM output, and if resources were available, it would be quite interesting to run additional future experiments with these patterns in order to gauge the resulting sensitivity. One could also compute CGCM SST delta fields as a function of ENSO phase, and apply the deltas accordingly (perhaps what you are suggesting). We have added mention these possible future research directions.

   References:
   Mizuta, R., O. Arakawa, T. Ose, S. Kusunoki, H. Endo, and A. Kitoh, 2014: Classification of CMIP5 future climate responses by the tropical sea surface temperature changes. SOLA, 10, 167-171, https://doi.org/10.2151/sola.2014-035

   Vecchi, G. A. and B. J. Soden, 2007: Effect of remote sea surface temperature change on tropical cyclone potential intensity. Nature, 450, p.1066.

   Xie, S.-P., C. Deser, G. A. Vecchi, J. Ma, H. Teng, and A. T. Wittenberg, 2010: Global warming pattern formation: Sea surface temperature and rainfall. J. Climate, 23, 966−986, doi:10.1175/2009JCLI3329.1

3. Please explain why this study chose to sample ENSO phases and not phases of some other interannual of decadal mode of climate variability?

   ENSO is a global forcing phenomenon, and is often used as a basis for seasonal forecasts of tropical cyclone activity (e.g., Klotzbach and Gray, 2004; Camargo et al., 2007; Klotzbach, 2007) due to its strong connection with global tropical cyclone activity (e.g., Gray, 1984; Chan, 1985; Lander, 1994; Chu and Wang, 1997; Kossin et al., 2010). Since our initial motivation for conducting these simulations was to examine TCs and extratropical transition events, we wanted to sample a mode of variability with a large influence on global tropical cyclone activity, and therefore, chose ENSO. We have added this reasoning and associated references to L15–17 on page 5.

   References:
   Camargo, S. J., Barnston, A. G., Klotzbach, P. J., and Landsea, C. W.: Seasonal tropical cyclone forecasts, World Meteorological Organization (WMO) Bulletin, 56, 297–309, 2007.

Chan, J. C. L.: Tropical cyclone activity in the Northwest Pacific in relation to the El Niño/Southern Oscillation phenomenon, Mon. Weather Rev., 113, 599–606, doi:10.1175/1520-0493(1985)113<0599:TCAITN>2.0.CO%3B2, 1985.

Chu, P.-S., and Wang, J.: Tropical cyclone occurrences in the vicinity of Hawaii: Are the differences between El Niño and Non-El Niño years significant?, J. Climate, 10, 2683–2689, doi:10.1175/1520-0442%281997%29010<2683%3ATCOITV>2.0.CO%3B2, 1997.

Gray, W. M.: Atlantic season hurricane frequency. Part I: El Niño and 30 mb Quasi-Biennial Oscillation influences, Mon. Weather Rev., 112, 1649–1668, doi:10.1175/1520-0493(1984)112<1649:ASHFPI>2.0.CO;2, 1984.

Klotzbach, P. J.: Recent developments in statistical prediction of seasonal Atlantic basin tropical cyclone activity, Tellus A, 59, 511–518, doi:10.1111/j.1600-0870.2007.00239.x, 2007.

Klotzbach, P. J., and Gray, W. M.: Updated 6–11 month prediction of Atlantic basin seasonal hurricane activity, Mon. Weather Rev., 19, 917–934, doi: 10.1175/1520-0434%282004%29019<0917%3AUMPOAB>2.0.CO%3B2, 2004.

Kossin, J. P., Camargo, S. J., and Sitkowski, M.: Climate modulation of North Atlantic hurricane tracks, J. Climate, 23, 3057–3076, doi:10.1175/2010JCLI3497.1, 2010.

Lander, M. A.: An exploratory analysis of the relationship between tropical storm formation in the Western North Pacific and ENSO, Mon. Weather Rev., 122, 636–651, doi:10.1175/1520-0493(1994)122<0636:AEAOTR>2.0.CO;2, 1994.

4. The discussion of missing cool-wakes in the simulations on page 8 is well stated. But that is only half the story. The reanalysis SSTs will contain the cool wakes of observed TCs, so this could unphysically dampen simulated TCs that cross these ‘phantom’ cool wakes.

   This is a good point; however, the cold wakes in the OSTIA SST analysis field do not appear to be very strong. For example, Figs. 3 and 4 show a sample of times during the 2010 Atlantic TC season when two major hurricanes (Fig. 3: Danielle from 21–30 August and Fig. 4: Earl from 25 August–4 September) were present. While these are just two instances, it does illustrate that the phantom cold wakes in the SST analysis may not substantially impact the simulated TCs. The documentation of OSTIA does not explicitly state how SSTs are handled during TC conditions, but their interpolation to remove diurnal variability could be a factor in smoothing out the effects of cold wakes from TCs. Nevertheless, thank you for raising a valid point; we have added L18–20 and L26–27 on page 8 to acknowledge this limitation.

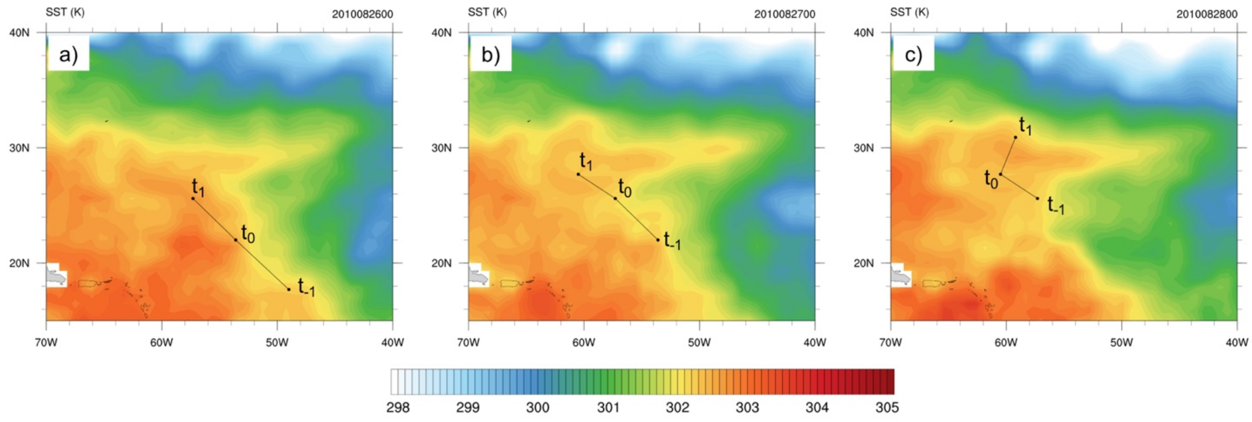

Figure 3: SST (K; shaded every 0.1 K) for (a) 00 UTC 26 August 2010, (b) 00 UTC 27 August 2010, and (c) 00 UTC 28 August 2010 when Hurricane Danielle was present in the North Atlantic. A subset of the storm track is shown in black with t_0 in each panel marking the position at the time in the top right corner, and t_1 and t_{-1} marking the positions 24-h before and after, respectively. Danielle was classified as a Major Hurricane in panel (b).

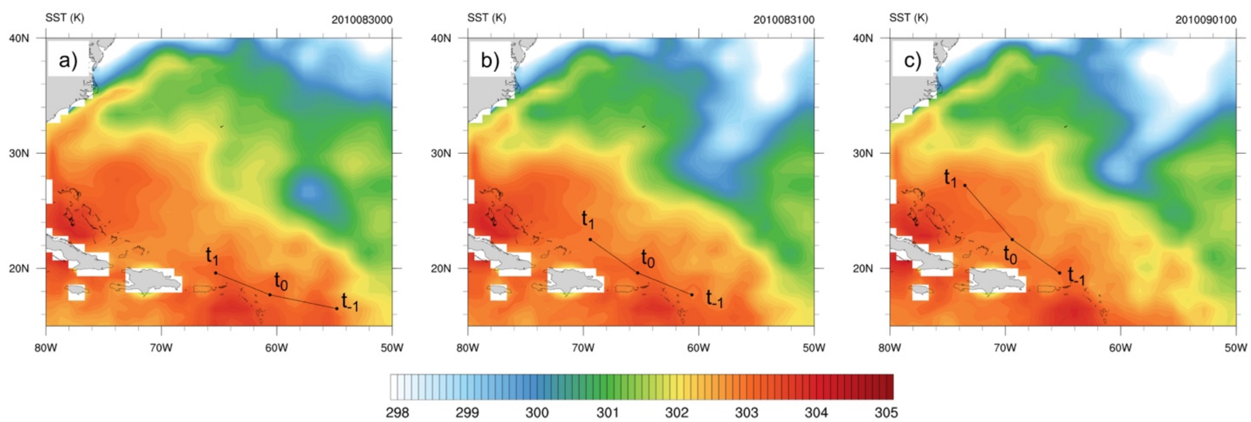

Figure 4: SST (K; shaded every 0.1 K) for (a) 00 UTC 30 August 2010, (b) 00 UTC 31 August 2010, and (c) 00 UTC 01 September 2010 when Hurricane Earl was present in the North Atlantic. A subset of the storm track is shown in black with t_0 in each panel marking the position at the time in the top right corner, and t_1 and t_{-1} marking the positions 24-h before and after, respectively. Earl was classified as a Major Hurricane in panels (b) and (c).

5. The evaluation shows MPAS does a reasonable job at capturing the climatological spatial distribution of TCs. Given that your approach emphasizes ENSO phases, can you also check whether MPAS captures the observed TC response to ENSO phase in the Atlantic and Pacific.

   Histograms of TC count by year are shown in Fig. 5. For the North Atlantic (Fig. 5a–b), MPAS does simulate a negative trend in TC activity indicating, in general, higher levels of TC activity during La Niña years and lower levels during El Niño years, consistent with expectations and the IBTrACS trend. For the Western North Pacific (Fig. 5c–d), however, MPAS simulates a *negative trend in TC activity, contrary to expectations and the positive trend exhibited by the IBTrACS. The trends in both data sets for the Western North Pacific, though, are not particularly strong.*

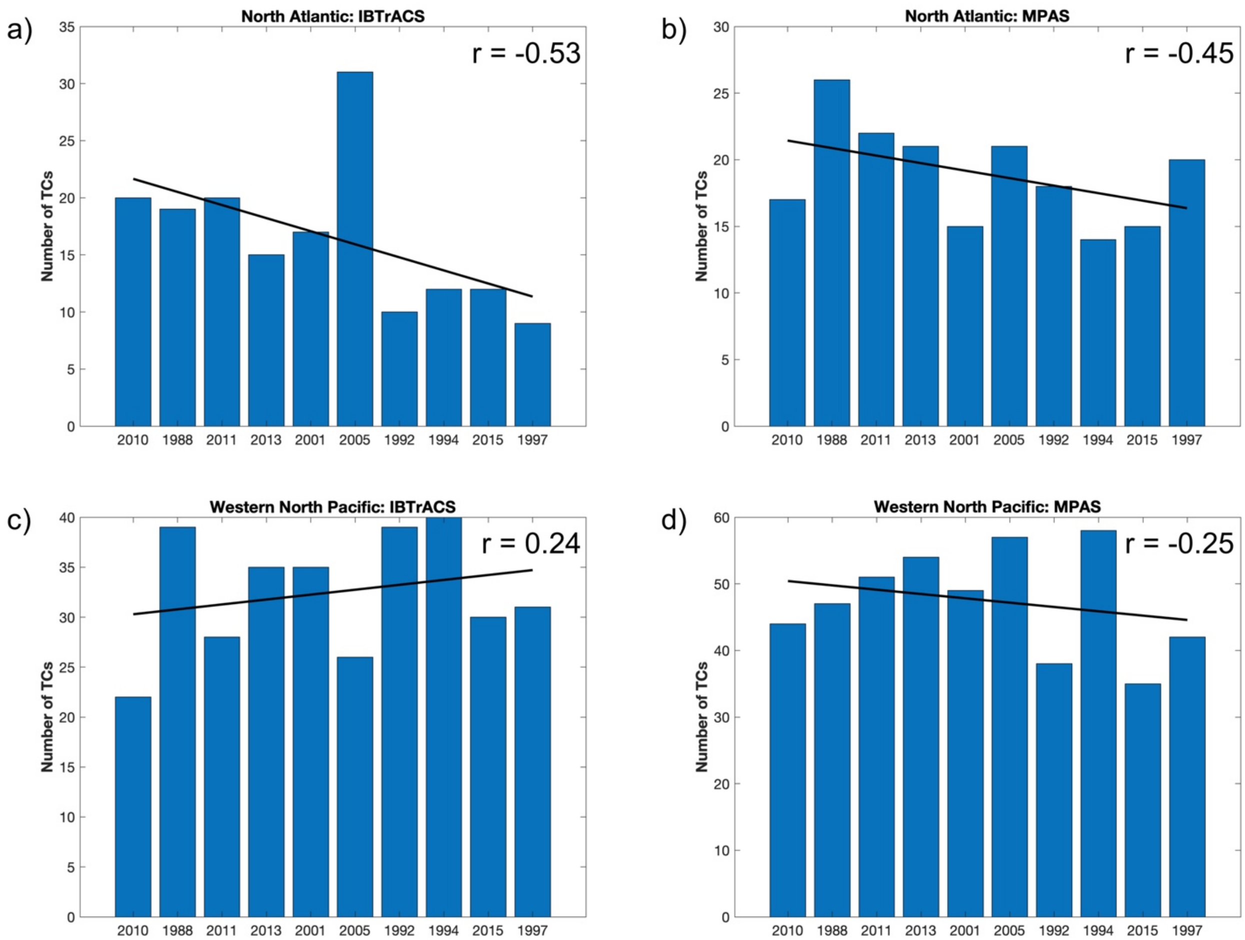

Figure 5: Histograms of TC count as a function of year for the North Atlantic basin for the (a) IBTrACS and the (b) MPAS simulations and for the Western North Pacific basin for the (c) IBTrACS and the (d) MPAS simulations. The years on the abscissa are ordered from the strongest La Niña year (2010) to the strongest El Niño year (1997). Linear regression lines are shown in black with the correlation coefficient in the top right corner in all panels.

6. MPAS misses TCs that develop from easterly waves over the eastern North Atlantic. I agree that one possible reason is enhanced shear. Another reason could be that the westward shift in the African Easterly Jet means that the wave energy accumulation zone (as discussed in Done et al. 2011) is also shifted westward.

   Done, J.M., Holland, G.J. and Webster, P.J., 2011. The role of wave energy accumulation in tropical cyclogenesis over the tropical North Atlantic. Climate dynamics, 36(3-4), pp.753-767.

   Thank you for this additional reference and reason for our displaced Atlantic TCs, this information has been added to L30–32 on page 11.

7. The purpose of Fig. 10 is to compare seasonal cycles of TCs. I suggest plotting the normalized distributions to remove the effect of differences in absolute numbers.

   Thank you for this suggestion. Fig. 10 has been updated with normalized probability histograms.

8. The motivation for this experimental design is to assess future changes in high- impact weather. But no results on future changes to high impact weather are presented. I read in the final section that results will be forthcoming in a separate publication. To set readers expectations I suggest adding a note earlier in the manuscript, perhaps at the end of the introduction.

   Thank you for this suggestion. A version of L34–35 from page 13 has been added to L21–23 at the end of the Introduction on page 4.

Response to Anonymous Reviewer 2

We'd like thank Reviewer 2 for their insightful comments and suggestions on our manuscript. Responses to their suggestions are included below.

1. The authors show the TC characteristics for the present climate, but not for the future climate, whereas they show in section 5 plots of the future climate. I was disappointed that nothing was said about TCs in the future climate. I assume that it will be analysed in a future paper, but for the reader it feels very disappointing. Why not postpone the climate change simulations altogether to a following paper? The added value of section 5 is very minor. Concentrating on the present climate and showing the quality of the model as in sections 3 and 4 should be fine for presenting the model and the experimental set-up. If you want to discuss climate change I would like it to be more than what is in the paper and include TC changes, or make it explicitly that they will discussed in a a future paper.

   Since we plan for climate change effects on TCs, and other high-impact weather events, to be the subject of future publications (several of which are either submitted or in preparation), we feel this manuscript should focus solely on the credibility of the current and future simulations. Therefore, we will retain Section 5 in this manuscript in order to avoid repeating it in each future publication. However, as you've suggested, we have added L21–23 at the end of the Introduction on page 4 and L21–22 on page 11 to help make this point as explicit as possible.

2. The experimental set-up is new in certain aspects, but I had the feeling that the authors over state a bit the uniqueness. They present it as global PGW simulations, whereas to me they are a clever way of time slice experiments with prescribed SSTs, with a spin-up to let the model come into equilibrium with the SST and CO2 concentrations. I found the PGW term therefore confusing and suggest to remove it.

   You are correct, our simulations are better described as time-slice experiments rather than global PGW experiments. We will retain L8–20 on page 3 in the Introduction as it is relevant background information. We also opt to retain L28–30 on page 6 as we feel comparing and contrasting our methods to the PGW technique is necessary. However, we had added L34 on page 6 to explicitly state that our future simulations are not global PGW experiments, but rather time-slice experiments. Additionally, our experiments are described as time-slice runs in L31 on page 3. It is our application of MPAS for climate-change purposes that we feel is truly unique, rather than the time-slice experiment method.

3. Ordering the simulations by El-Nino strength is clever, but anomalies outside El-Nino can occur, which can be substantial, and still have to be adjusted. I am not sure if one month spin-up is sufficient for these regions.

   We acknowledge that after one month, the spin-up to the new boundary conditions may be incomplete; however, during our initial evaluation, no obvious "shocks" appears at these inter-year boundaries after a couple of weeks, leading us to believe that one month was a sufficient amount of spin-up time. Additionally, other users who wish to use the output could *discard more of the beginning portions of the simulations if they feel more spin-up time is needed; discarding the first 2.5 months would still yield full yearly simulations.*

4. The use of high resolution SST data and computation of future SST by taking the delta of an ensemble of low resolution GCMs is a good approach. However, the delta SST can be influenced by the biases in SST. This can be particular large at the western boundary currents, that are incorrectly represented by low-resolution GCMs.

   Because CMIP5 models do not resolve ocean fronts (e.g., associated with western boundary currents and eddies), there are likely to be systematic errors in their depiction of future changes in SST. They offer, however, the best estimates of such changes that are currently available, and our method retains realistically strong SST gradients.

<u>Response to Executive Editor David Ham</u>

1. The code and data availability section refers to MPAS, and to various data sets. However this is insufficient to allow the reader to reproduce the results in the manuscript. For this, the reader needs the configuration files or run scripts used to run each experiment, and any scripts which were used to post process or analyse the model data. The exact version of these should be persistently and publicly archived (for example on Zenodo) and cited from the manuscript.

   The initialisation files and surface update files containing sea-surface temperature and sea ice information are too large to upload to Zenodo given the 50GB limit (these files are ~30GB for each run and there are 20 experiments in total). Sample run scripts and sample post-processing scripts, however, have been uploaded to Zenodo; the "Data and code availability" section has been updated to reflect this new citation.

2. The reference to the MPAS code used is on GitHub. This is both impermanent, as MPAS might move off GitHub in the future, and fails to identify the exact version of the code which was used (the model text refers to version 5.1, but was this the 5.1 release, or just a version of the master branch taken when 5.1 was the release number?) To remedy this, the exact version of MPAS used should be publicly and persistently archived. Since MPAS is developed on GitHub, the GitHub-Zenodo integration may be the easiest way to accomplish this. See: https://guides.github.com/activities/ citable-code/.
   Similar issues apply to Tempest Extremes.

   Because we do not host the MPAS or TempestExtremes GitHub repositories, we are unable to use the GitHub-Zenodo integration to make these repositories citable. However, we have updated the MPAS website in the "Data and code availability section" to point directly to the version used in our study, and specified that we used the release version 5.1. We have also updated the citation of TempestExtremes in the "Data and code availability" section following the TempestExtremes source manuscript (Ullrich and Zarzycki, 2017).

3. The result data is only available "on request". If at all possible, this data should be persistently and publicly archived so that the reader who wishes to investigate a result in the paper can do so directly. However, if this is not possible for licence reasons or because of the data volume, the data needs to be sufficiently precisely identified that it will still be retrievable if the authors have moved on from the host institution.

   The data volume is too large (~50TB) to host on a public repository like Zenodo. We have spoken with our host institutions at North Carolina State University and University of California San Diego, but they are unable to provide a permanent volume with a DOI for data storage. We will seek to store these data on the U.S. National Center for Atmospheric Research (NCAR) Computational and Information Systems Lab (CISL) archive. We welcome suggestions on how the data can be better identified in the manuscript.

4. The external data sets used are incorrectly cited, and in some cases poorly identified. Specifically:
   ERA-Interim data. The NCAR data archive provides DOIs for this data, and even tells you what to write in the bibliography. Please cite this in accordance with their instructions.

   This citation has been updated in the references and the link in the "Data and code availability" section has also been updated as noted by NCAR.

   TRMM data. Similarly, there is a data citation tab which tells you how to cite this data correctly.

   This citation has been updated as prescribed by GES DISC, and the link in the "Data and code availability" section has been updated to point directly to the version used.

   PRISM appear not to have precise data citation instructions, however they do have a precise convention for identifying the exact data set(s) used.

   Reference to this data set is a remnant from a previous version of the manuscript, and has therefore been removed from the "Data and code availability" section.