Referees' comments:
Bold-author response
Grey-revised manuscript changes

Review 2:

Author Comments:

We thank the reviewer for his/her helpful suggestions and comments on the manuscript. Below we include responses and revisions based off of the reviewer comments.

Concerns:
1. I am confused by the use of fluxes from a CCSM3 experiment in the new version of the model. If the CCSM3 experiment was unable in itself to reduce the SST gradient sufficiently, using the fluxes from that experiment in the slab ocean version of the new CESM1.0 seems to bias the new model towards not attaining the kind of gradient reductions required to match the observations. It is any surprise then that the new model cannot do what the observations, on the face of it, seem to require?

This comment presumes that the atmosphere has no significant mechanisms for polar amplification by itself, in the absence of increased ocean heat transport (OHT) which is counter to what is known about the subject (Holland and Bitz, 2003; Langen and Alexeev, 2007; Graversen et al., 2008; Graversen and Wang, 2009; Abbot et al., 2009). If positive OHT feedbacks were strong and atmospheric polar amplification was weak, then this might be an issue. On the contrary, every study that has been performed studying the sensitivity of zonal mean OHT to changes in a wide range of boundary condition alterations occurring throughout the Cenozoic robustly shows small changes in OHT, none of which go far to explaining high latitude warmth (Huber and Sloan, 2001; Krapp and Jungclaus, 2011; Zhang et al., 2011; Hill et al., 2013). Perhaps all of the coupled ocean-atmosphere models are incorrect in their prediction, but that is rather the point of our study. If OHT is anything close to what models predict (and they all give similar values for the Miocene relative to today), then atmospheric processes as represented in a state-of-the-art AR5 model are not sufficient to explain the past warm climate of the MMCO.

To be more clear about what we did with the slab ocean and why, we will include additional text explaining the slab ocean model framework and methodology in methods section 2.2.

Text added to experimental design section 2.2:
We include fluxes from a CCSM3.0 run at 560 ppm CO_2, which is a better match compared against the proxy records in comparison to the lower CO_2 simulation (Herold et al., 2011). We admit that this simulation has CO_2 levels above reconstructed values, but this is a better background climate state for conducting our simulations because of CCSM3's known low sensitivity to CO_2 forcing (Otto-Bliesner et al., 2006).

2. Overall I feel there needs to be a far better appreciation of the bias introduced into the comparison of observations and model output by being so inclusive with the observations. I do not see why the climate simulation should match observations from 3 million years before the MMCO when the model has been set up with, presumably, a best guess for the MMCO itself. It seems to me that a lot could change over those kinds of timescales which would be highly relevant in terms of the forcings given to the model.

The reviewer questions the comparison of our simulations with MMCO boundary conditions with some data that lie outside the MMCO. Generally speaking, I think that the reviewer's confusion may arise from considering the spatial and temporal variability of the Quaternary in temperature as representative of the lower gradient, less dynamic climate of the Miocene. Put simply, the temporal sampling and spatial sampling issues diminish in these warm climates, which can be gleaned by looking at proxy time series or by subsampling climate model output.

Our specific response is two-fold. First, the vast majority of the data lie within the MMCO, but we did include some data primarily from the tropics which is exceptionally data poor and also (by virtue of being the tropics) relatively homogeneous temporally and spatially. This homogeneity is easily seen in the data for the time interval that do exist (See Figure 1). The important issue remains explaining the pronounced global warmth and the high latitude records. In our current proxy record compilation the high latitude points are all within the time frame for the MMCO.

Second, the boundary condition changes between the early Miocene and the late middle Miocene are small (primarily vegetation, since OHT changes very little) and not generally important for the major paleoclimate questions being asked here. To verify this statement we ran additional simulations more representative of the late Miocene including imposing a realistic AIS reconstruction and altering vegetation including adding the Saharan Desert (Micheels et al., 2009). This simulation in comparison to the MMCO produces negligible changes in global MAT (-0.33˚C) (Figure 2) and small changes in the tropics. Our late Miocene simulation MAT is quite consistent with Bradshaw et al., 2012, which conducted an extensive comparison with records against late Miocene simulations. Important points are, 1) that tropical temperatures between the MMCO and late Miocene within this model framework are not very different. 2) The tropics are where the MMCO simulations matched the data the best in both the MMCO and late Miocene simulations. Thus including tropical data in our compilation which may lie outside the MMCO is reasonable.

In the revised text we will add section 5.3 in the discussion, describing additional biases and caveats produced by our approach.
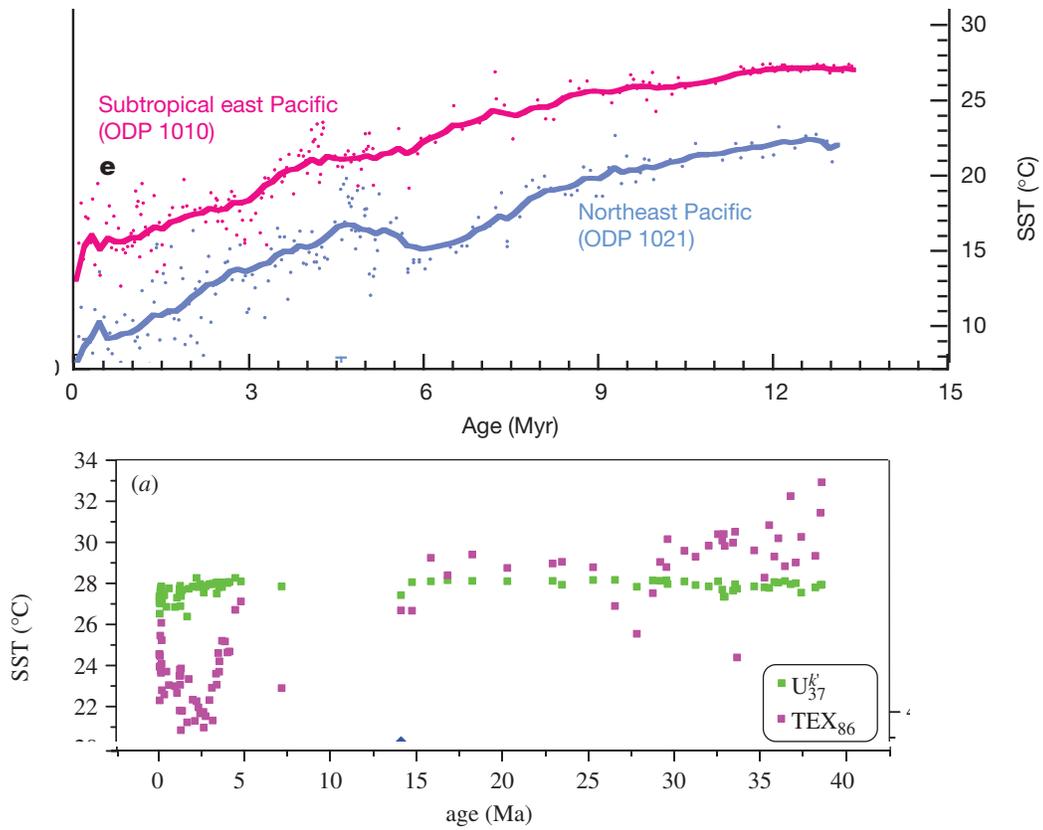
Figure 1. Tropical and mid-latitude SST compilation from LaRiviere et al., 2012 (top) and tropical SST compilation from Zhang et al., 2013 (bottom).
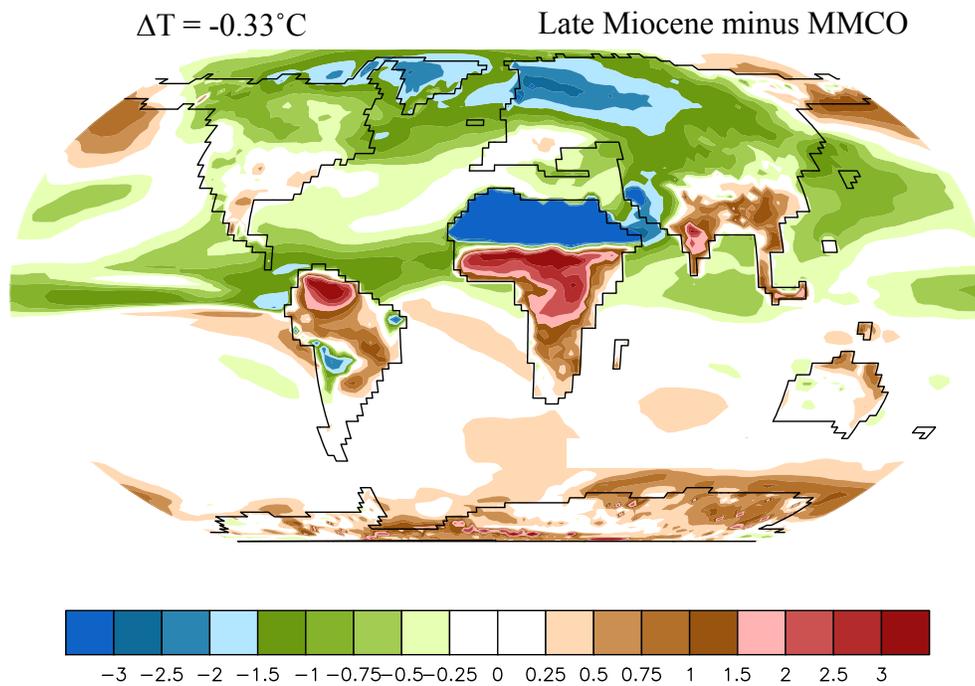

Figure 2. Surface temperature change (°C) between late Miocene and MMCO.

3. The attribution of the same tropical state in the Miocene as in the Pliocene is certainly conjecture in the absence of sufficient observations to support the premise. The reference to Pliocene tropical SSTs being 4 to 6 C warmer than present day does not match my own understanding of those papers which present data specific to upwelling zones in the tropics but that is not all of the tropics.

Our wording in this section was confusing and strained. The main point we were trying to convey is that the mid-Miocene, based on all existing qualitative and quantitative indications was characterized by a tropics/subtropics as least as warm as the Pliocene, if not warmer (Figure 1), and perhaps more importantly where we have records they are not wildly variable. This characterization is supported by the recently published SST data set of Zhang et al., 2013. This provides confidence that choosing SSTs outside of the MMCO strictly speaking is not likely to skew our results. While sampling is sparse and we reach slightly previous to or after the MMCO the regions for which we have data suggest that this introduces a sampling error which is smaller than the error bars we have placed on the data.

Revised text in section 2.3:
Given that the Pliocene tropical SSTs were ~ 4–6°C (Brierley et al., 2009; Dekens et al., 2007; Ravelo et al., 2006; Fedorov et al., 2013) above modern in the tropical and sub-tropical upwelling zones and the late Miocene were ~7–9°C above modern (LaRiviere et al., 2012) it is reasonable to conjecture MMCO tropical SSTs were this warm or warmer. This characterization is supported by the recently published SST data set of Zhang et al., 2013.

4. In section 3.1 reference is made to a PI simulation - what PI simulation is this?

Experimental design section 2.2 has a discussion of the Pre-Industrial (PI) simulation. The simulation is used as a baseline for comparing our globally averaged MMCO MAT to PI. Model statistics for this simulation are presented in Table 1.

2.2 Experimental design

The control Pre-industrial (PI) simulation employs the modelling components described above in standard configuration and with CO_2 concentrations set at 287 ppm. The slab ocean forcing file for the PI case has heat fluxes, salinity, and density inputs from a fully coupled atmosphere, ocean, ice, and land simulation (Bitz et al., 2012). Additionally we run a PI simulation at 400 ppm CO_2 (PI400) to compare with our MMCO simulation (also at 400 ppm CO_2). This high CO_2 PI configuration allows us to isolate the temperature effect of including MMCO boundary conditions at constant CO_2.

5. I am not sure that 51 sites with a poor global coverage encompassing the last 6 million years would actually give a valid estimate of modern global mean temperatures. Therefore, the use of this technique to support the method of deriving observational-based estimates of Miocene global temperature seems very unproven to me.

We realize that the technique is new, but this approach was recently published in a study conducted by Caballero and Huber, 2013, exploring Eocene paleo-sensitivity (see methods from this study). Our validation of this technique is covered in the text. We reiterate the main validation points here. We have proven that our technique (using the same geo-localities as the records) can reproduce modern climate using an independent modern observational dataset (ERA40, or using a CESM1 modern simulation) to within a tolerance of less than 2.2°C. This is a crude estimate by the standards of late Holocene climatology, but is more than sufficient to allow us to make the case we make in this study.

We have proven that the same technique can reproduce the global MAT of our Miocene simulations, i.e. spatial sampling is introducing less than 1.2°C of error in a perfectly known distribution. We note, that, given that the model, to the degree that is it known to be biases, over-predicts spatial gradients in temperature, the technique is probably doing an even better job at reconstructing global means than is indicated by this model-derived test.

These two independent tests highlight that the technique can recreate a global MAT value compared to the globally weighted MAT we calculate from the modern observations and model within a tolerance of 2.2°C. We understand that there are biases in the proxy records, but we account for this in the calculation of our standard errors.

6. I see nothing convincing that the error bars given on the observations will encompass all the temporal variability over such a large timeframe. This seems more like a statement of hope rather than proven statement.

From the reviewer comment it sounds like we agree that proxy record bias is difficult to constrain and in some cases underrepresented (for many different reasons). For this reason we intentionally inflated the error bars on the proxy data reconstructions to encompass not only the stated error of the methodologies but to represent the 'unknown unknowns' in the proxy interpretations. Our justification for making the error bars larger and within the range recent published studies (Royer et al., 2012, Grimm and Denk, 2012) was not out of hope, but rather to encompass the temporal variability and other sources of uncertainty the reviewer suggests. The magnitude of this uncertainty is still boundable by comparison with the relatively well understood and constrained benthic $D^{18}O$ records which are compatible with the error bars we have used in this approach.

To address the reviewer question about the temporal variability across the MMCO we use a stacked oxygen isotope record for the early to middle Miocene (Zachos et al., 2008) and calculate the approximate global mean temperature variations, relative to the MMCO. Based on the known temperature relationship to oxygen isotopes and applying a factor of 0.5 to account for polar amplification we find that the difference in temperature before and after the MMCO is -0.836˚C and -0.997˚C, respectively. This is likely to be an overestimate since some fraction of the isotope record (normally one assumes something like half) is representing ice volume, not temperature. This interpretation is buttressed by more sophisticated studies like that of Liebrand et al., 2011. Thus, our error bars encompass the likely variation in global mean temperature that may be present in the records used here.

Added text in section 2.3

To address the temporal variability of including records across a long timescale using a stacked oxygen isotope record for the early to middle Miocene (Zachos et al., 2008) we calculate approximate global mean temperatures, relative to the MMCO, based on the known temperature relationship to oxygen isotopes and applying a factor of 0.5 to account for polar amplification. We find that the difference in temperature before and after the MMCO is -0.836˚C and -0.997˚C, respectively. Thus, our error bars encompass the likely variation in global mean temperature that may be present in the records used here. This interpretation is buttressed by more sophisticated studies like that of Liebrand et al., 2011.

My view is that uncertainties in the observations are too lightly treated. Some of the fundamental questions are ignored, for example are the observations really of MAT or a growing season? This is especially important for high latitude sites where it seems to me far more likely that you will have a seasonal bias in the observations, rather than any representation of MAT. If so the provided observational error could fall short of reality at such locations.

During our compilation of proxy records we took special care to compile estimates from the studies which aimed to reconstruct MAT and also took records that were thoroughly vetted in published work (Pound et al., 2012, Herold et al., 2011). Although we realize there can be signals included which may be more representative of "growing seasons" or seasonal cycles we did many types of model data comparisons before choosing to show the MAT pointwise comparisons (We show these below). We found that cool biases existed in the model simulations whether we compared records against the modelled seasonal temperature variations or MAT. With this being said we hope to remedy this in the text by further describing the possible proxy record biases and how we accounted for them in methods and in the discussion section 5.3. Overall our attempt to model the MMCO and compare with available records should be considered a best attempt for comparing a dataset of this size against a model simulation.

Added text section 5.3:

The issue of some terrestrial records especially in mid-to-high latitude regions having a difficult time matching modelling simulations has been pointed out more extensively in a model data comparison of the Eocene conducted by Huber and Caballero, 2011. Possible reasons behind the model data mismatch include under sampling for leaf physiognomic techniques (Wilf, 1997), skew of high latitude temperatures due to the "toothiness" (Boyd et al., 1994), and other systematic biases mentioned below (Burnham et al., 1989; Spicer et al., 2005; Peppe et al., 2010). Additionally, due to the length of the period we explore there is the question of whether the record is recording a seasonal or MAT signal.

Other factors related to the model boundary conditions could be contributing to the modelling simulations being too cold compared to the terrestrial records. The modelling framework is limited to the fact that paleoelevation in some areas is simplified due to resolution and also poorly constrained in other regions (Herold et al., 2008). This could significantly effect MAT temperatures in the mid-latitude regions depending on the imposed elevation especially if specific locations are too high. Additionally, further altering orbital variations (Haywood et al., 2013) and Miocene vegetation cover (Herold et al., 2010) in the high latitudes can have a significant effect on the regional temperature over the land surface. Finally, we point out that a generally good match for terrestrial records occurs at 800 ppm CO_2 (Figure 3c,d) as there is enhanced high latitude warmth compared to the tropical regions. This again highlights that we are missing a forcing equal to a doubling of CO_2 to explain MMCO warmth. Thus finding the right combination of boundary condition changes and the right model framework will continue to improve the model-data mismatch regionally.

7. The method of DMC seems strange. By plotting absolute temperatures the effect of latitude on temperature will blur the assessment of model performance (make it look better than it actually is). A better test might be compare anomalies in the observations versus the anomalies from the model.

We agree that there are various approaches to presenting the models versus the data, although we note that making the models look too good, as the reviewer comments is not a problem that we have. Here we present the comparison in a couple of ways for reference. First we show below the various ways that we tried to plot the data before deciding to present the work as is shown in the manuscript.
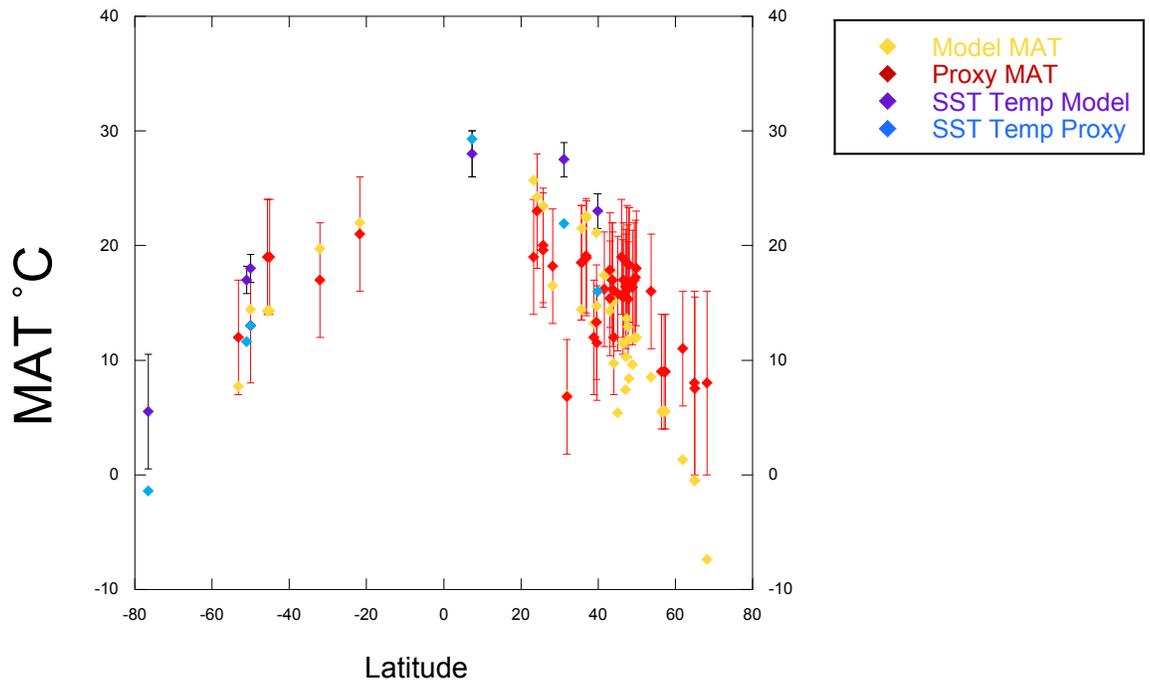
Figure 3. Model data comparison for terrestrial and SST records across latitudes for the MMCO simulation. Model MAT (yellow diamonds), proxy MAT (red diamonds) with error uncertainty as described in submitted manuscript. Same for SST records.
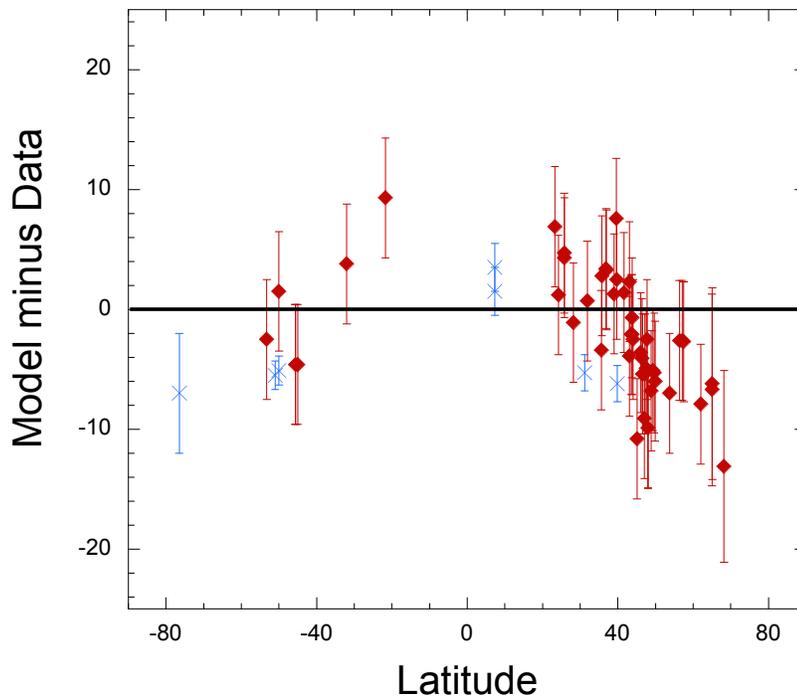

Figure 4. Modelled temperature (MMCO simulation) minus the proxy record MAT across all latitudes. Error bars represent proxy record uncertainty discussed in the paper.

There were two issues with the model data comparisons described above (Fig 3, Fig 4). First plotting points across latitudes causes clutter in the areas where there is more geospatial representation (Fig 3, Fig 4). Second, by plotting the model and data on the same axis we are unable to represent error/uncertainty for two variables. Significant effort was devoted to representing uncertainty in our proxy records, but also for our modelling simulations. We point out that we ran additional simulations changing orbital parameters to present the possible minimum and maximum temporal temperature changed induced by orbital uncertainty.

The plots the reviewer brings up are useful (Fig 5), but there are a couple of issues with this approach. First it makes it hard to isolate what specific geo localities are a poor match against the records. Second issue is what to choose for the control simulation, as the MMCO simulation is compared against the PI simulation whereas the proxy records are compared against observations. This means that the baseline control simulation is different between the comparison.

The model data comparison in the paper currently helps isolate regions where the match is poor without clutter, includes error bars for the both the proxy records and the modelled temperature variations (including orbital variations).
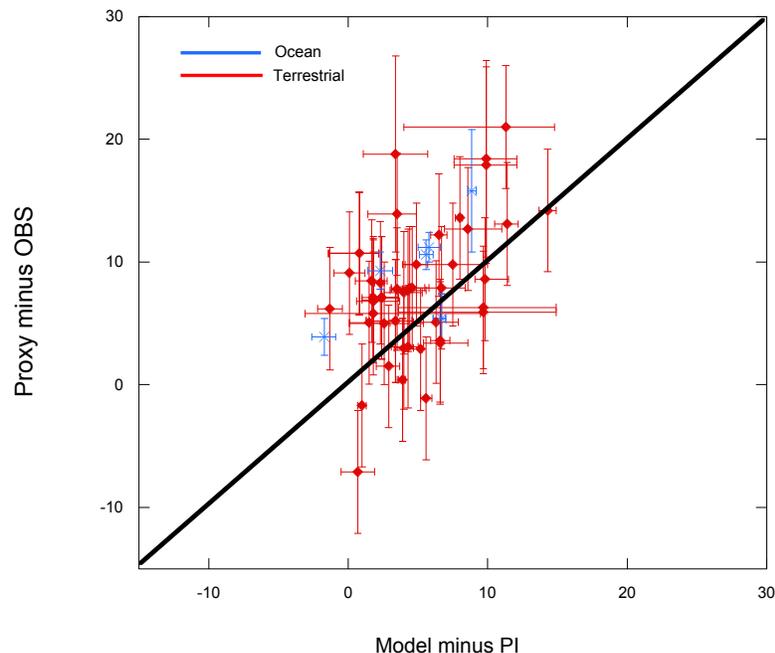

Figure 5. Modelled temperature for MMCO simulation minus PI simulation and proxy records minus observations. Error bars represent proxy record and modelled uncertainty discussed in the paper.

8. The combination of forcings in one of the sensitivity tests is interesting but I am left wondering how realistic that combination of forcings would actually be in the context of the MMCO. It seems like a 'throw everything at it including the kitchen sink'-type approach, but the validity of such an approach is not discussed

The literature (for example Fedorov and Philander 2003) has suggested that some of these features such as a permanent El Niño characterized all pre-Quaternary climates so we think it is reasonable to add the various forcings in tandem. We add some discussion about why we choose to explore the different sensitivity tests. Lowering the Antarctic Ice Sheet and adding an El Niño are two hypothesis which exist in the literature for the Miocene. Whereas the El Niño, lowered ice sheet, and increasing obliquity may be less justified. We show this combination of forcings because all increase high latitude warmth (an area where the model has not performed well in comparison to records).

Added text in results section 4.2
These boundary condition changes in conjunction with one another are not clearly detected in proxy reconstructions and we present this simulation as an example of including a suite of forcings aimed at increasing temperature in high latitude regions. Although higher than modern obliquity is reconstructed in orbital configurations calculated for the MMCO (Laskar et al., 2004).

9. Given the inherent uncertainties in the observations, which are not fully accounted for (truthfully because it is probably not possible to do that at the current time – if so the conclusions need to be more temperate to take account for that).

We find it unlikely that we are focusing on results that are not robust, given that exact same biases are known in all models for the Miocene, and that the same biases have been identified in the Cretaceous, Eocene, and Pliocene (consider the recent paper by Salzman et al., 2013). Nevertheless, we will caveat some of the conclusions in the submitted manuscript to make it clear that 1) we use one model framework, 2) are missing data from important regions globally, 3) that the records may be recording signals besides MAT.

The incomplete nature of the sensitivity tests, the use of a slab ocean model, and the fact that the entire analysis has been conducted with just a single model (thus model dependency is not explored), leaves me unconvinced about the basic conclusion that the model is too insensitive.

There seems to be some logical inconsistencies in the reviewers 3 points.

(1) Previously the reviewer found, the 'kitchen sink' approach to be unjustified, but at the same time the 'incomplete' nature of the sensitivity tests is a flaw. Would the reviewer advise more sensitivity tests or fewer to resolve this problem? We agree that there are

unfinished sensitivity tests and other model frameworks needed to continue understanding the MMCO. One reason we chose to use slab configuration (among others listed in the text) is to be able to explore a wide set of experiments to isolate which changes contribute to substantial warmth. We have tested the most common (and not so common) boundary condition changes understood to have a likely impact on MAT and even with all of these changes there is a negative bias in the model. We have done everything conceivable at this time to address these issues.

(2) We note that models routinely use slab model experiments as the operational definition of sensitivity (Gettelman et al., 2012, Shell et al., 2008) and that these produce very similar values to equilibrated coupled runs (Danabasoglu and Gent 2009, Caballero and Huber, 2013) so we do not think that is a relevant issue.

(3) Nor is it a problem to analyze just one model to to say that model itself is too insensitive. We can't say that all models are too insensitive from this one study (although a review of the literature on past warm climates suggests this is very much the case).

Added text in conclusions:
We understand that we are also limited in this study to the use of one model configuration and our conclusions should be understood within this context. We have tested the most common (and not so common) boundary condition changes understood to have a likely impact on MAT and even with all of these changes there is a negative bias in the model.

The point about missing forcings I can agree with, as this seems sensible given the wide timeframe considered, but there seems little within the paper to support the premise that the sensitivity of CESM1 is somehow fundamentally flawed.

We want to be clear that we do not pinpoint model sensitivity to CO_2 forcing (and weak sensitivity) as the sole explanation for the model data mismatch. It is included in a list of possible reasons for explaining the discrepancies we find in our study including missing other forcings like aerosols, vegetation, or other trace gases. To clarify our stance, if we look back at all the fully coupled simulations that did comparisons with records for the MMCO (Herold et al., 2011, Krapp and Jungclaus, 2011, Henrot et al., 2010) and for the Eocene (Huber and Caballero, 2011, Lunt et al., 2012), these studies come to a similar conclusions about model simulations being too cold compared against the records at reconstructed CO_2 levels.

It requires double the reconstructed CO_2 levels to match the data in both Eocene (Huber and Caballero, 2011) and Miocene (this result is presented in Fig 3). The models used in previous studies and in this study have similar sensitivity to CO_2 forcing (2-3°C per doubling). If we believe the CO_2 reconstructions for the Eocene and miocene, than it is

possible that the lack of simulated warmth may be caused by the model's being insensitive to CO₂ forcing.

Given how poorly constrained fast climate sensitivity is today, and how minor changes in tunable parameters (Jackson et al., 2008) alter the value of a model within modern constraints it seems quite plausible to us that even minor changes in model parameters that enhance climate sensitivity provide a parsimonious resolution to a range of paleoclimate 'problems' that extend from the Pliocene on back.

Also, I think the paper would benefit by having real input from the scientists who created the observational datasets that the model is compared to. That way a more convincing case can be constructed that the error bars are reasonable.

The authors have a long history of working closely with the data record generating and proxy development community. We are in daily communication with the generators of these records, for better or for worse. The authors have taken great care to discuss the model data comparisons with a variety of modeler's and geologists. The authors also have published a variety of papers with the data community who create the types of datasets and reconstructions used in our proxy record compilation (Lyle et al, 2008; Herold et al., 2011, Liu et al., 2009, Eldrett et al., 2009, Galeotti et al., 2010; Taylor et al., 2013). Many of the model data comparison techniques and error uncertainty are a result of this collaboration.

References:

Abbot, D. S., Huber, M., Bousquet, G., & Walker, C. C.: High−CO2 cloud radiative forcing feedback over both land and ocean in a global climate model. *Geophysical Research Letters*, *36*, 2009.

Bradshaw CD; Lunt DJ; Flecker R; Salzmann U; Pound MJ; Haywood AM; Eronen JT, Erratim: The relative roles of CO and palaeogeography in determining late Miocene climate: Results from a terrestrial model-data comparison, *Climate of the Past*, 8, 1301-1307, 2012.

Caballero, R., and Huber, M.: State-dependent climate sensitivity in past warm climates and its implications for future climate projections PNAS, 110, 14162-14167, 2013.

Danabasoglu, G., and Gent, P. R.: Equilibrium climate sensitivity: Is it accurate to use a slab ocean model?. *Journal of Climate*, *22*, 2494-2499, 2009.

Eldrett, JS, DR Greenwood, IC Harding, M Huber, Increased seasonality through the Eocene to Oligocene transition in northern high latitudes, Nature 459, 969-973, 2009.

Fedorov, A. V., Brierley, C. M., Lawrence, K. T., Liu, Z., Dekens, P. S., and Ravelo, A. C.: Patterns and mechanisms of early Pliocene warmth, Nature, 496, 43–49, doi:10.1038/nature12003, 2013.

Galeotti, S., A von der Heydt, M Huber, D Bice, H Dijkstra, T Jilbert, L Lanci, Evidence for active El Niño Southern Oscillation variability in the Late Miocene greenhouse climate, Geology 38, 419-422, 2010.

Gettelman, A., Kay, J. E., & Shell, K. M.: The evolution of climate sensitivity and climate feedbacks in the community atmosphere model. *Journal of Climate*, *25*, 1453-1469, 2012.

Graversen, R. G., Mauritsen, T., Tjernström, M., Källén, E., & Svensson, G.: Vertical structure of recent Arctic warming. *Nature*, *451*, 53-56, 2008.

Graversen, R. G., and Wang, M.: Polar amplification in a coupled climate model with locked albedo. *Climate Dynamics*, *33*, 629-643, 2009.

Grimm, G. W. and Denk, T.: Reliability and Resolution – a revalidation of the coexistence approach using modern-day data, Rev. Palaeobot. Palynol., 172, 33–47, doi:10.1016/j.revpalbo.2012.01.006, 2012.

Herold, N., Huber, M., and Müller, R. D.: Modeling the Miocene Climatic Optimum. Part I: Land and Atmosphere, J. Climate, 24, 6353–6372, doi:10.1029/2010PA002041, 2011.

Herold, N. K., Huber, M., Greenwood, D. R.; Müller, R. D. and M. Seton, Early to middle Miocene monsoon climate in Australia, Geology, 39, 3-6, doi: 10.1130/G31208.1, 2011.

Hill et al., Paleogeographic controls on the onset of the Antarctic Circumpolar Current, GRL, in press, 2013.

Holland, M. M. and Bitz, C. M.: Polar amplification of climate change in coupled models. *Climate Dynamics*, *21*, 221-232, 2003.

Huber, M., & Sloan, L. C.: Heat transport, deep waters, and thermal gradients: Coupled simulation of an Eocene greenhouse climate. *Geophysical Research Letters*, *28*, 3481-3484, 2001.

Jackson, C. S., Sen, M. K., Huerta, G., Deng, Y., and Bowman, K. P.: Error reduction and convergence in climate prediction. *Journal of Climate*, *21*, 6698-6709, 2008.

Krapp, M., & Jungclaus, J. H.: The Middle Miocene climate as modelled in an atmosphere-ocean-biosphere model. *Climate of the Past*, *7*, 1169-1188, 2011.

Langen, P. L., & Alexeev, V. A.: Polar amplification as a preferred response in an idealized aquaplanet GCM. *Climate dynamics*, *29*, 305-317, 2007.

LaRiviere, J. P. Ravelo, A. C., Crimmins, A., Dekens, P. S., Ford, H. L., Lyle, M., and Wara, M. W.: Late Miocene decoupling of oceanic warmth and atmospheric carbon dioxide forcing, Nature, 486, 97–100, doi:10.1038/nature11200, 2012.

Laskar, J., Robutel, P., Joutel, F., Gastineau, M., Correia, . C. M., and Levrard, B.: A long-term numerical solution for the insolation quantities of the Earth. Astronomy & Astrophysics, 428, 261-285, 2004.

Liebrand, D., Lourens, L. J., Hodell, D. A., De Boer, B., Van de Wal, R. S. W., and Pälike, H.: Antarctic ice sheet and oceanographic response to eccentricity forcing during the early Miocene. *Climate of the Past*, *7*(3), 869-880, 2011.

Liu, Z, M Pagani, D Zinniker, R DeConto, M Huber, H Brinkhuis, S.R. Shah, R.M.: Global cooling during the Eocene-Oligocene climate transition, Science, 323, 1187-1190, 2009.

Lunt, D. J., Dunkley Jones, T., Heinemann, M., Huber, M., LeGrande, A., Winguth, A., Loptson, C., Marotzke, J., Roberts, C. D., Tindall, J., Valdes, P., and Winguth, C.: A model – data comparison for a multi-model ensemble of early Eocene atmosphere--ocean simulations: EoMIP, Clim. Past, 8, 1717–1736, doi:10.5194/cp-8-1717-2012, 2012.

Micheels, A., Eronen, J., and Mosbrugger, V.: The Late Miocene climate response to a modern Sahara desert, Global Planet. Change, 67, 193–204, 2009.

Philander, S. G., and Fedorov, A. V.: Role of tropics in changing the response to Milankovich forcing some three million years ago. *Paleoceanography*, *18*, 1045, 2003.

Pound, M., Haywood, A. M., Salzman, U., and Riding J. B.: Global vegetation dynamics and latitudinal temperature gradients during the Mid to LateMiocene (15.97–5.33Ma), Earth-Sci. Rev., 112, 1–22, doi:10.1016/j.earscirev.2012.02.005, 2012.

Shell, K. M., Kiehl, J. T., & Shields, C. A.: Using the radiative kernel technique to calculate climate feedbacks in NCAR's Community Atmospheric Model. *Journal of Climate*, *21*, 2269-2282, 2008.

Taylor, K. W. R., M. Huber, C. J. Hollis, M. T. Hernandez-Sanchez, and Pancost, R.D.: Re-evaluating modern and Palaeogene GDGT distributions: Implications for SST reconstructions, Glob. Planet. Change, 108, 158–174, 2013.

Zachos, J.C., Dickens, G.R., and Zeebe, R.E.: An early Cenozoic perspective on greenhouse warming and carbon cycle dynamics. Nature, 451, 279-283, 2008.

Zhang, Z., Nisancioglu, K. H., Flatøy, F., Bentsen, M., Bethke, I., and Wang, H.: Tropical seaways played a more important role than high latitude seaways in Cenozoic cooling. *Climate of the Past Discussions*, *7*, 965-996, 2011.

Zhang, Y. G., Pagani, M., Liu, Z., Bohaty, M., and DeConto, R.: A 40-million-year history of atmospheric CO_2, Phil. Trans. R. Soc. A, 371, 2013.