Climate sensitivity is a key parameter in the understanding of the climate behavior and therefore in the prediction/projection of our future climate. Such a paper dealing with this topic is therefore very welcome. This paper is in addition dealing with the climate sensitivity as a function of the background climate state, a research that started worldwide a few years ago and must be encouraged. Climate sensitivity must definitely be differentiated between warm and cold climates. Finally this paper couples data and models to show the state-dependency of climate sensitivity over a very long period (5 million years) which includes a large number of extreme climate situations. All these made the review favorable to the publication of such a paper, but with revisions of some points discussed here under.

General remarks

What is important for the future is to know whether the increase of temperature due to a doubling of the present-day (pre-industrial) CO2 concentration is equivalent or larger or smaller than a similar doubling during the previous interglacials (times when ice was similar as to-day). Such a climate sensitivity is different from the one used in this paper ($K/(Wm^{-2})$). For example, using a climate sensitivity restricted to the change of global temperature for a doubling of CO2, Yin and Berger (2012, Climate Dynamics) have stressed : "Within the range of the interglacial variability with the CO2eq concentration going from 234 to 300 ppmv, our climate sensitivity is shown to generally decrease with increasing temperature: MIS-9 has the lowest sensitivity and MIS-13 the highest. The sensitivity at MIS-5 is 10% lower than at Pre-Industrial time". The same results transferred in $K/(Wm^{-2})$ gives a decrease from 0.41 (MIS-13) to 0.37 (MIS-9) (if ΔT is divided by 5.35*ln(2), i.e. 3.71)

Moreover, if the climate sensitivity (mainly to CO2) is indeed depending on the climate background, the results obtained from cold climates can hardly be used for improving the projection of our future climate (see page 3042 lines 17-19).

Previous works on climate sensitivity, Page 3022 lines 26 and mainly page 3023 line 4, conclusion page 3028 line : "during Pleistocene warm period S was about 45% larger than during the Pleistocene cold periods" and page 3041 lines 9-10 plead for Kohler et al. discussing such climate sensitivity considering only the interglacials/warm periods and only CO2 if possible(more detailed discussions than what is done in sections 2.3 and 3.3) .

This remark leads to the following recommendations. The authors say on purpose that their analysis is going beyond what has been done before. It would therefore be interesting to see the relative importance of each individual improvement to explain the differences from previous studies.

By introducing new data and calculations (see page 3023 bottom and page 3024 top), the authors introduce unintentionally also new hypotheses and sources of uncertainties. They discuss these uncertainties in section 2.5 and some other places in the paper, but what are the impact on the calculation of the climate sensitivity itself? Some conclusions are drawn in section 3.3 but it would be interesting to know, for example, which of the change of time series or resampling of CO2 data (page 3038 lines14-15) has the largest impact on S. This is very important for recommending in which direction studies must continue to be done to improve our knowledge.

Along the same lines:

What is the impact of the uncertainties of the reconstruction of paleoclimate data of the last 5 million years (in particular of ΔT_g) ?

What is the impact on the calculated radiative forcing of the land ice albedo from a 3-D ice sheet model uncoupled (?) to the rest of the climate system? Can the authors be a little bit more explicit on how they calculate ΔR_{LI} ? What else more than surface albedo, TOA and changes in ice-sheet area is needed to "estimate" ΔR_{LI} ? What is the relative impact of this "technique" on climate sensitivity?

What is the impact of fixing the value of a polar amplification factor as a function of the climate state itsef (page 3026 line 6; to which extend is it not a circular reasoning by claiming finally that the climate sensitivity –which depends on polar amplification-is climate state-dependent). What is the importance of such polar amplification factor on the climate sensitivity? There are finally few figures showing the influence of different parameters on S (only Figures 8 b and e).The importance/meaning of the linearity or non-linearity of the relationship between $\Delta T_g and \Delta R$ must be better explained. Does the fact that "if the fit follows a linear function, its value might be determined from the slope of the regression line…" (page 3031 line 8) imply that a state-dependency is absolutely requesting a non-linear relationship between $\Delta \mathrm{T}$ and ΔR as the authors seem to let it assume page 3024 line 7 , page 3031 lines 1 and 2 and page 3035 line 5.

I think that what is missing the most in the paper is a figure with S_{CO_2} , S_{LI} and $S_{CO_2,LI}$ as a function of ΔT_g showing clearly (?) the state-dependency of S which is the purpose of the paper.

More specific remarks:

1. P 3021 line 23: What the authors mean by "These details" when speaking about the astronomical forcing? Is that statement not opposed to what they say page 3025 line 20. There the authors claim that they use the long term variations of the solar radiation input. It is true that these variations can hardly be visible on figure 1a. Is it due to a lack of resolution or are these variations negligible? The second possibility is probably true as the authors use annual mean isolation which variations are indeed very small (their figure 4c, black curve). This raises a real problem because the insolation forcing is not totally negligible for calculating the temperature changes, but provided the seasonal variations are used in the response of the climate system. (see the relative contribution of insolation and cO2 in Yin and Berger, 2012)
2. Page 3022 line 26: Is the linear combination of ΔR_{LI} and ΔR_{CO_2} giving the same weight for the two? At least this is what can de deduced from the numerical values given page 3034. Would it not be better to give them a weight depending on their relative uncertainty.
3. Page 3025 line 1: what is the exact meaning of eustatic here (is it total sea level variations both mass and steric components?)
4. Page 3026 section 2.2: what is the impact of neglecting changes of temperature in the SH?
5. Page 3027 line 4: what are the two choices mentioned: are they -4.6±0.8 and -5.7±0.6 or -5.7-0.6 and -5.7+0.6?
6. All the reconstructed CO2 values are far from being homogeneous (see pages 3029 and 3030). This discussion is very welcome but what is the final impact on the climate sensitivity?
7. Section 3.2 is discussing the relationship between ΔT and ΔR looking for non-linearity. This is an excellent point, but I have difficulties with figure 7, namely to understand the fitting lines of figure 7e and 7g. In particular I do not see the inverse slope in the points of Fig 7e. If the black line is a fit I do not see how it can be obtained.

8. Page 3044 line 3: another earlier and still valid reference is Berger and Loutre (2002, Science) who were the firsts to come with such a result.