Reply: Reviewer 1

In their study Hoogakker et al. nicely compile BIOME maps using pollen records and model simulations. From the simulated biosphere changes they infer plant productivity and terrestrial carbon storage, and by using budget equations finally the ocean $\delta_{13}C$. While I very much appreciate their effort in compiling BIOME maps and vegetation distributions over the past 120 kyr, I am not convinced by their conclusions on $\delta_{13}C$ changes. The manuscript itself is well organised and written. I thus encourage publication in CP after a major revision.

General:
The compiled BIOME maps from pollen data sets are very informative and useful for the paleo data and model community. In additon there are very few simulations over the past glacial-interglcial cycles with a fully coupled model. So, it is good to see that these simulations are evaluated against paleo data.
My main critics concern the interpretation of the model results. In my opinion the approach to reconstruct ocean $\delta_{13}C$ is too simplistic and without recognition of available evidence.
Therefore, I am not convinced that terrestrial carbon stock changes have the dominant role in ocean $\delta_{13}C$ changes over the past 120 kyr. It could well be true, but there are a lot of assumptions involved and other mechanism, e.g. ocean water mass changes (Bereiter et al., 2012) that possibly could explain the observations.

This is an interesting comment. With regards to the last paragraph, we are not dismissing water mass changes at all! We are fully aware of the changes in d13C and their link with nutrient inventories / water mass change interpretations, as well as inferences made using εNd. What we mean is that changes in the vegetation were likely to have changed ocean δ13C (as has been extensively shown for the LGM). Our study basically provides the first estimates of the possible extent of these changes. They are to be seen as baseline changes; water mass changes are on top of this. In our simulations we have not modelled the effect of meltwater release and millennial scale cooling events. In order to make this clearer we have added the following line to the introduction:

(Our) line 171-172 we added: ' The effects of millennial scale climate fluctuations were not simulated.'
The reason we are comparing with records that do not include the N Atlantic is because there dramatic changes do occur in bottom water, where AABW may have shoaled to about 2 km in the North Atlantic.

As a neutral reader I would expect a more thorough calculation and uncertainty consideration of carbon stock and $\delta_{13}C$ changes. Here my suggestions:

1. The Vostok CO2 record is outdated and often lower by 10-20 ppm than newer data. Use a composit record of newer data sets (see Figure below). Also use the common timescale AICC2012 (Veres et al. 2012). 20 ppm can greatly affect NPP and thus carbon storage.

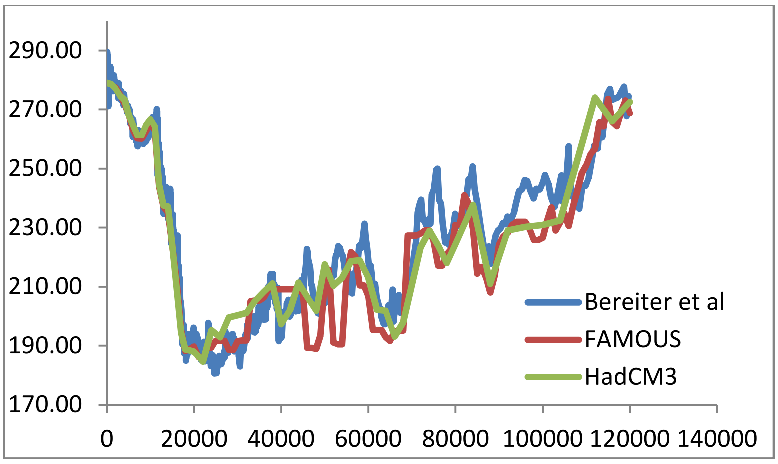

Fig 1.

Differences between prescribed CO_2 in FAMOUS and HadCM3 simulations. FAMOUS used EPICA and HadCM3 runs used Vostok ice core reconstructions with the newer Bereiter composite curve. The largest discrepancies with the composite are between 70 and 100 kyr BP, and as the reviewer says, these are up to 20ppm and usually the composite is higher than the input CO_2 to either HadCM3 or FAMOUS models and biome4. The timing of the peaks in CO_2 in the new curve is more similar to the HadCM3 inputs than FAMOUS.
It is not feasible on the timescale and resources available to rerun the climate model simulations with the later composite CO_2 as inputs. However, one could make the assumptions that 10-20ppm higher CO_2 would not alter the simulated climate significantly to modify the NPP of the terrestrial biosphere in comparison with the CO_2 fertilization effect on NPP. With this in mind we performed a sensitivity test whereby the CO_2 prescribed to biome4 run with FAMOUS climate was increased by 20ppm (for all time steps). The following curve shows the increase in terrestrial carbon storage for the whole glacial cycle from the resulting CO_2 fertilization effect.

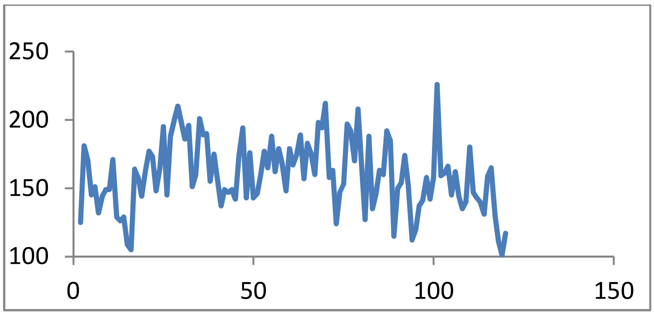

Fig. 2

The numbers are most relevant between 70 and 100 kyr, when the FAMOUS-biome4 CO_2 input curves are most different from the composite. There is a ~150GtC increase (varying between 100 and 200GtC) or around 10% (average total modelled carbon storage during this time is ~1500GtC in the original FAMOUS-driven simulations). During the other time periods covered by the model simulations the CO_2 inputs are similar to the composite and so there would be much smaller or negligible changes to NPP or carbon storage from running biome4 with the composite record.

We understand and agree with the reviewer that it would be a good idea to rerun the simulations with the composite CO2 curve forcing both the climate and vegetation, for completeness. However, as it is not possible to do this for the climate simulations (especially HadCM3) and as the sensitivity tests show that the impact would probably be maximum 10% increase in carbon storage and only really significant during MIS4 we would prefer to maintain the consistency between CO2 curves used for the climate forcing and the BIOME4 simulations presented in our original results

2. Show the uncertainty and impact of atmospheric δ13C on ocean δ13C using ice core records for the last 20 kyr (Schmitt et al., 2012) and MIS 5 (Schneider et al., 2013). Even though the majority of the δ13C signal is transferred to the ocean it could give you an indication on the direction of change.

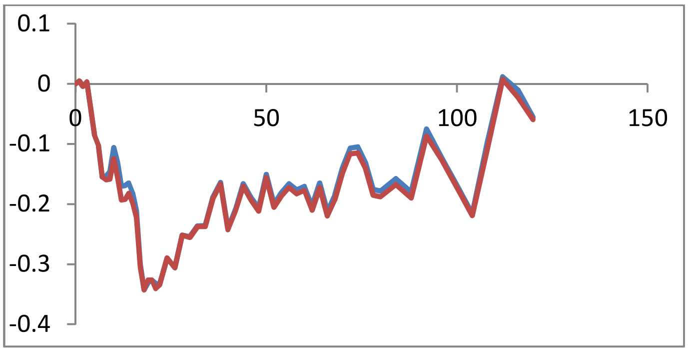

Fig. 3

Figure shows the impact of including time varying atmospheric δ13C based on ice core records on global ocean δ13C. The red line uses a constant δ13C of -6.5 per mil, and the blue line uses data from Lourantou et al, Schmitt et al, and Schneider et al with interpolation for the missing time period from 105 – 22 kyr BP. The difference is very small on the long term trends and while there could be some larger impacts on the smaller scale sub-glacial cycle patterning between ~90 and 40 kyr BP if there are also similar scale variations in δ13C, however there is no data for this time period.

3. One shortcoming of the δ13C analysis is the missing of peatlands and permafrost carbon stocks in the model, as mentioned in the beginning of the paper. They act excactly on these long time scales that matter for the terrestrial carbon change over the observed period, and also have an opposite effect on carbon storage compared to e.g. forest ecosystems. According to Ciais et al., 2012, as you write, the inert carbon stock was larger by ~ 700 PgC during the LGM compared to PI. Assuming a linear increase in permafrost carbon between the previous interglacial and the LGM: How would the increasing carbon storage in permafrost areas affect the δ13C budget? Please discuss and see below for more specific comments.

We have added to the discussion and included some additional exploratory plots on the impact of an increasing inert carbon pool during the glacial. See below for details.

4. The equations used to estimate carbon storage from NPP has underlying assumptions that are questionable. It is assumed that soil carbon is in steady state

at each time step in the past, which is wrong for ecosystems with a turnover time larger than the model time step (1000 years for FAMOUS I guess), i.e. for wetland ecosystems or again permafrost areas. Further, turnover times are estimated from present day soil carbon storage. Again it is assumed that for current conditions soils are in steady state, in a time of rising temperature, CO_2, and nutrient input. A discussion of the implications is needed here.

Although it is true that the heterogeneity of soil organic matter means that some soil carbon varies on millennial timescales, the soil response to changes in climate tends to be dominated by the more labile carbon pools. Effective residence times for soil carbon have been estimated as being decades near the equator and ~250 years in the high northern latitudes (Carvalhais et al., 2014). Moreover, our models do not include components such as wetlands or permafrost with the very long timescales suggested by the reviewer. The steady-state soil carbon assumption used in our study thus neglects a lag in total biosphere carbon response, although on the millennial timescales analysed here it is unlikely to introduce major inaccuracy compared to other assumptions used in the study.

The use of current, non-equilibrium soil conditions to estimate the turnover times for each biome is perhaps more of an issue. The difference in turnover constants (now listed in table 2) derived for the FAMOUS and HadCM3-forced runs partly reflects differences in assumptions for what level of simulated NPP is appropriate to compare with these modern carbon stocks. The uncertainty in turnover constants can introduce further uncertainty to the carbon storage calculations, on a similar scale (~10%) to those seen above related to the details of the CO_2 forcing curve used for the glacial cycle. On further investigation the rather large drop in pre-industrial to LGM terrestrial carbon storage seen in the FAMOUS-forced biome simulations seems to lie at the upper end of this uncertainty range. Some discussion of these issues is now included alongside our summary of the method of Wang et al. (2011).

Given the richness of BIOME data and the complexity of the climate models used in this study I think the analysis of $\delta_{13}C$ falls short. I don't think new climate simulations are needed, but additional simulations with BIOME4 for sensitivity tests and a thorough uncertainty estimate (1 sigma band for land and ocean $\delta_{13}C$) would considerably improve the statement of the paper.

We've added an uncertainty estimate in the $\delta13C$ ocean bit – where the largest uncertainty arises from the inert terrestrial carbon pool.

Specific:
p. 1039, l. 2: The Vostok CO_2 record is at least 16 years old and outdated by records with higher temporal resolution and measured by more accurate techniques (e.g. direct measurements by ice sublimation). Please replace it with data from newer ice cores (see Figure and References below).

See reply to comment above (estimated differences from using different CO_2, inclusion of comment that we are aware of newer CO_2 data available with higher temporal resolution). However, we also note that for the purpose of our paper, which primarily addresses smooth glacial-interglacial changes, significantly higher temporal resolution is not really needed.

p. 1943, l.2: should this be Fig. 1 or Table 1? Could not find reference to Fig. 1 in text.

P 1037, l 16. The reference to Table 2 is correct, as this also provides details of site elevations. Figure 1 shows the locations of the various pollen sites, with site details provided in Table 2. We have added a reference to Figure 1 in line 158 '(for locations see Figure 1)'.

p. 1049, l.11: Are HadCM3 surface temperatures absolutely 1°C colder at present or is the LGM-present anomaly 1°C colder? In the first case this should not matter, when you use anomalies for BIOME4. Please clarify.

The HadCM3 paleoclimate anomalies are, in general, a degree or so colder than the FAMOUS anomalies, so this is a significant difference in forcing.

p.1052, l. 19: The interpretation of the "sahara greening" in the model is at its limit, when only a hand full of grid cells swap color at this coarse resolution. In general the description for comparing model grid cell changes could be shortened and less speculative.

The Sahara greening reference has been scaled back.

p. 1060, l. 23: Is the CO_2 fertilization effect or the CO_2 climate effect more dominant for NPP? Could this be tested with a BIOME4 simulation with constant CO_2?

We have performed sensitivity tests where BIOME4 is driven by glacial-interglacial climate changes both with and without changes to CO_2 to influence CO_2 fertilization. The outcome is that CO_2 fertilization is the predominant driver of lower glacial NPP (around 85% of the impact at the LGM). We have added a sentence to reflect this in this paragraph.

p. 1061, l. 7: These numbers directly depend on prescribed CO_2. Using an updated CO_2 record (see Figure below) should result in e.g. smaller differences between the Eemian and the Holocene.

We have estimated the impact that using the composite CO_2 curve would have on the FAMOUS-forced BIOME4 simulations. Using the sensitivity experiment (described earlier) where CO_2 input to BIOME4 was increased by 20ppm for each timeslice we approximate the sensitivity to CO_2 of NPP and then use this to estimate NPP had the composite curve been used rather than the original CO_2 forcing in FAMOUS. This captures the impact of CO_2 fertilization, which is the primary driver (but can be considered conservative given that we do not have the CO_2-forced climate changes). The comparison can be seen below:

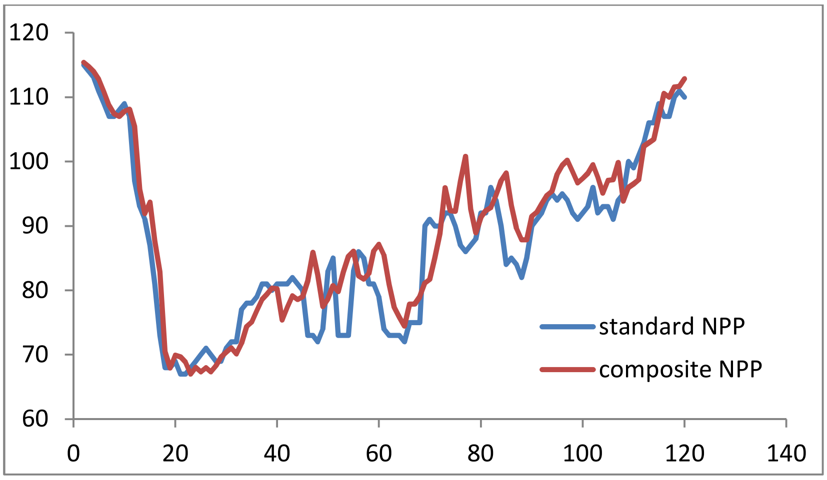

Fig. 4

The details described in the paragraph highlighted by the reviewer do change somewhat depending on CO2, as he/she says. The Eemian (at ~120kyr) is more similar to the Holocene, although still lower. The timing and magnitude of the large drops in productivity are modified slightly. We have altered the paragraph in the text to draw attention to the idea that the NPP will be quite sensitive to the CO2 inputs. However, we would prefer to leave the data as it is rather than use the composite curve so as to maintain climate and biome CO2-forcing consistency in the results we present.

p. 1092, l. 24: Which soil carbon data has been used for the calibration of the turnover times?

Modern soil and vegetation carbon inventory data by megabiome was taken from Table 3.2 of the IPCC TAR WG1 (Prentice et al. 2001). Prentice et al. cite Mooney, Roy and Saugier (2001) for vegetation carbon, and the IGBP-DIS soil carbon layer (Carter and Scholes, 2000) overlaid with De Fries et al. (1999) current vegetation map for average ecosystem soil carbon. As noted above, the turnover times and resultant terrestrial carbon storage are sensitive to the choice of carbon datasets – this has been noted in the paper.

Carter, A.J. and R.J. Scholes, 2000: Spatial Global Database of Soil Properties. IGBP Global Soil Data Task CD-ROM. International Geosphere-Biosphere Programme (IGBP) Data Information Systems. Toulouse, France.

De Fries, R.S., C.B. Field, I. Fung, G.J. Collatz, and L. Bounoua, 1999: Combining satellite data and biogeochemical models to estimate global effects of human-induced land cover change on carbon emissions and primary productivity. Global Biogeochemical Cycles, 13, 803-815

Mooney, H., J. Roy and B. Saugier (eds.) 2001. Terrestrial Global Productivity: Past, Present and Future, Academic Press, San Diego

p. 1063, l22: If I would argue that NPP is dominated by CO2 fertilization, as the curve in Fig. 5 visually correlates with the CO2 record, would you still get a precessional

cycle in terrestrial carbon storage with constant CO2 in BIOME4? Using the updated CO2 record may change the periodicity. Please reassess.
We have used the sensitivity tests with higher CO_2 to produce a rough curve for terrestrial carbon storage had the Bereiter composite curve been used to drive BIOME4:

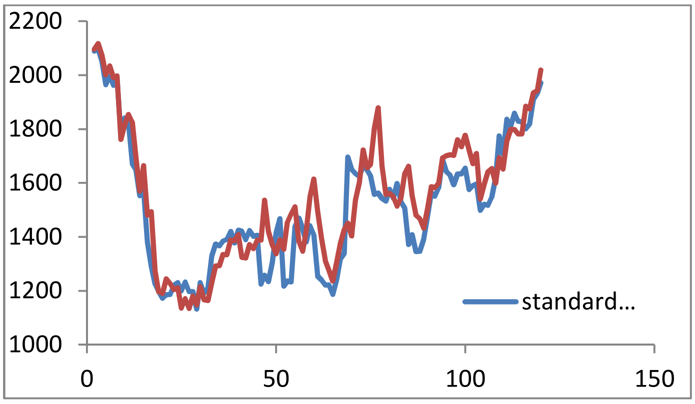

Fig. 5

Carbon storage has a greater proportional impact from climate than NPP – particularly the soil temperature dependence by the method of calculation used in this paper. The timing of the peaks and troughs in carbon storage are a balance between productivity and increased respiration (carbon release) with warmer temperatures. The timing and periodicity of these is roughly similar for BIOME4 in both scenarios, as in the figure. However, as we do not have the resources to rerun the climate model with the new composite curve, and because climate is more important in this case, we would prefer again to keep to the current CO_2 inputs for the climate and vegetation models.

p. 1064, l. 19: Replace 'decrease' by 'difference' as the former has a time direction associated. Time runs from LGM to PI.

We have changed this in the text.

p. 1065, l. 6: Please use updated CO2 (see Figure below) and δ_{13}C (Schmitt et al., 2012; Schneider et al., 2013) records for the atmospheric part of the budget.

We have used the updated and interpolated d13C records in the calculation, which makes very little difference to the calculation for the ocean budget (see figure 3). We have then estimated the changes to d13C for the global ocean from using the Bereiter et al composite curve by using the sensitivity experiment with FAMOUS-forced BIOME4 driven with higher CO2 to both as the atmospheric inventory input as well as the terrestrial carbon storage input. The result is given in Figure 6 below. The difference, as with terrestrial NPP and carbon storage, is largest between 70 and 100 kyr BP, where the Bereiter input produces a slightly isotopically heavier ocean at some time points (up to 0.05 per mil increase in most cases but larger than this at a couple of time slices only). However, it does not change the glacial-interglacial maximum change or the general characteristics of the glacial decline. The figure is produced for the reviewer but as before we would like to have the CO2 inputs to the climate, vegetation, and d13C budget calculations be consistent, and so keep them as they were in the paper.

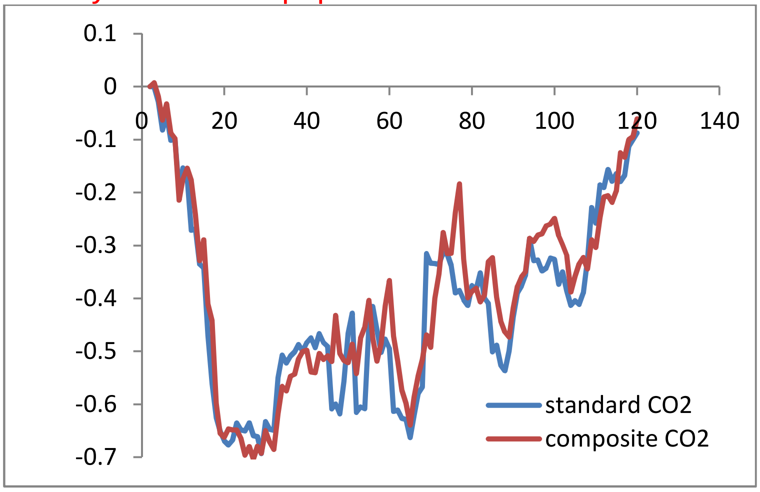

Fig. 6

p. 1066, l. 2: You could mention that biomes do not include permafrost (normally C3 plants) and peatlands (C3 plants and sphagnum moss with $\delta_{13}C$ = ~ -30 per mill). Having said that, please also clarify that the variability of terrestrial $\delta_{13}C$ in Fig. 6a is of secondary importance for ocean $\delta_{13}C$. What matters are terrestrial carbon storage changes.

We have added a statement to this paragraph to reflect that changes in δ13C in the terrestrial biosphere are not as crucial as changes to the size of the terrestrial carbon pool.

p. 1066, l. 14: This is correct, but only because both models lack inert carbon pools. If you include them like in Ciais et al., 2012, then the FAMOUS model would agree better (see paragraph 4.3. in your own words).

We have estimated the effect of including an inert terrestrial carbon pool (permafrost, peatlands) of the size inferred by Ciais et al (2012). We used the figures of 1600GtC at pre-industrial and 2300GtC at LGM and then interpolated between. Given the lack of estimates for the Eemian, we used the same value as the pre-industrial at 120 ka BP and then interpolated linearly to the LGM value. Assuming a lighter δ13C signature (of -27 per mil) for the inert pool the impact on the FAMOUS global ocean δ13C is displayed in Fig. 7:

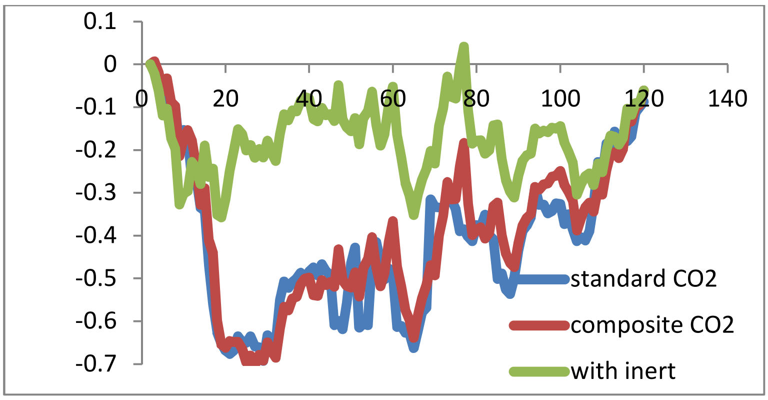

Fig. 7
For the FAMOUS-driven BIOME4 simulations, the inclusion of an inert terrestrial carbon pool improves the comparison of the ocean d13C with the data compilation by Oliver et al., and would make it match the data better than HadCM3-driven BIOME4 simulations, as the reviewer suggested. We have included discussion about this neglected carbon reservoir in section 4.4.

p. 1066, l. 24: Please also cite Bereiter et al., 2012.

OK

p. 1067, l. 5ff: This statement is too strong, I'm not convinced. I believe that the trend in modelled ocean δ_{13}C from MIS 5 to MIS 2 may be robust, but not the variability in between, e.g. the variability from HadCM3 climate is rather small.

We have toned this down and changed 'are' to 'may be'.

p. 1068, l. 25: Again, I'm not convinced by the presented material that the role of land δ_{13}C is “dominant” for ocean δ_{13}C. See General comments.

We toned this down and changed 'dominant' to 'important'.

p. 1068, l. 13: This is very valuable and a good reason this paper deserves publication after a revision.

References: Ciais et al., 2011 should be Ciais et al., 2012 in the entire text.

We have changed this throughout the text.

Figure 2: What does (a) and (b) signify? Is there any difference between plots on top rightand left? Please enlarge this figure panel in two figures for better visibility.

Indeed, we requested that these panels be reproduced on separate pages to improve visibility in our initial submission to CPD - we will again ask the editor to make the panels as large as possible. a) simply denoted the panel of the more recent set of timeslices and b) to the earlier ones, but the distinction is not in fact used in the text and could be removed. The figure caption has been edited for clarity.

'Reconstructed biomes (defined through highest affinity score) superimposed on simulated biomes using FAMOUS (B4F, left) and HadCM3 (B4H, right) climates for selected marine isotope stages (denoted in ka BP). '

References:
Ahn J., Brook E. J. (2008) Atmospheric CO2 and climate on millennial time scales during the last glacial period. *Science* 322:83–85.
Ahn, J., et al. (2012), Abrupt change in atmospheric CO2 during the last ice age, Geophys. Res. Lett., 39, L18711, doi:10.1029/2012GL053018.
Bereiter, B., et al., (2012) Mode change of millennial CO2 variability during the last glacial cycle associated with a bipolar marine carbon seesaw, *Proceedings of the National Academy of Sciences of The United States of America, 109/25*, 9755-9760

Ciais, P., et al. (2012) Large inert carbon pool in the terrestrial biosphere during the Last Glacial Maximum, *Nat. Geosci.*, 5, 74–79, 2012.
MacFarling Meure, C., et al. (2006), Law Dome CO2, CH4 and N2O ice core records extended to 2000 years BP, *Geophys. Res. Lett.*, 33, L14810, doi:10.1029/2006GL026152.
Monnin E., et al. (2001) Atmospheric CO2 concentrations over the last glacial termination. *Science* 291:112–114.
Schmitt, J., et al. (2012) Carbon isotope constraints on the deglacial CO2 rise from ice cores, *Science*, 336, 711–714.
Schneider, R., et al. (2013) A reconstruction of atmospheric carbon dioxide and its stable carbon isotopic composition from the penultimate glacial maximum to the last glacialinception, *Climate of the Past, 9*, 2507-2523.
Veres, D., et al. (2013) The Antarctic ice core chronology (AICC2012): an optimized multiparameter and multi-site dating approach for the last 120 thousand years, Clim. Past, 9, 1733-1748, doi:10.5194/cp-9-1733-2013

Figure: Composit record following Bereiter et al., 2012.

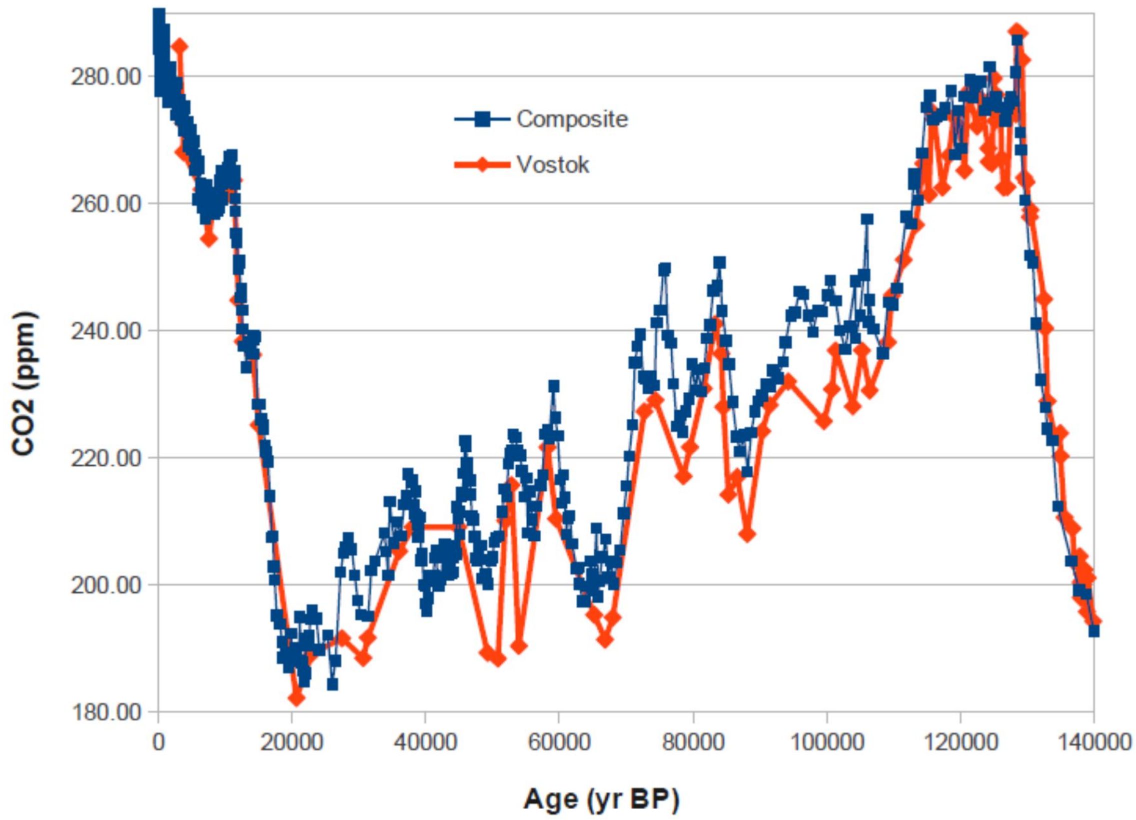

Figure: Composite record following Bereiter et al., 2012.
Composite CO2 record on AICC2012 (Veres et al., 2012)

-46 - 10 yr BP: Law Dome (MacFarling Meure et al., 2006)
0 - 1 kyr BP: WAIS (Ahn et al., 2012)
1 - 2 kyr BP: Law Dome (MacFarling Meure et al., 2006)
0 - 22 kyr BP: Dome C (Monnin et al. 2001)
22 - 24 kyr BP: Dome C Sublimation (Schmitt et al., 2011)
24 - 38 kyr BP: Byrd (Ahn et al., 2008)
38 - 60 kyr BP: TALDICE (Bereiter et al., 2012)
60 - 115 kyr BP: EDML (Bereiter et al.,2012)
105 - 155 kyr BP: Dome C Sublimation (Schneider et al., 2013