Response to referee #1

Specific comments:

1. 2 Data and methods 2.1 Model description. For the ocean carbon cycle component, how are the production and dissolution of $CaCO_3$ (the hard tissue pump) simulated? Please elaborate.
   We provide more details on this in the revised version:

   "Biogenic calcification is implemented as being proportional to a fraction of small phytoplankton production, which is temperature-dependent. An exponential curve is prescribed to simulate the dissolution of sinking $CaCO_3$ (Moore et al., 2004}. There exists no dependence of calcification-dissolution rates on saturation state."

2. 3 General climate and carbon cycle evolution 3.1 Temperature Page 363: "However, the uncertainties in the early period of the reconstructions prohibits to robustly answer the question whether the models are too global in their response to external forcing." "too global in their response to external forcing". What does it mean exactly? Please clarify.
   We expanded with the following paragraph to hopefully clarify this:

   "A lingering question of climate modeling in general is whether models are too global in their response to external forcing. That is, they might show too little regional variability that is independent from the global mean response during a forced period. However, the uncertainties in the early period of the reconstructions prohibits to robustly answer this question."

3. 3.3 Carbon cycle Page 366 "The prognostic atmospheric CO_2 increases to 1156 ppm by 2100 CE. This would imply a forcing of 7.6 Wm^{-2} from CO_2 relative to 850 CE" Please clarify how the CO_2 radiative forcing was calculated.
   In the same way as the radiative forcing in Fig. 1c, namely according to IPCC (2001). We clarify this in the revised manuscript.

4. "Together with the underestimated oceanic uptake this leads to the roughly 20% larger airborne fraction in CESM as compared to the RCP8.5." What does airborne fraction of RCP8.5 refer to and how was it calculated?
   We are referring to the CO_2 concentration that is prescribed in the radiative code according to the RCP8.5 scenario (so this is not calculated, but prescribed). To clarify we expanded to:

   "Together with the underestimated oceanic uptake this leads to the roughly 20% larger airborne fraction in CESM as compared to what is actually prescribed as atmospheric concentration in the radiative code according to the RCP 8.5"

5. 5 Volcanic forcing Page 372 "Although carbon loss due to fire increases" This should be elaborated a bit more. How is the effect of fire accounted for between the period of 850–2100CE?
   Fire activity in the model is prognostic and depends on drought conditions (soil moisture mainly) and the availability of material to burn. We changed the paragraph in the volcanic section to read:

   "Due to the dry conditions and availability of dead biomass there is increased fire activity, leading to increased carbon loss due from land. However, the fire cannot get rid of the large amount of dead biomass immediately..."

   Further, we give some details on the fire module in the Methods section:

“Further, it includes a prognostic fire module, which is governed by near-surface soil moisture conditions and fuel availability.”

Response to referee #2

General comments

Referee: This paper presents a new CESM past to future model simulation and is generally within the remit of ESD. I found this paper difficult to follow because it lacked a clear direction and purpose. The abstract suggests that the originality of this work is the continuity of the simulation from 1000 years before present to 100 years into the future, and using a different solar irradiance reconstruction. However, it is not obvious how the aims of the paper: to detect large-scale forced variability; forcing vs. structural uncertainty; and provide context to future projections (p.355 l.3-7), are novel or can be addressed with this simulation. Providing context, in particular, is a rather vague aim.

Reply: (1) The main novelty is the interactive carbon cycle over the last millennium. We rephrased the abstract to stress this. (2) The referee comment on “structural uncertainty” is addressed in our answer to the next comment. (3) Some of the aims in the introduction have been rephrased to hopefully appear less vague.

Referee: The paper goes on to compare this new simulation to a mix of previously published data and model simulations with different components, different forcings applied, and different resolutions. Because of these differences, I found the comments about structural vs. forcing uncertainty rather less credible. The paper ‘fails’ to find any large-scale variability, and it was unclear what the contribution on past context was. So, the claimed originality doesn’t have much in common with the aims, the aims only loosely tie with the results presented, and the conclusion is that it is a null result.

Reply: We agree that there are multiple components that contribute to what we summarize as “structural model uncertainty”. Some of them may rather be referred to as “differences in implementing a given forcing”, some as “differences in resolution”, some as “differences in climate sensitivity”, some as “differences in magnitude of internal variability” or others. The point here is to highlight the implications these differences can have in presence of supposedly identical forcing across models. To provide an in-depth discussion and dissection of the underlying causes of all the model differences is beyond the scope of this paper. We revised parts of the introduction and conclusions in light of this comment.

Referee: I think perhaps that the basic issue with this paper is that it tries to cover much. There are references to millennial timescale, pre-industrial, future, comparison between CCSM4 and CESM, comparison between CESM and MPI-ESM, comparison with other PMIP model simulations, comparison with other CMIP model simulations, comparison with data, orbital forcing, climate sensitivity, carbon cycle feedbacks, and carbon cycle response to volcanic forcing. Consequently each section of results is quite superficial. The paper is quite long and not clear in its overarching aim or aspect of novelty. This would be a more useful paper if the scope were reduced and there was a definite focus on what the scientific contribution of this work is.

Reply: We agree that there are a lot of different topics addressed in this paper, which reflects the overview character of the paper as well as the contributions from many different co-authors. We think, however, that each section provides a new result, even if one can certainly argue that each of those results could be investigated in more depth in a subsequent study. We revised the abstract and discussion to hopefully reflect this better.

Specific comments:

Referee: My specific comments do not address sections 3.2 or 5. Given the large range of scope of the sections in the paper, I do feel well qualified to assess these.

Referee: Is the control simulation properly spun up? A supplementary figure would help in this case. And the soil carbon storage units need checking on P.357 L.21

Reply: As stated in the Experimental Setup section, the model is not in equilibrium and in response to another referee comment we now give more information on model drift in that section. The units have been revised as well.

Referee: The methods and experimental set-up desperately needs a table with a clear table of the features, and forcings of the models that are mainly used in the paper. A simple explanation of why these model simulations were chosen, why the authors consider them comparable despite their noted differences would help readers.

Reply: We implemented a table and a more detailed reasoning for the model choices.

Referee: The first part of the results seems to be a competent description of this model simulation over this time period. There are a few null results, it's more or less in line with the results from other models, it doesn't always agree with the data (but then neither do the other models). I don't know what the objective or hypothesis for this model simulation was, and I'm not totally convinced that the authors know either.

Reply: As stated in the introduction, we are interested in the carbon cycle sensitivity and climate variability in presence of altered forcings as compared to traditional PMIP3 protocols. The fact that some of the altered forcing do not result in a discernible effect on the simulated climate might be the confirmation of a null hypothesis, yet it is still a valuable result. Further, having the orbital forcing fixed allowed to study its influence when comparing with other simulations.

Referee: Section 6. I'm a bit dubious about the methods used here. "Mimicking to some extent" other methods is rather vague, more clarity would be helpful here. The methods section doesn't say whether dynamic vegetation is turned on in the simulation, but presuming that it is, I'd be surprised if the low pass filter didn't obscure the reaction of the C3 grasses and other quick growing vegetation types to temperature increases.

Reply: We rephrased the whole section to reflect this and comments by other referees.

Given the experimental setup, we cannot apply the identical methods as in Frank et al. or Jungclaus et al. We follow them as much as possible and describe in detail what we did.

Following from that, it is worth stressing that the intention of the low-pass filter is exactly to filter out interannual variability due to volcanoes, ENSO, and other short-lived temperature variations, since these variations would not allow for a robust estimate of the climate-carbon cycle sensitivity (which we state in that paragraph). In fact, we show that it is difficult enough with the low-pass filtered data to find robust results. We would apply a very different analysis if we were interested in the interannual variability of the carbon cycle and the land vegetation. Due to the prescribed land use changes, there is no dynamic vegetation active in the model. We clarify this in the Methods section now. We encourage the referee to consider section 5, despite not being an expert on volcanic forcing. This section is looking in some detail on the interannual variability (i.e., without filtering) of the carbon cycle in response to volcanoes.

Referee: Similarly, selecting only the northern hemisphere biases the results because of the smaller amount of ocean. The oceans are obviously a huge part of the carbon cycle, particularly over longer time periods and it seems quite possible to me that the global sensitivity could be different to that of the northern hemisphere. If considering the global CO2, surely you need to consider global temperature, else you could be attributing a CO2 change originating in a S hemisphere ocean circulation change to a N hemisphere temperature change.

Reply: This a good point. With this approach, we were again following Frank et al. and Jungclaus et al., who use global CO2 and NH temperature. We double-checked the results using global CO2 and global temperature, which results in median values of 1.7 ppm/K for the transient simulation and 2.3 ppm/K for the control simulation. The values for using global CO2 and NH temperature were 1.3 and 2.3. This result is not surprising, as including the SH smoothes the temperature time series, especially after volcanoes or solar minima (due to the ocean damping). Therefore, temperature variations are reduced while the CO2 time series remains the same, resulting in a higher sensitivity estimate. In contrast, for the control simulation there are no volcanoes or solar variations, so that the inclusion of the SH has little effect on the climate-carbon cycle sensitivity estimate from the control simulation. We include the following paragraph in the revised manuscript:

"Note, that we use NH SAT in order to be comparable with existing studies. Using global instead of NH SAT can influence the estimate of gamma, especially for the forced simulation: including the wast ocean area of the SH tends to dampen temperature variability induced by volcanoes and TSI variations. With temperature variability dampened, gamma increases to 1.7 ppm/K (1.4-2.1). For the CTRL, on the other hand, which does not see volcanoes or TSI variations, using global SAT has no discernible effect (2.3 ppm/K)."

Referee: p.375 l.1-20 To say that the c cycle sensitivity is "comparably low" is just not the case. The median value is outside of the reconstructed range, so "very low" would be a better way of describing the sensitivity. I find the sentence about Arora et al rather misleading, since the model is "in agreement" with other CMIP5 models, but the positioning of the sentence makes it seem as though it is in agreement with data.

Reply: We changed this to "very low" as suggested by the referee. Further, we changed the sentence discussing other model studies:

"This low sensitivity of CESM was found in other model studies as well, e.g., Arora et al. (2013)."

Similarly, in the conclusions we now state:

"Generally, the sensitivity of the carbon cycle to temperature variations in CESM is very low compared to observations..."

Referee: The discussion in general doesn't add to the paper as it reiterates the findings. It also gives general advice about how paleoclimate modelling can be better conducted. This advice is not (so far as I can see) novel, and the last paragraph of the paper is particularly galling, since it calls for ensembles with properly separated forcings, which is what the rest of the paleoclimate community usually already do, and what probably should have been done to address the aims of this paper.

Reply: While we agree that an ensemble of simulations and/or a number of single forcing simulations would have been helpful, this was far beyond the resources of our institution. Upon the start of this study there were simulations from 9 different models available within the PMIP3 framework, but only one of them had single forcing simulations and an actual ensemble (i.e., more than one simulation). So, we disagree that this approach is already the "standard" in paleoclimate modeling. For those reasons it is in our view worth stressing the need for such ensembles. We expanded that part of the discussion to highlight the problem of optimizing the usage of computing resources.

Referee: The figures are nicely presented.

Response to referee #3

The manuscript 'Climate and carbon cycle dynamics in a CESM simulation from 850-2100 CE' by Lehner et al. describes the evolution of climate and the carbon cycle from the last millennium to the end of the current century as simulated by CESM model. The authors investigate the response of the climate and the global carbon cycle to the role of orbital forcing and volcanic eruption.   They take advantage of this modelling framework to determine climate-carbon cycle sensitivity over several periods. The authors employ a quantitative methodology comparing the response of CESM model to previous simulations of CCSM and MPI-ESM and to available reconstruction and observational data. This manuscript is well written and the analyses are sounds. As such, this manuscript is a good documentation of the climate and carbon cycle evolution during the last millennium as simulated by CESM. Therefore, I recommend its publication after the following minor issues are addressed.

General comments:

Referee: (1) The paper is too long and might be shorten if results & discussion are re-arrange.

Reply: Some sections have been condensed, some needed to be expanded in order to satisfy referee comments.

Referee: (2) Several mechanisms rely on the role of the ocean. However, few analyses are provided in terms of ocean physics and ocean marine biogeochemistry.

Reply: In response to a number of referee comments we provide more details on the some of the processes (see specific comments).

Referee: (3) It is unclear if the ocean component of the CESM model has benefited from a proper spin-up.

Reply: We have mentioned in the original version of the manuscript that the ocean is likely not in equilibrium. We have now expanded the discussion on model drift and provide more diagnostics on this topic (see specific comments).

Specific comments:

Referee: P352 L14 what do you mean by "potentially" ?

Reply: What we intend to say is that only because we cannot detect a forced signal that does not mean there might not be one. The sample size might be too small to detect it. We changed the text to "might mask" instead of "potentially masks".

Referee: P352 L16 please cite the adequate references here.

Reply: Including references in the abstract is to our knowledge not common practice. However, the adequate references (Kaufman et al., Esper et al.) are given in the introduction and respective section.

Referee: P352 L17-18 in regards of the results/discussion section, few words are needed to indicate that the climate-carbon sensitivity in CESM is lower than that estimated by Frank et al., 2010.

Reply: We extended the sentence to read "The climate-carbon cycle sensitivity in CESM during the last millennium is estimated to be about 1.3 ppm/°C, lower than recent proxy-based estimates."

Referee: P353 L24 usually the envelope refers to 1xsd (66% confidence interval) while that used in the manuscript is 2xsd (95% CI).

Reply: It is unclear to us why the referee notes this for this specific location in the manuscript (maybe there was a mix-up in page and line reference?). We usually apply the same uncertainty estimate as the reconstruction we compare to.

Referee: P354 L21 please add (Tjiputra and Otterå, 2011) to the reference list

Reply: Done.

Referee: P355 L9 please remove 'fully'. Your experimental design implies that the carbon cycle is coupled only with biogeochemical components not the climate. Or, maybe add few lines on how biogeochemical responses of the interactive carbon cycle may impact the climate (e.g., evapotranspiration in response to rising xCO2 in CLM4). I seems this setup might bias the determination of climate-carbon sensitivity. Maybe add few words on this in the discussion.

Reply: We agree with the referee, removed "fully" and included the following sentence into the discussion of these results:

"Further uncertainty arises from the experimental setup used here that does not incorporate feedbacks from the carbon cycle to the climate, such as changed surface energy and water fluxes due to local changes in atmospheric CO2."

Referee: P357 2.2 experimental setup I think that description of the ocean biogeochemical inititial condition is omitted here. Please provide a description. What are the drift in ocean transport metrics like the AMOC, ACC, AABW flow in CESM?

Reply: The linear trends over the whole control simulation for those three quantities are:

AMOC: -0.22 Sv 100 yr^{-1}

ACC: 0.70 Sv 100 yr^{-1}

AABW: 0.01 Sv 100 yr^{-1}

We mention them in the Experimental Setup section along with numbers for the DIC drift (-0.01% 100 yr^{-1}).

Referee: P358 L18 you mean that there is no background volcanoes over the future scenario period ? How does this impact the simulated natural variability compared to previous period (in terms of detrended signal) ?

Reply: It will reduce natural variability. This is why we did not analyze the 21th century in terms of synchronization (Figs. 5, 7, 8). We extended the already existing reasoning for this to read:

"Thereby, we focus on the preindustrial period, as the twentieth and twenty-first century are dominated by anthropogenic trends, which are non-trivial to remove for a proper correlation analysis. Also, the omission of volcanic forcing during the twenty-first century would likely bias the natural variability estimate low."

Referee: P360 L3 If I'm right, the experimental design in IPSL model is not similar to yours since impacts of volcanoes is computed offline and added to the variation of the solar constant (see Dufresne et al., (2013; Swingedouw et al., (2013)).

Reply: Yes, indeed. But the forcing is based on the same reconstruction. The point we want to make is that there are such large differences in how models implement the same reconstructed volcanoes (what summarize as structural model differences) that it becomes difficult to separate uncertainty due to forcing from uncertainty due to forcing implementation. We included the following sentence to clarify this:

"Note, however, that the technical implementation of those forcings into the two models are different, giving rise to structural model uncertainty even in presence of identical forcing timeseries."

Referee: P364 L15 please provide quantitative information here. A table might help.

Reply: We now provide a table with the cumulative carbon fluxes over different time segments. Upon doing that we discovered a small bug in our previous summing over the years 1750-2011, which is why those numbers (which were already in the text before) changed slightly, although not discernable (see the new Table 3).

Referee: P365 L28 please cite (Schwinger et al., 2014)

Reply: Done.

Referee: P366 L5 please cite (Wunsch and Heimbach, 2007; 2008). Quantitative information on the Southern Ocean ventilation might help (AABW flows, winter mixed volume etc...)

Reply: We now cite Wunsch's work and further make reference to Long et al. (2013), who provide extensive documentation of CESM's ventilation and mixed layer depth bias.

Referee: P368 L17 Weaker correlations in the high latitudes domains were expected since you apply a 5-year smoothing filter. You could eventually assess the correlation in high-latitude domains with filter bow larger than 5 years.

Reply: We tried filters of 10 years (moving window length = 200 years) and 20 years (moving window length = 400 years), the conclusions are not affected. With larger filters there start to be too few independent values to reach significant correlations anymore. Or in other words, the length of the moving window reaches the length of the entire simulation.

Referee: P369 L5 please cite Geoffroy et al. (2015) which show how land-sea ratio warming differs between CCSM4 and MPI-ESM.

Reply: Done.

Referee: P369 L17 To my point of view the penetration depth of the signal must refers to heat fluxes not solely to changes in ocean temperature. Please check whether the results are consistent using the ratio between OH [W m-3] and Hflx [W m-2].

Reply: We are not entirely sure how the proposed analysis resolves this issue. Maybe we also do not fully understand the reviewer's comment. Apologies if this is the case. In a global mean perspective, as we have it in Figure 7, the surface heat fluxes will predominantly determine the decadal temperature anomalies and their penetration depth after volcanoes. Circulation changes will play a minor role at best. Of course, regionally this can be a different story. However, we do not see significant changes/phasing in, for example, the AMOC. However, we constructed a composite of the strongest three volcanoes in CESM and show the heat flux at different depth in the ocean in Figure 1 (attached to review response). It confirms that there are significant changes in heat flux across a surface of 50 m, 100 m, 150 m, and 200 m. In the tropical Pacific there is increased heat loss (positive anomalies) in the upper layers, while there is reduced heat uptake (negative anomalies) in the Southern Ocean and North Atlantic. Both of these processes will act to cool the global ocean down to depth of at least 200 m.

Referee: P369 L 26 you may also refer to Swingedouw et al., (2015)

Reply: Done.

Referee: P370 section 4.2 Further details are needed here. First the use of DIC anomaly with respect to 850-1849 might be clearly state in the text. Then, It is unclear to me whether the evolution of the distribution of the DIC anomalies in function of time is an artifact of the anomaly calculation or an effective difference of behavior between the two models.

Reply: We now emphasize the reference period at the beginning of section 4. The apparent differences between panels a and b of Fig. 8 are certainly influenced by different low-frequency trends in the two models (see also response to next comment). This is exactly why the running window correlation in panel c is useful for highlighting periods of coherent model behavior, as it is largely independent from differences in mean or millennial-scale drift in the two models.

Referee: If control simulation is available over such period, please assess if the patterns shown on Figure 8 also emerge after correcting the drift in DIC. Since most of the difference are due to various behavior in Southern Ocean mixed-layer depth, it might be interesting to illustrate these latter with an additional Figure. If models are identical, you could eventually refer to (Resplandy et al., 2015) which provide a quantitative comparison of several CMIP5 model including CESM-BGC and MPI-ESM over the preindustrial control simulation.

Reply: Thanks for this good comment and reference. It seems from the analysis in Resplandy et al. that CESM and MPI have a comparably weak variability in the Southern Ocean CO2 fluxes. The MPI simulation used here bases on a very long control and should be largely without drift, according to Jungclaus et al. (2010). Unfortunately, the CESM control simulation is not long enough to calculate the drift for the whole transient simulation. But we redid the analysis for Figure 8 on the part of the CESM transient simulation for which we have a corresponding control simulation. The drift in deep ocean DIC in CESM is removed, however the correlation pattern between the two models in the upper ocean remains largely the same (see Figure 2; attached to review response) and so do our conclusions. However, we make note of these new results in the revised text and refer to Resplandy et al. in the discussion.

"There appear to exist spurious trends in CESM, likely related to model drift. We repeated the analysis, but with the CESM output detrended in each grid cell by subtracting the CTRL over the corresponding period 850-1372 CE. Due to the shortness of CTRL, we cannot apply this to the whole simulation. However, these tests showed that the correlation between the two simulation is largely insensitive to the drift in CESM."

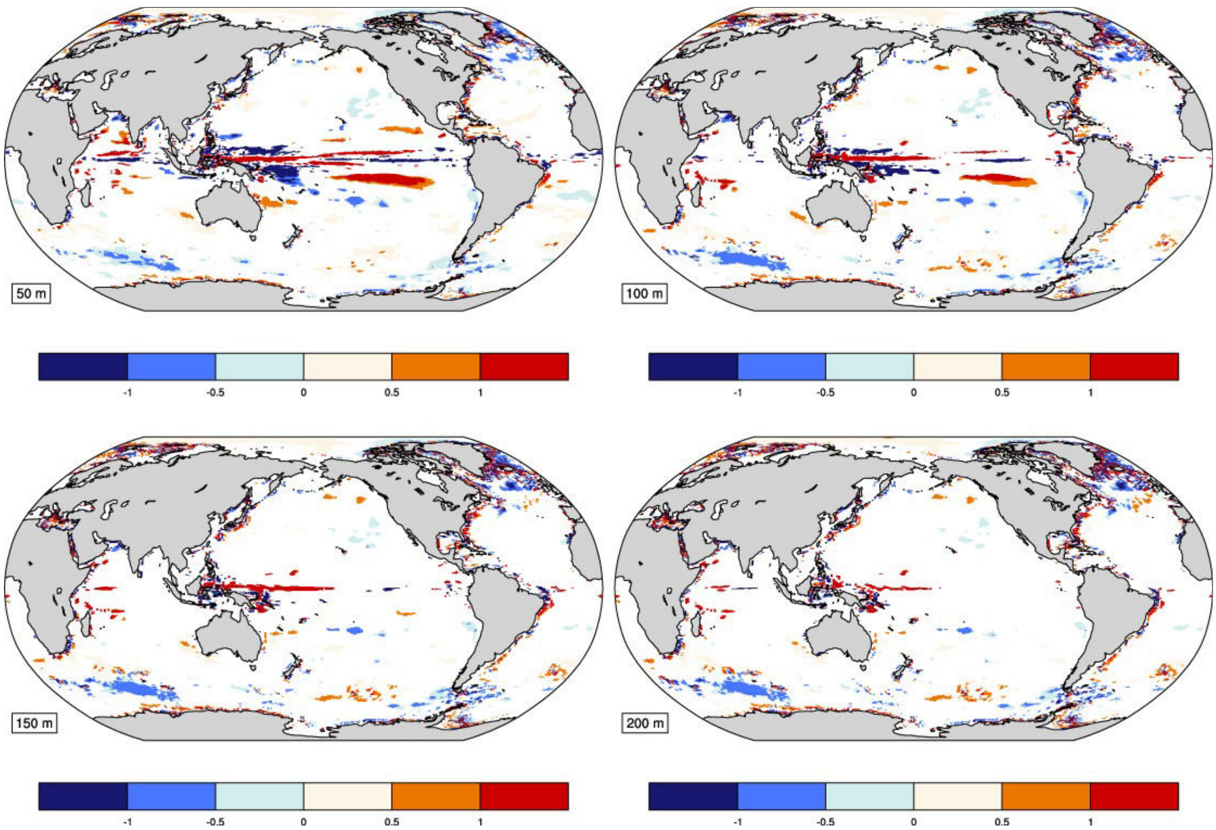
Referee: P378 L4 please mention that the Time of Emergence framework address solely direct changes not climate-carbon cycle feedbacks.

Reply: Given the small carbon-cycle sensitivity in CESM we do not expect this to alter the conclusions discernably. We nevertheless clarify by adding the sentence:

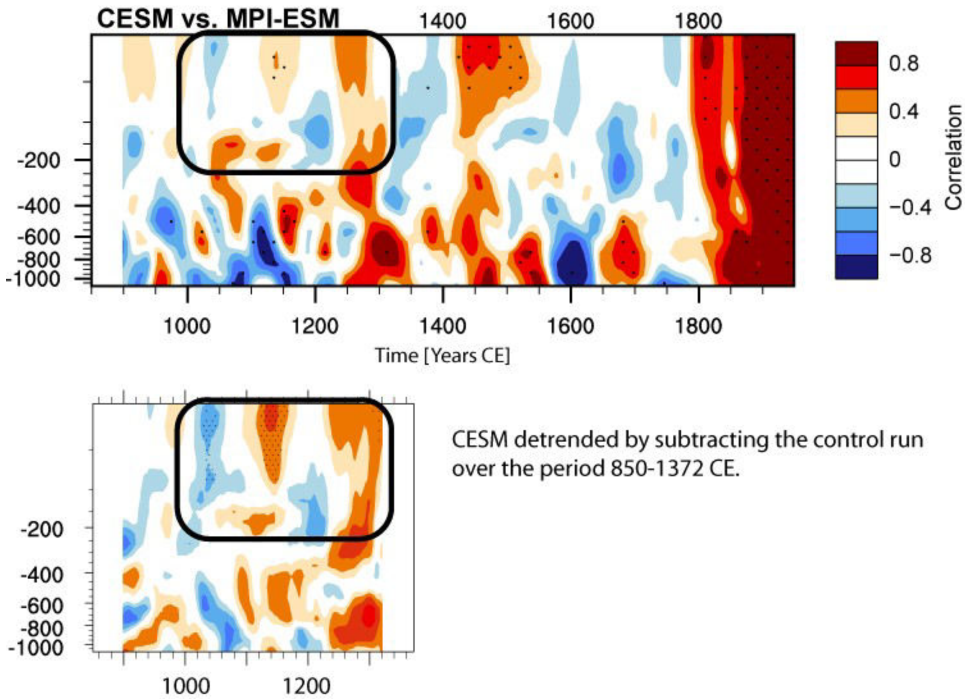
“Note, that these estimates might differ slightly for a radiatively interactive carbon cycle setup.”

Referee: Figure 4 caption: change ‘observational’ by ‘observation-derived’ since GCP data are a combination of several observational source of data plus process-based model reconstruction.

Reply: Done.



Caption Figure 1: Superposed Epoch Analysis on the strongest three volcanic eruptions in CESM for heat flux (W/m2) across different depth in the ocean. Depth labeled in bottom left corner of each panel. Here, the 5 years following an eruption are subtracted from the 5 years preceding an eruption. Only values significant at the 5% value are shaded.



Caption Figure 2: (upper panel) same correlation anaysis as in Fig. 8 of the main manuscript. (lower panel) same as (upper panel) but with the CESM values detrended by the respective segment of the control simulation.

Response to referee #4

This manuscript presents an analysis of CESM Last-Millenium simulations, compares them with other models (esp. MPI) and looks in more depth particularly at aspects of forced variability and carbon cycle response/sensitivity.

I found the manuscript interesting and well written with some good points made. With one exception I found very little to comment on beyond minor presentational aspects, and therefore recommend publication after some minor revisions.

The aspect I would like to dwell on though is the presentation of the carbon cycle "sensitivity". There has been much made on the diagnosis and constrain of this quantity from both short term and long term observations and the area is extremely important, but often controversial or not treated consistently. Hence some caution is required to make sure that the work shown here is not misinterpreted.

I don't disagree with your analysis, nor your figure 13, but there are two main things I think which need to be brought out much more clearly.

1. The quantity "carbon cycle sensitivity" (both sensitivity to temperature and sensitivity to CO2), is not a fundamental quantity that can be measured in one context and applied in another. You acknowledge this briefly in p.375, line 20, but I think it needs more discussion. The processes behind any response are many and varied and have different magnitudes and timescales - hence the global sum of these varies a lot across timescales. For this very reason we assembled a table of processes and timescales in AR5 carbon cycle chapter (Ch. 6, table 6.10). So the key thing to bring out in the discussion is that this is an interesting metric to measure (in obs and models), but the value cannot be compared across timescales or used to infer future behaviour. As a model diagnostic, experiments should be designed so that model behaviour cn be compared with observations sampled in the same way. I'd recommend Friedlingstein and Prentice paper on this: Current Opinion in Env. Sustainability, vol. 2, issue 4, 2010.

Reply: We agree that this might cause confusion and have now extended the explanation to hopefully clarify the distinction of our metric with respect to other studies as well as potential implications.

2. perhaps more importantly, the quantity itself you present here is subtly different from anything I've seen elsewhere, and is really quite different from "gamma" as used to quantify carbon cycle feedback metrics in Friedlingstein 2006 or Arora 2013.

- firstly, their definition of gamma is indeed a true sensitivity - i.e. atmospheric CO2 is held fixed, climate is allowed to vary, and then you can diagnose the impact this has (as an isolated forcing) on land/ocean carbon stores. This is what C4MIP regard as "gamma".

- in observations, and fully coupled models (i.e. with interactive atmospheric CO2), changes in carbon divided by changes in temperature ARE NOT gamma. They fold in the feedback responses - the initial sensitivity to climate is modified because the atmospheric CO2 has changed, and hence the climate changes further, and the carbon stores respond to this. So, for example, the quantity presented in Frank et al is not the C4MIP gamma. That's not to say it's not a good quantity to calculate - but please don't call it gamma. For obvious reasons this gets confusing and wrong comparisons are made between it and other studies.

- in your study here, you present something SIMILAR to Frank et al, but the experiment design is such that it's not exactly the same. By having an interactive CO_2 which the carbon cycle sees, but a prescribed CO_2 which the climate sees, then you have a brand new experimental design. It sits somewhere between the "COUPLED" and "UNCOUPLED" designs of C4MIP. Your carbon cycle can therefore respond to changes in atmospheric CO_2 caused by the climate effect on carbon stores. But the climate itself will not further respond. Hence i would expect exactly what you see - a relatively low value of gamma, because any increase in CO_2 in the atmosphere will be offset by increased uptake (i.e. the C4MIP "beta" term kicks in).

So in summary for the carbon cycle sensitivity section:
i. DONT call it gamma. It's not.

Reply: Ok, done.

ii. DO explain how/why it differs from Friedlingstein gamma, and the Frank quantity

Reply: We included a paragraph contrasting the different sensitivity quantities that exist and how ours fits in.

iii. DO stress more clearly that it can't be extrapolated across timescales due to many different processes

Reply: We do that now in section 6.

iv. also, consider splitting into land/ocean values - these should be readily available from model results and would be interesting to see how the magnitude and lags vary

Reply: We calculated and included those values now.

v. can you also explain why you use NH temperature to define it? Again, this creates a difference from Friedlingstein definition. In observational reconstructions maybe NH is better constrained? but in model results at least global T is available. given you don't compare this result to observations, should you therefore use a global T? or at least justify why not.

Reply: Frank et al. and Jungclaus et al. used NH temperature; to be as comparable as possible we did the same. In response to other referee comments we repeated the analysis with global temperature and included the results in the paper. The conclusions remain unchanged.

some more minor comments follow

- methods section - as this model has a nitrogen cycle, can you mention how you treat N-deposition as a forcing? I imagine there is no standard PMIP protocol for this. e.g. Is anthropogenic N-deposition assumed zero until 20th century?

Reply: For pre-1850 anthropogenic deposition is included in the fixed 1850 prescribed nitrogen deposition values. For post-1850 it follows the references given under Experimental Setup (Lamarque et al.). We now clarify this in that section.

- sec 4.2. You say carbon cycle variability hinders the analysis of phasing between models. Could you remove some of this using a simple regression to ENSO (as you do later for your Pinatubo CO2 figure). This may be an easy way to remove some internal variability in CO2 in the model to let the forced responses show through a bit more. (e.g. if you look at fig 2 of Jones and Cox 2001, the volcanic signal is very clear once a Nino3-regression is removed)

Reply: We tried that without success. The reason this works in Jones and Cox or our Fig. 12 is because they both look at global CO2. We, on the other hand, attempted here to find spatially coherent changes between models (we essentially produced something like Fig. 5 for land and ocean fluxes and found no significant correlations). Regional fluxes are much more noisy and regressing out ENSO did not help that. One reason could be that ENSO has a different influence on the carbon cycle in the two models. In the section on volcanoes, however, we discuss the differences between the models in terms of globally intergrated or averaged quantities. These results are much more robust (even though they still reveal fundamental differences between the two models considered).

- can you check units on figure 9[c]? it says both PgC and ppm? (I assume PgC by looking at panels d and e)

Reply: This is intended and additionally stressed in the caption, but you are right, we should have indicated that it is “Pg C” and “ppm CO2” in the panel title. Done.

- I did like the analysis of the volcanic response of CO2. I wondered if something similar for the climate sensitivity of Carbon would be possible (not here, but as a later study). Looking at the mechanisms which control the changes in your (soon-to-be-renamed!) gamma, would this throw up a clue on how we can evaluate models better? Why do the different timeperiods have different sensitivities? can the model be used to figure out why? There has been a lot of interest in using short term interannual variability to try to constrain carbon cycle sensitivity (Cox et al 2013). There must be some more constrain from palaeo runs/data too, and the first step would be to find the model processes which lead to these time-changes in gamma. A process-understanding of a palaeo carbon cycle constraint would be very powerful!

Reply: Thanks for this valuable comment! There might indeed be merit in proceeding this line of thought in a dedicated study. In particular making use of a larger number of models as well as new paleo proxies.

Manuscript prepared for Earth Syst. Dynam.
with version 2014/09/16 7.15 Copernicus papers of the LaTeX class copernicus.cls.
Date: 18 May 2015

Climate and carbon cycle dynamics in a CESM simulation from 850-2100 CE

Flavio Lehner[1,2,*], Fortunat Joos[1,2], Christoph C. Raible[1,2], Juliette Mignot[1,2,3], Andreas Born[1,2], Kathrin M. Keller[1,2], and Thomas F. Stocker[1,2]

[1]Climate and Environmental Physics, University of Bern, Switzerland
[2]Oeschger Centre for Climate Change Research, University of Bern, Switzerland
[*]now at National Center for Atmospheric Research, Boulder, USA
[3]LOCEAN Laboratory, Sorbonne Universités, France

Correspondence to: Flavio Lehner (lehner@climate.unibe.ch)

Abstract. Under the protocols of the Paleoclimate and Coupled Modelling Intercomparison Projects a number of simulations were produced that provide a range of potential climate evolutions from the last millennium to the end of the current century. Here, we present the first simulation with the Community Earth System Model (CESM), which includes an interactive carbon cycle, that ~~continuously~~ covers the last millennium~~, the historical period, and~~ . The simulation is continued to the end of the twenty-first century. Besides state-of-the-art forcing reconstructions, we apply a modified reconstruction of total solar irradiance to shed light on the issue of forcing uncertainty in the context of the last millennium. Nevertheless, we find that structural uncertainties between different models can still dominate over forcing uncertainty for quantities such as hemispheric temperatures or the land and ocean carbon cycle response. Comparing with other model simulations we find forced decadal-scale variability to occur mainly after volcanic eruptions, while during other periods internal variability masks potentially forced signals and calls for larger ensembles in paleoclimate modeling studies. At the same time, we fail to attribute millennial temperature trends to orbital forcing, as has been suggested recently. The climate-carbon cycle sensitivity in CESM during the last millennium is estimated to be ~~about 1.3~~between 1.0-2.1 ppm °C $^{-1}$. However, the dependence of this sensitivity on the exact time period and scale illustrates the prevailing challenge of deriving robust constrains on this quantity from paleoclimate proxies. In particular, the response of the land carbon cycle to volcanic forcing shows fundamental differences between different models. In CESM the tropical land dictates the response to volcanoes with a distinct behavior for large and moderate eruptions. Under anthropogenic emissions, global land and ocean carbon uptake rates emerge from the envelope of interannual natural variability as simulated for the last millennium by about year 1947 and 1877, respectively.

1 Introduction

The last about 1,000 years constitute the best opportunity previous to the instrumental period to study the transient interaction of external forcing and internal variability in climate, atmospheric CO_2, and the carbon cycle on interannual to multi-decadal time scales. In fact, the instrumental record is often too short to draw strong conclusions on multi-decadal variability. The relatively stable climate together with the abundance of high-resolution climate proxy and ice core data makes the last millennium an interesting target and testbed for modeling studies. Yet, the large and sometimes controversial body of literature on the magnitude and impact of solar and volcanic forcing on interannual to multi-decadal climate variability illustrates the challenges inherent in extracting a robust understanding from a period that is characterized by a small signal-to-noise ratio in many quantities and for which uncertainties in the external forcing remain (e.g., Wanner et al., 2008; PAGES 2k network, 2013; Schurer et al., 2013). In addition, a process-based quantitative explanation of the reconstructed preindustrial variability in atmospheric CO_2 and carbon fluxes is largely missing.

Compared to glacial-interglacial climate change, the last millennium experienced little climate variability, yet there is evidence for distinct climate states during that period (e.g., Lehner et al., 2012b; Keller et al., 2015). Within the last millennium the Medieval Climate Anomaly (MCA, $\sim$950-1250 AD) and the Little Ice Age (LIA, $\sim$1400-1700 AD) are two key periods of documented regional or global temperature excursions suggested to be driven by a combination of stronger solar irradiance and reduced volcanic activity and vice versa, respectively (e.g., Crowley, 2000; Mann et al., 2009; PAGES 2k network, 2013). Despite large efforts in reconstructing (e.g., PAGES 2k network, 2013) and simulating (e.g., Fernandez-Donado et al., 2013; Masson-Delmotte et al., 2013) the transition from the MCA to the LIA, substantial uncertainties remain with respect to the mechanisms at play. Recent studies point towards solar insolation playing a minor role for climate over the last millennium (Ammann et al., 2007; Schurer et al., 2014), while in turn regional feedback processes in response to volcanic eruptions and solar variability need to be considered to explain decadal-scale climate variability (e.g., Lehner et al., 2013; Moffa-Sanchez et al., 2014). At high northern latitudes, the importance of millennial-scale orbital forcing is another debated issue (e.g., Kaufman et al., 2009).

Further, the last millennium offers the possibility to study the natural variability of the carbon cycle and its response to external forcing. Models with a carbon cycle module are extensively tested against present-day observations and widely used for emission-driven future projections (e.g., Hoffman et al., 2014). Yet, there are only few last millennium simulations including a carbon cycle (e.g., Gerber et al., 2003; Jungclaus et al., 2010; Brovkin et al., 2010; Friedrich et al., 2012). The sensitivity of the carbon cycle to climate has been shown to be mostly positive, i.e., with warming additional CO_2 is released to the atmosphere (Ciais et al., 2013). However, the magnitude of this feedback remains poorly constrained by observations (Frank et al., 2010) and models (e.g. Friedlingstein et al., 2006). In particular, determining the role of the land in past and future carbon cycle variability and trends is still challenging. In both idealized (Doney et al., 2006; Joos et al., 2013) and scenario-guided multi-model studies (Friedlingstein et al., 2014) the land constitutes the largest relative uncertainty in terms of intermediate- to long-term carbon uptake.

As for physical climate quantities, explosive volcanic eruptions constitute an important forcing for the carbon cycle. The sensitivity of the carbon cycle to such eruptions have been investigated by Jones and Cox (2001), Frölicher et al. (2011), ~~or Brovkin et al. (2010)~~ Brovkin et al. (2010) , or Tjiputra and Otter (2011) using different Earth System Models. For this short-lived forcing, the land response appears to be the driver of most post-eruption carbon cycle changes, with a range of magnitudes and time horizons associated with the different models. Further, Frölicher et al. (2013) pointed out that the magnitude of the carbon cycle response to volcanoes depends critically on the climate state during the eruption.

The third Paleoclimate Modelling Intercomparison Project (PMIP3; Schmidt et al., 2011) and fifth Coupled Model Intercomparison Project (CMIP5; Taylor et al., 2012) represent joint efforts, in which different modeling groups perform identical experiments, allowing for a systematic comparison of the models (e.g., Schmidt et al., 2014). Here we contribute to the existing set of simulations an integration from 850-2100 CE with the Community Earth System Model, including a carbon cycle module. The aims of this study are (i) to detect coherent large-scale features of forced variability in temperature and carbon cycle quantities, in particular in response to volcanic eruptions, (ii) to investigate the relative role of forcing uncertainty and model structural uncertainty, and (iii) to provide ~~a preindustrial context to the future projections of~~ an estimate of preindustrial variability and the time of emergence from it under anthropogenic climate change. The setup chosen here is unique in a number of ways and taylored to address the aims mentioned before: first, the carbon cycle is fully interactive with the other model components with the exception of the radiation code, which is fed by reconstructed CO_2. This allows us to study the isolated effect of climate on the carbon cycle, while guaranteeing an external forcing consistent with existing reconstructions. Second, the orbital parameters are held constant to study their importance relative to simulations with transient orbital parameters. Third, the solar forcing incorporated in the simulation has a larger amplitude than the majority of PMIP3 simulations and hence enables us to investigate whether the results are sensitive to this amplitude.

This paper is structured as follows: a description of the model and experimental setup is presented in Section 2. In Sections 3 and 4 we address the general simulated climate and carbon cycle evolution and investigate forced and unforced variability of the simulated climate by comparing models to reconstructions and models to models. Sections 5 focuses on the response of the climate and carbon cycle to volcanic forcing. Section 6 deals with estimating the climate-carbon cycle sensitivity in CESM. A discussion and conclusions follow in Section 7.

2 Data and methods

2.1 Model description

The Community Earth System Model (CESM; Hurrell et al., 2013) is a fully-coupled state-of-the-art Earth System Model developed by the National Center for Atmospheric Research (NCAR) and was released in 2010. In terms of physics, CESM relies on the fourth version of the Community Climate System Model (CCSM4; Gent et al., 2011). Additionally, a carbon cycle module is included in CESM's atmosphere, land, and ocean components. The CESM version used here is release 1.0.1 in the so-called 1° version and includes components for the atmosphere, land, ocean, and sea ice, all coupled through a flux coupler.

The atmospheric component of CESM 1.0.1 is the Community Atmosphere Model version 4 (CAM4; Neale et al., 2010), which has a finite volume core with a uniform horizontal resolution of $1.25° \times 0.9°$ at 26 vertical levels. The land component is the Community Land Model version 4 (CLM4; Lawrence et al., 2011), which operates on the same horizontal grid as CAM4 and includes a prognostic carbon-nitrogen cycle that calculates vegetation, litter, soil carbon, vegetation phenology, and nitrogen states. Further, it includes a prognostic fire module, which is governed by near-surface soil moisture conditions and fuel availability.

The ocean component is the Parallel Ocean Program version 2 (POP2; Smith et al., 2010; Danabasoglu et al., 2012) with an nominal 1° horizontal resolution and 60 depth levels. The horizontal resolution varies and is higher around Greenland, to where the North Pole is displaced, as well as around the equator. Embedded in POP2 is the Biogeochemical Elemental Cycle model (BEC; Moore et al., 2004) that builds on a nutrient-phytoplankton-zooplankton-detritus food web model and distinguishes three phytoplankton functional types (Long et al., 2013). Carbon export and remineralization are parameterized according to Armstrong et al. (2002). Alkalinity, pH, partial pressure of CO_2, and concentrations of bicarbonate, and carbonate ions are diagnosed from prognostic dissolved inorganic carbon, alkalinity, and temperature- and salinity-dependent equilibrium coefficients. Biogenic calcification is implemented as being proportional to a fraction of small phytoplankton production, which is temperature-dependent. An exponential curve is prescribed to simulate the dissolution of sinking $CaCO_3$ (Moore et al., 2004) . There exists no dependence of calcification-dissolution rates on saturation state. Organic material reaching the ocean floor is remineralized instantaneously, i.e., no sediment module is included. River discharge from CLM4 does not carry dissolved tracers but nitrogen deposition to the ocean surface has been prescribed. The sea ice component is the Community Ice Code version 4 (CICE4) from the Los Alamos National Laboratories (Hunke and Lipscomb, 2010), including elastic-viscous-plastic dynamics, energy-conserving thermodynamics, and a subgridscale ice thickness distribution. It operates on the same horizontal resolution as POP2.

2.2 Experimental setup

Table 1 provides an overview of the simulations conducted for this study. First, a 500-year control simulation with perpetual 850 CE forcing (hereafter CTRL) was branched from a 1850 CE control simulation with CCSM4 (Gent et al., 2011). However, restart files for the land component were taken from a 850 CE control simulation, kindly provided by the NCAR, in which the land use maps by Pongratz et al. (2008) were applied. This procedure has the advantage that the slow-reacting soil and ecosystem carbon stocks are closer to 850 CE conditions than in the 1850 CE control simulation. A transient simulation covering the period 850-2099 CE was then branched from year 258 of CTRL. Despite the shortness of CTRL leading up to the start of the transient simulation, most quantities of the surface climate, such as air temperature, sea ice, or upper ocean temperature, can be considered reasonably equilibrated at the start of the transient simulation, as the forcing levels due to TSI and most greenhouse gases are similar between 1850 and 850 CE (Landrum et al., 2013). However, weak trends in CTRL are still detectable in slow-reacting quantities such as deep ocean temperature (below 2,000 m; $\sim -0.04\,°\mathrm{C}\ 100\,\mathrm{yr}^{-1}$), Atlantic Meridional Overturning Circulation (~ -0.22 Sv 100 yr^{-1}), Antarctic Circumpolar Current (∼0.70 Sv 100 yr^{-1}), dissolved inorganic carbon in the ocean (∼-0.01 % 100 yr^{-1}), or soil carbon storage (∼~~4~~0.2 ~~Pg~~% C 100 yr^{-1}). The Antarctic Bottom Water formation rate shows no drift.

The applied transient forcing largely follows the PMIP3 protocols (Schmidt et al., 2011) and the Coupled Model Intercomparison Project 5 (CMIP5; Taylor et al., 2012), consisting of total solar irradiance (TSI), greenhouse gases (GHGs), volcanic and anthropogenic aerosols, and land use changes (Fig. 1). Here, the TSI reconstruction by Vieira and Solanki (2010, TSI$_{VS09}$) is used, to which a synthetic 11-year solar cycle is added (Schmidt et al., 2011). In light of the recently enlarged envelope of reconstructed TSI amplitude (Schmidt et al., 2012), we scale TSI by a factor of 2.2635 to have an amplitude of 0.2% between present-day (1961-1990 CE) and the late Maunder Minimum (1675-1704 CE), which is about twice as large as the 0.1 % used in most PMIP3 simulations:

$$TSI = 2.2635 \cdot (TSI_{VS09} - \overline{TSI_{VS09}}) + \overline{TSI_{VS09}}. \quad (1)$$

Fig. 1a shows that the TSI used here lies in between the large-amplitude reconstruction by Shapiro et al. (2011) and the bulk of small-amplitude reconstructions of the original PMIP3 protocol (Schmidt et al., 2011). Note, that a recent detection and attribution study indicates small amplitude TSI reconstructions to agree better with temperature reconstructions over the last millennium than large amplitude reconstructions (Schurer et al., 2014), in agreement with Ammann et al. (2007). For the twenty-first century the last three solar cycles of the data set are repeated continuously. The insolation due to Earth's orbital configuration is calculated according to Berger (1978) with the orbital parameters held constant at 1990 CE values.

The volcanic forcing follows Gao et al. (2008) from 850-2001 CE. It provides estimates of the stratospheric sulfate aerosol loadings from volcanic eruptions as a function of latitude, altitude, and month and is implemented in CESM as a fixed single-size distribution in the three layers in the lower stratosphere (Neale et al., 2010). Post-2001 CE volcanic forcing remains zero.

Land use and land use changes (LULUC) are based on Pongratz et al. (2008) from 850 to 1500 CE, when this dataset is splined into Hurtt et al. (2011), a synthesis dataset that extends into the future. The two datasets do not join smoothly but exhibit a small step-wise change in the distribution of crop land and pasture at the year 1500 CE. Up until about 1850 CE global anthropogenic LULUC are small, however, can be significant regionally (Hurtt et al., 2011). Towards the industrial era LULUC accelerate, dominated by the expansion of crop land and pasture. Here, only net changes in land use area are considered. The impact of shifting cultivation and wood harvest on carbon emissions from land use is neglected; these processes are estimated to have contributed order 30% to the total carbon emissions from land use (Shevliakova et al., 2009; Houghton, 2010; Stocker et al., 2014).

The temporal evolution of long-lived greenhouse gases (GHGs: CO_2, CH_4, and N_2O) is prescribed based on estimates from high-resolution Antarctic ice cores that are joined with measurements at mid-twentieth century (Schmidt et al., 2011, and references therein). While the carbon cycle module of CESM interactively calculates the CO_2 concentration originating from land use changes, fossil fuel emissions (post-1750 CE, following Andres et al., 2012), and carbon cycle-climate feedbacks, it is radiatively inactive. Instead, the ice core and measured data are prescribed in the radiative code, keeping the physical model as close to reality as possible. As a result, the impact of the interactive coupling of the carbon cycle module is minor for simulated climate, and limited to changes in surface conditions due to changing vegetation. For the extension of the simulation post-2005 CE the Representative Concentration Pathway 8.5 (RCP 8.5) is used, representing the unmitigated "business-as-usual" emission scenario, corresponding to a forcing of approximately 8.5 W m^{-2} at the year 2100 (Moss et al., 2010).

Aerosols such as sulfate, black and organic carbon, dust, and sea salt are implemented as non-time-varying up to 1850 CE, perpetually inducing the spatial distributions of the 1850 CE control simulation during this time (Landrum et al., 2013). Post-1850 CE, the time-varying aerosol datasets provided by Lamarque et al. (2010, 2011) are used, whereby CESM only includes a representation of direct aerosol effects. Similarly, nitrogen (NH_x and NO_y) input ~~to the ocean~~ is held constant until it starts to be time-varying from 1850 CE onwards, also following Lamarque et al. (2010, 2011). Iron fluxes from sediments are held fixed (Moore and Braucher, 2008).

2.3 Other model simulations

Besides to output from current Model Intercomparison Projects, we compare CESM results to those from a similar simulation with CCSM4 (Landrum et al., 2013) and IPSL-CM5A-LR (Sicre et al., 2013), two simulations without interactive carbon cycle. ~~Further, we compare to MPI-ESM (ECHAM5/MPIOM; Jungclaus et al., 201~~

~~assess the robustness of the simulated climate and carbon cycle variations in response to external forcing.~~ The solar and volcanic forcing ~~of~~ reconstructions applied to CCSM4 and IPSL-CM5A-LR are identical to ours with the exception of the scaling of TSI that we applied to CESM. The goal here is to investigate the question whether different solar forcing amplitudes applied to the same physical model (CESM vs. CCSM4) have a larger effect than applying the same solar forcing to two different physical model (IPSL-CM5A-LR vs. CCSM4).

Further, we compare CESM to MPI-ESM (ECHAM5/MPIOM; Jungclaus et al., 2010; Friedrich et al., 2012) , a model that includes an interactive carbon cycle, to assess the robustness of the simulated climate and carbon cycle variations in response to external forcing. MPI-ESM uses Krivova et al. (2007) as TSI forcing and Crowley et al. (2008) as volcanic forcing. These differ from the CESM forcing in amplitude much more than in timing and therefore allow for a comparison of the forced response. All simulations, except ours, apply transient orbital forcing and are summarized in Table 2. If not using the full ensemble of MPI-ESM, we focus on the member "mil0021". Another difference in terms of experimental setup between CESM and MPI-ESM is that MPI-ESM was run with a fully interactive carbon cycle, i.e., the prognostic CO_2 interacts with the radiation and through that again influences climate, while in our setup this is a one-directional interaction only. Further, MPI-ESM is coarser resolved than CESM in both ocean and atmosphere and applies the A1B scenario for the twenty-first century (IPCC, 2000), which corresponds roughly to the current intermediate scenario as compared to the high scenario RCP8.5 used in CESM.

3 General climate and carbon cycle evolution

3.1 Temperature

The simulated annual mean Northern Hemisphere (NH) surface air temperature (SAT) follows the general evolution of proxy reconstructions: a warm Medieval Climate Anomaly (MCA, ~950-1250 CE), a transition into the colder Little Ice Age (LIA, ~1400-1700 CE), followed by the anthropogenically driven warming of the nineteenth and twentieth century (Fig. 3). The NH MCA-to-LIA cooling amounts to 0.26 ± 0.18°C (taking the time periods defined above, which are as in Mann et al., 2009), placing it in the lower half of reconstructed amplitudes that range from about 0.1 to 0.7°C (IPCC, 2013). The subsequent warming from 1851-1880 CE to 1981-2010 CE amounts to 1.23 ± 0.15°C , while observations report only 0.71 ± 0.13°C (Cowtan and Way, 2014). This overestimation by CESM takes place almost entirely after 1960 and arises largely from missing negative forcing from the indirect aerosol effect, which is not implemented in CAM4 (Meehl et al., 2012). The late twentieth century being the warmest period in the NH in the past millennium is consistent with reconstructions (e.g., PAGES 2k network, 2013).

In CESM, the inception of the NH LIA occurs in concert with decreasing TSI and a sequence of strong volcanic eruptions during the thirteenth century. Reconstructions differ substantially in this matter and start to cool as early as 1100 CE or as late as 1400 CE. Further, new regional multi-proxy reconstructions of temperature provide no support for a hemispherical or globally synchronous MCA or LIA but show a clear tendency towards colder temperatures and exceptionally cold decades over most continents in the second half of the millennium (PAGES 2k network, 2013; Neukom et al., 2014).

The last millennium simulation with CCSM4 shows a largely coherent behavior with CESM in terms of amplitude and decadal variability of NH SAT (850-1850 CE correlation of 5-year filtered annual means $r = 0.88$, p<0.001). The difference in NH SAT due to the different TSI amplitudes in CESM and CCSM4 scales roughly with the regression slope of NH SAT vs. TSI of both CESM and CCSM4 ($\sim$0.13 °C per W m^{-2}), although internal variability can easily mask this effect at times. For example, the Maunder Minimum (1675-1704 CE), the 30-year period with the lowest TSI values and – by construction of the TSI scaling – with the largest difference between CESM and CCSM4 (1.5 W m^{-2}), is only 0.14°C cooler than in CCSM4 and not 0.20°C as expected from the regression.

The NH temperature evolution of additional PMIP3 and CMIP5 simulations shows that the multi-model range is within the one of the reconstructions and encompasses the instrumental-based observations (Fig. 3). Disagreement between models and reconstructions exists in particular on the magnitude of response to the eruptions at 1258 CE and around 1350 CE. The 1258 CE eruption is the largest volcanic event recorded for the last millennium and its climatic impact was likely enhanced through the cumulative effect of three smaller eruptions following shortly after (Gao et al., 2008; Crowley et al., 2008; Lehner et al., 2013). However, the pronounced cooling that is simulated by the models for this cluster of eruptions is largely absent in temperature reconstructions. Conversely, around 1350 CE temperature reconstructions show a decadal-scale cooling presumably due to volcanoes that is absent in the models, as the reconstructed volcanic forcing shows only two relatively small eruptions around that time. Part of this incoherent picture may arise from the unknown aerosol size distribution (Timmreck et al., 2010) and geographic location of past volcanic eruptions (Schneider et al., 2009), and differences in reconstruction methods. As many proxy reconstructions of temperature rely heavily on tree ring data it is worth noting that the dendrochronology community currently debates whether the trees' response to volcanic eruptions resembles the true magnitude of the eruption (Mann et al., 2012; Anchukaitis et al., 2012; Tingley et al., 2014).

Disagreement among the models exists on the relative amplitude of the MCA, where most models show colder conditions than CESM and CCSM4. Remarkably, the simulation by IPSL-CM5A-LR applied the same TSI and volcanic forcing as CCSM4, yet it comes to lie at the lower end of the PMIP3 model range during the MCA. In other words, the way how models respond to variations in TSI and other forcings can still make a larger difference in the simulated amplitude than the scaling of TSI by a factor of 2, which in turn complicates a proper detection and attribution of solar forcing during the last millennium (Servonnat et al., 2010; Schurer et al., 2014). Further disagreement among the models exists on the response to volcanic eruptions, where CESM and CCSM4 are among the more sensitive models (an oversensitivity of CCSM4 to volcanoes based on twentieth century simulations was reported by Meehl et al., 2012). Turning to the century-scale change over the industrial era, CESM and CCSM4 are on the upper end of the CMIP5 range and show an overestimation of the observed warming.

The simulated mean SAT of the Southern Hemisphere (SH) generally shows a similar evolution as for the NH with the signature of the MCA and LIA superimposed on a weak millennial cooling trend. Models and reconstructions disagree to a larger extent in the SH than in the NH, in particular regarding cold excursions due to large volcanic eruptions, which are largely absent in the reconstructions. Similar results have been reported in a recent study on interhemispheric temperature variations that finds much less phasing of the two hemispheres in reconstructions than in models, potentially related to underestimated internal variability on the SH in models (Neukom et al., 2014). A lingering question of climate modeling in general is whether models are too global in their response to external forcing. That is, they might show too little regional variability that is independent from the global mean response during a forced period. However, the uncertainties in the early period of the reconstructions prohibits to robustly answer ~~the questionwhether the models are too global in their response to external forcing~~this question. Similar to the NH, the industrial warming in the SH from 1851-1880 CE to 1981-2010 CE (0.53 ± 0.07°C) is overestimated by CESM (0.71 ± 0.13°C).

The differential warming between the hemispheres in CESM is among the smallest among CMIP5 models (not shown). This is mainly due to the underestimated deep water formation in the Southern Ocean, leading to a comparably strong warming of the SH and likely an underestimation of the oceanic uptake of anthropogenic carbon (Long et al., 2013). With a transient climate response of 1.73°C and an equilibrium climate sensitivity of 3.20°C (Meehl et al., 2012), CESM lies in the middle of recent estimates of 1.0 to 2.5°C and 1.5 to 4.5°C , respectively (IPCC, 2013).

3.2 Orbital forcing

To detect and attribute the influence of orbital forcing on SAT trends during the last millennium, we compare our simulation with fixed orbital parameters to the CCSM4 simulation with time-varying orbital parameters (Fig. 2). While both models experience a negative long term trend in global TSI until about 1850 CE (Fig. 1), the difference arising from the different orbital setup can be seen best in Arctic summer land insolation (Fig. 2). Hence, Arctic summer land SAT has been proposed as a quantity to be affected by orbital forcing already on time scales of centuries to millennia (Kaufman et al., 2009). However, we find no detectable difference between the two simulations in the trend of Arctic summer land SAT (Fig. 2b). In fact, the Arctic multi-decadal to centennial summer land SAT anomalies in CESM span a very similar range as in CCSM4, despite CESM not accounting for time-varying orbital parameters: Fig. 2c shows non-overlapping 100- and 200-year mean SAT anomalies plotted against the corresponding mean solar insolation. The results from CCSM4 suggest a clear relation of the two quantities, however, the results of CESM show that nearly identical SAT anomalies are possible without orbital forcing. In other words, while we detect a long-term cooling trend in Arctic summer SAT in both CESM and CCSM4, we fail to attribute this trend to orbital forcing alone, as suggested by Kaufman et al. (2009). This is confirmed in new simulations with decomposed forcing, again comparing simulations with fixed and time-varying orbital parameters (B. Otto-Bliesner, personal communication).

3.3 Carbon cycle

The prognostic carbon cycle module in CESM allows us to study the response of the carbon cycle to transient external forcing. The land biosphere is a carbon sink during most of the first half of the last millennium, but becomes a source as anthropogenic land cover changes start to have a large-scale impact on the carbon cycle (Table 3). The ocean is a carbon source at the beginning and becomes a sink in the second half of the last millennium~~(not shown)~~. The residual of these fluxes represents changes in the atmospheric reservoir of carbon, illustrated in Fig. 3c by the prognostic CO_2 concentration. The amplitude of the simulated concentration does not resemble the one reconstructed from ice cores (i.e., imposed on the radiative code of CESM), in particular the prominent CO_2 drop in the seventeenth century is not captured by CESM. This raises the question whether the sensitivity of the carbon cycle to external forcing is too weak in CESM, whether the imposed land use changes are too modest (Kaplan et al., 2011; Pongratz et al., 2011), whether major changes in ocean circulation are not captured by models (Neukom et al., 2014), or whether the ice core records are affected by uncertainties due to in-situ production of CO_2 (Tschumi and Stauffer, 2000). Ensemble simulations with MPI-ESM also do not reproduce the reconstructed amplitudes or the drop (Jungclaus et al., 2010). Further, Earth System Models of Intermediate Complexity or vegetation models driven by GCM output do not reproduce the uptake of carbon by either ocean or land needed to explain the reconstructed amplitudes (Stocker et al., 2011; Gerber et al., 2003).

The rise in atmospheric CO_2 due to fossil fuel combustion is in good agreement with ice cores until about the 1940s. After that, a growing offset exists, leading to an overestimation of about 20 ppm by 2005 in CESM, qualitatively similar to the CMIP5 multi-model mean (Hoffman et al., 2014). From the observational estimates one can diagnose that the discrepancy arises primarily from overestimated carbon release from land (Fig. 4a; see also Hoffman et al., 2014; Lindsay et al., 2014). From 1750 to 2011 CE the cumulative total land release (including LULUC) is ~~80~~83 Pg C (compared to 30±45 Pg C from observational estimates Ciais et al., 2013), while the cumulative net land uptake is ~~101~~95 Pg C (160±90 Pg C Ciais et al., 2013). The ocean cumulative uptake of ~~154~~151 Pg C compares more favorably to current estimates of 155±30 Pg C (Ciais et al., 2013). Note, however, that given the overestimation of atmospheric CO_2, one would expect a higher ocean uptake. This bias originates largerly from an underestimation of the uptake in the Southern Ocean (Long et al., 2013). Along with this goes an underestimated seasonal cycle in CESM, originating from a too weak growing season net flux in CLM4 (Keppel-Aleks et al., 2013). MPI-ESM, on the other hand, underestimates atmospheric CO_2 due to weak emissions from LULUC (Pongratz et al., 2008).

The twenty-first century sees substantial emissions from fossil fuel burning under RCP 8.5 (Fig. 3c). In addition, LULUC is associated with a positive flux into the atmosphere, particularly until around 2050 CE (~~not shown~~Table 3). After accounting for LULUC (which constitutes a carbon loss for the land) the net land sink increases to about 7 Pg C yr^{-1} at the end of the twenty-first century (Fig. 4a). The rate of ocean uptake, on the other hand, peaks around 2070 at about 5 Pg C yr^{-1}, despite that atmospheric CO_2 continues to rise (Fig. 3c). This decoupling of the trends in atmospheric CO_2 growth and ocean uptake flux is linked to non-linearities in the carbon chemistry (Schwinger et al., 2014) . The change in dissolved inorganic carbon per unit change in the partial pressure of CO_2 decreases with increasing CO_2, and thus the uptake capacity of the ocean. Additionally, differences in the ventilation time scales of the upper and the deep ocean likely play a role. While the surface ocean and the thermocline exchanges carbon on annual-to-multi-decadal time scales with the atmosphere, it takes century to ventilate the deep ocean as evidenced by chlorofluorocarbon and radiocarbon data ~~(Key et al., 2004) .~~ (Key et al., 2004; Wunsch and Heimbach, 2007, 2008) . CESM has a documented low bias in Southern Ocean ventilation due to too shallow mixed layer depths, contributing to the underestimated carbon uptake of the ocean (Long et al., 2013) .

The prognostic atmospheric CO_2 increases to 1,156 ppm by 2100 CE. This would imply a forcing of 7.6 W m^{-2} from CO_2 relative to 850 CE, significantly more than the approximately 6.5 W m^{-2} that are imposed by the radiative code ~~(see Fig. 1c).~~ (calculated according to IPCC, 2001, , see also Fig. 1c) . This propagation of the twentieth century bias is consistent with the CMIP5 multi-model mean (Friedlingstein et al., 2014) and has motivated attempts to reduce such biases by using observational constraints for ocean ventilation (Matsumoto et al., 2004), the tropical land carbon storage sensitivity to temperature variations (Cox et al., 2013; Wenzel et al., 2014; Wang et al., 2014) and for the oceanic and terrestrial carbon fluxes (Steinacher et al., 2013). CESM with CLM4, however, shows very little sensitivity in tropical land carbon, in part due to the inclusion of an interactive nitrogen cycle, which – through enhanced photosynthetic uptake due to nitrogen fertilization – tends to counteract accelerated soil decomposition from warming (Lawrence et al., 2012; Wenzel et al., 2014). Together with the underestimated oceanic uptake this leads to the roughly 20% larger airborne fraction in CESM as compared to ~~the~~ what is actually prescribed as atmospheric concentration in the radiative code according to the RCP 8.5.

Fig. 4 puts the current and projected changes into perspective of preindustrial variability. Estimated interannual variability prior to 1750 CE is ± 0.94 Pg C yr^{-1} (1 standard deviation) for the net atmosphere-land and ± 0.42 Pg C yr^{-1} for the net atmosphere-ocean flux. The much larger interannual variability in land than ocean flux is consistent with independent estimates and results from other models (e.g., Ciais et al., 2013). Large volcanic eruptions, as they have occurred in the last millennium, cause anomalously high uptake rates that for a short period of time are on par with current uptake rates (Fig. 4a and b, full range). We estimate when the anthropogenically forced, global-mean land and ocean uptake fluxes leave the bound of preindustrial natural variability (Hawkins and Sutton, 2012; Keller et al., 2014). As a threshold criteria, it is required that the decadal-smoothed uptake fluxes are larger than the upper bound of 2 standard deviations of the annual fluxes prior to 1750 CE. Then, the simulated global-mean land and ocean uptake fluxes have emerged from natural interannual variability by 1947 CE and by 1877 CE, respectively. The prognostic atmospheric CO_2 concentration emerges already in 1755 CE, while the simulated global-mean temperature does not emerge until 1966 CE.

4 Model-model coherence

A classical approach to assess the robustness of model results is to rely on the multi-model mean response to a given forcing (IPCC, 2013). However, as there are only very few last millennium simulations with comprehensive Earth System Models to date, this approach is not feasible to investigate the decadal-scale climate-carbon cycle responses to external forcing in the period before 1850 CE. Instead, we estimate periods of forced variability with a 100-year running-window correlation of CESM and MPI-ESM, indicating phasing of the two models. The time series are anomalies to their 850-1849 mean and are smoothed with a 5-year local regression filter before calculating the correlation. Thereby, we focus on the preindustrial period, as the twentieth and twenty-first century are dominated by anthropogenic trends, which are non-trivial to remove for a proper correlation analysis. In addition, regression analysis is used.

4.1 Temperature

Fig. 5a and b show anomalies of zonal mean annual SAT from CESM and MPI-ESM. In both models the northern high latitudes show the strongest trend from positive anomalies during the MCA to negative anomalies during the LIA. This is consistent with the current understanding of polar amplification during either warm or cold phases (Holland and Bitz, 2003; Lehner et al., 2013). The twentieth and twenty-first century then see the strong anthropogenic warming, although this occurs earlier in CESM due to missing negative forcings from indirect aerosol effects (section 2). Superimposed on the preindustrial long-term negative trend are volcanic cooling events. In CESM many of these are global and are able to considerably cool the SH extra-tropics around 60° S, while in MPI-ESM the SH extra-tropics are only weakly affected. These differences are likely related to the Southern Ocean heat uptake rates in the two models (arising from under- and overestimation of Southern Ocean mixed layer depths in CESM and MPI-ESM, respectively; Danabasoglu et al., 2012; Marsland et al., 2003). This is evident also in the delayed warming at these latitudes in the twenty-first century in MPI-ESM as compared to CESM. The consistent SH high latitude positive anomalies before the thirteenth century, on the other hand, appear to be related to a positive phase of the Southern Annular Mode (SAM) in both models (not shown), a behavior common to most PMIP3 models. Note, however, that a recent reconstruction of the SAM finds the models to lack amplitude in their simulated variability, challenging the models' capabilities to represent SAM (Abram et al., 2014).

The phasing on interannual to decadal scales between the two models is largely restricted to periods of volcanic activity and within those mainly to land-dominated latitudes (except Antarctica, which shows no forced variability on these time scales; Fig. 5c). Despite the largest absolute temperature anomalies occurring in the Arctic, the correlations are highest in the subtropics, due to the smaller interannual variability there. Periods of centennial trends, such as the MCA or the Arctic cooling during the Maunder Minimum around 1700 CE, do not show up in the correlation analysis that focuses on 100-year windows, suggesting multi-decadal low-frequency forcing, such as centennial TSI trends, or internal feedback mechanisms to be responsible for the missing correlation. A regression analysis between the 5-year filtered annual TSI and SAT at each gridpoint (different filter lengths of up to 50 years have been tested as well without changing the results) reveals a clear link of the two quantities at high latitudes. In CESM this seems to be driven primarily by a displacement of the sea ice edge (Arctic) and Southern Ocean heat uptake (Fig. 6a). As the sea ice response has not been detected in an earlier model version (Ammann et al., 2007, their Fig. 4), it warrants the questions whether the regression of SAT on TSI might be biased by imprints of volcanoes (Lehner et al., 2013), even when the timeseries are filtered, especially in a model like CESM that has a very strong volcanic imprint.

Forthcoming simulations with solar-only forcing will be able to answer that question. MPI-ESM, on the other hand, shows a similar polar amplification signal from solar forcing, but not as clearly linked to sea ice (Fig. 6b). MPI-ESM also displays a stronger land-ocean contrast than CESM (see also Geoffroy et al., 2015) .

In addition to the comparison with MPI-ESM, Fig. 5d shows results from the correlation analysis between CESM and CCSM4, two simulations that in terms of physics differ only in their applied TSI amplitude and orbital parameters. Not unexpected, there are generally more robust signals of forced variability as compared to CESM vs. MPI-ESM (Fig. 5c), very likely due to the identical physical model components in CESM and CCSM4. Similarly, global mean SAT shows generally stronger phasing between CESM and CCSM4 (Fig. 5e). However, the latitudinal and temporal pattern of the CESM vs. CCSM4 analysis agrees well with the one arising from CESM vs. MPI-ESM (Fig. 5c; with exception of the much stronger phasing in CESM and CCSM4 during the volcanic eruptions in the 1450s) and suggest the physical mechanism behind periods of phasing to be robust across the two models.

Applied to ocean temperature, the above approach enables us to investigate the penetration depth of a forced signal seen at the surface (Fig. 7). Indeed, most of the surface signals also show up as significant correlations down to depths of about 150-200 m, whereby their timing suggests again volcanic forcing as the origin. Reduced heat loss from the tropical equatorial Pacific together with reduced heat uptake at high latitudes are responsible for ocean cooling after volcanoes (not shown). The Atlantic Meridional Overturning Circulations (AMOC) in the CESM and MPI-ESM shows no significant correlation, however, the highest correlation occurs during the thirteenth century and coincides with a phasing of the upper ocean temperatures due to strong volcanic forcing (Fig. 7d). The correlation between CESM and CCSM4 at that time is even higher and points to a significant imprint of the volcanic forcing on ocean circulation ~~(Otterå et al., 2010; Swingedouw et al., 2013)~~ (Otterå et al., 2010; Swingedouw et al., 20 However, during the remainder of the millennium, no phasing of the AMOC is found.

4.2 Carbon cycle

We apply the same correlation analysis to zonally integrated land and ocean carbon fluxes from the two models to detect forced variability in the carbon cycle. Compared to SAT hardly any phasing can be found between the models in atmosphere-to-land carbon fluxes (not shown), which is due to its large interannual variability and to distinctly different responses to external forcing in the two models, as will be illustrated in section 5. Similarly, there is little model phasing in net atmosphere-to-ocean carbon fluxes (not shown). Results become somewhat clearer when considering globally integrated upper-ocean dissolved inorganic carbon (DIC; Fig. 8). ~~While there~~ There appear to exist spurious trends ~~at depth in both models,~~ in CESM, likely related to model drift. We repeated the analysis, but with the CESM output detrended in each grid cell by subtracting the CTRL over the corresponding period 850-1372 CE. Due to the shortness of CTRL, we cannot apply this to the whole simulation. However, these tests showed that the correlation between the two simulation is largely insensitive to the drift in CESM. In Fig. 8c there are periods of coherent carbon draw-down coinciding with volcanic eruptions around 1450 CE and 1815 CE in response to temperature-driven solubility changes. Interestingly, MPI-ESM shows a distinct behavior for the strong eruption of 1258 CE, with a prolonged ocean carbon loss after a weak initial uptake. CESM shows a stronger and more sustained carbon uptake, leading to no correlation between the two models for this eruption. The reasons for this discrepancy are discussed in section 5.

Generally, the largest changes in upper-ocean carbon storage occur in response to volcanoes and take place in the tropical Pacific (Chikamoto et al., submitted), with other significant changes occurring in the North and South Pacific, the subtropical Atlantic and the Arctic (section 5). Within the tropical oceans, the models show different characteristics: CESM shows a larger variability in DIC than MPI-ESM and, when influenced by anthropogenic emissions in the twentieth and twenty-first century, takes up a larger portion of the total ocean carbon uptake than in MPI-ESM (not shown). In MPI-ESM, the Southern Ocean shows stronger variability and larger carbon uptake in the twenty-first century, illustrating the different behavior of the two models in terms of ocean carbon cycle variability and trend magnitude, closely related to the different mixed layer depth in the Southern Ocean region.

5 Volcanic forcing

To further isolate the response of the climate system and carbon cycle to volcanic eruptions, a Superposed Epoch Analysis is applied to both simulations. Thereby, composite time series for the strongest three (top3) and following strongest seven eruptions (top10), by measure of optical depth anomaly, over the period 850-1850 CE are calculated for the CESM and MPI-ESM (Fig. 9). The time series are calculated as deseasonalized monthly anomalies to the 5 years preceding an eruption.

The physical parameters global mean surface air temperature and global mean precipitation decrease in both models after volcanic eruptions, although the response of CESM is stronger by roughly a factor 2-2.5 (Fig. 9a, b, f, g). Consequently, CESM temperature and precipitation take longer ($\sim$15 years) to relax back to pre-eruption values than MPI-ESM ($\sim$9 years).

The atmospheric carbon inventory, on the other hand, shows a remarkably different response in the two models. In CESM the atmosphere initially looses about 2-3 Pg C, irrespectively of the eruption strength, with the minimum occurring after about 1-2 years. In the top10 case values return to normal after about 16 years, while in the top3 case they tend to return already after about six years, and overshoot. This overshoot is not straightforward to understand and did not seem to occur in earlier versions of the model ~~(Frölicher et al., 2011)~~ (Frölicher et al., 2011; Rothenberg et al., 2012) . In MPI-ESM the response is a priori more straightforward and slower: in the top10 case the atmosphere looses about 2.5 Pg C, reaches a minimum after 2-4 years, and returns to pre-eruption values after 10-16 years. The top3 case reaches its minimum (-6 Pg C) a bit faster, but then takes about 20 years to return to pre-eruption values (Brovkin et al., 2010).

Partitioning these atmospheric carbon changes into land and ocean changes indicates that the land is primarily responsible for the differing response behavior of the two models, confirming the findings in the previous section. While in both models the land drives the atmospheric change by taking up carbon initially, it is released back to the atmosphere within about 3 years in CESM, but kept in the land for at least 15 years in MPI-ESM (and up to 50 years for the 1258 CE eruption; Brovkin et al., 2010). In the top3 case of CESM the land starts to even loose carbon after about 5 years, causing the overshoot seen in the atmospheric carbon.

A closer look at CESM reveals a distinct response to the top3 and the top10 volcanoes. The response to top3 must be understood as an interplay of a number of processes: the initial global cooling triggers a La Niña-like response and a corresponding cloud and precipitation reduction that is particularly pronounced over tropical land, where also large changes in carbon storage occur (see Fig. 11a-c for the spatial pattern). Fig. 10 and the following analysis therefore focuses on tropical land. Direct solar radiation decreases, indirect radiation increases, with a net decrease (Fig. 10d). These unfavorable conditions cause a reduction in net primary productivity and a strong decrease of vegetation (-8 Pg C; Fig. 10a and e). At the same time, decomposition of dead biomass becomes less efficient due to reduced temperature (similar to, e.g., Frölicher et al., 2011). Despite the simultaneous decrease in net primary production this results in a build-up of dead biomass of about 5 Pg C

(Fig. 10b). ~~Although carbon loss due to fire increases, it~~ Due to the dry conditions and availability of dead biomass there is increased fire activity, leading to increased carbon loss from land. However, the fire cannot get rid of the large amount of dead biomass immediately (Fig. 10f). While vegetation decrease and dead biomass buildup balance each other, the soil takes up about 2 Pg C (Fig. 10c), stores it for at least 16 years, and is therefore responsible for the initial net land uptake seen in Fig. 9e (see also Fig. 11c left). After about two years, tropical precipitation increases again and puts a halt to the decrease in vegetation (Fig. 10a and Fig. 11b right). The vegetation does not recover fully for another about 20 years. The dead biomass, on the other hand, gets decomposed entirely within about 15 years and therefore turns the land into a carbon source, causing the overshoot in CO_2. In the top10 case, the precipitation and radiation response is about half of the top3 case, and so is the vegetation decrease. Consequently, vegetation recovers faster. The decomposition of dead biomass, however, takes about the same amount of time as in the top3 case as the decomposition rates are similar for both cases. Hence, the land acts as a more sustainable carbon sink in the top10 case. In MPI-ESM it is the soil as well which acts as main land carbon storage pool, while the vegetation decrease is significantly less than in CESM (Brovkin et al., 2010), leading to the different response behavior of the two land models, particularly striking in the top3 case. Note that there are subtle regional differences between CESM and the earlier version of the carbon cycle-enabled NCAR model CSM1.4-carbon (Frölicher et al., 2011): tropical Africa sees a reduction of land carbon in CESM, related to a persistent increase in cloud cover and precipitation after volcanoes, while CSM1.4-carbon saw a decrease in precipitation and an increase in land carbon.

The ocean, on the other hand, shows a qualitatively similar response in CESM and MPI-ESM with an uptake of carbon and a gradual relaxation back to pre-eruption values over 20 or more years. In CESM the radiative cooling leads to increased uptake in the Western Pacific, while in the Eastern Pacific, cooling is less as this region is more controlled by upwelling rather than direct radiative forcing, as suggested by Maher et al. (2014) (Fig. 11d). Two or more years after the volcano a La Niña-like pattern settles in both surface temperature as well as carbon uptake. Some model differences exist, e.g., in the top3 case of MPI-ESM the ocean starts to release carbon, compensating the persistent positive anomaly in the land inventory (imposed on the ocean via atmospheric CO_2 concentration Brovkin et al., 2010), a feature not present in CESM, in which the land does not store the anomalous carbon as long. In CESM the tropical oceans appear to be more sensitive to volcanic forcing than to TSI variations. The equatorial Pacific shows the strongest response in DIC to volcanoes (Fig. 11d), while the response to TSI variations of comparable radiative forcing is up to an order of magnitude weaker and confined to higher latitudes (not shown). Overall it seems therefore that the response of the land vegetation governs the overall different responses in the two models.

In an attempt to validate the two models, one is restrained to the well-observed eruption of Pinatubo in 1991 CE, as the CO_2 records from ice cores do not adequately resolve short-term variations induced by volcanoes over the last millennium. Fig. 12 shows the global temperature and atmospheric carbon response to Pinatubo as extracted from observations, CESM, and the 3-member ensemble of MPI-ESM. Note that the effects of El Niño-Southern Oscillation and anthropogenic emissions have been removed from the CO_2 observations to obtain a tentative estimate of the actual CO_2 response to the Pinatubo eruption (Frölicher et al., 2013). The initial cooling of about -0.5°C and the relaxation back to initial temperatures around 1998 CE is captured well by both models. The MPI-ESM ensemble, however, shows a large and robust variation around 1995 CE, seemingly related to a phasing of ENSO variability in response to the eruption (see also Zanchettin et al., 2012). Further, the magnitude of atmospheric carbon response matches better in CESM, although the overshoot of the observation-based estimate is not captured. CESM's response also falls within the range of the earlier model version (Frölicher et al., 2013). It remains unclear whether this mismatch reflects a model-deficiency or is due to uncertainties arising from removing the ENSO signal from the CO_2 observations. However, the mechanisms described above that lead to an atmospheric CO_2 overshoot for large eruptions in CESM offer an opportunity for reconciliation of this dispcrepancy. Further, the precipitation response (and therewith the cloud and surface short-wave response) to volcanic eruptions is not well constrained due to the small number of observed eruptions (Trenberth and Dai, 2007). Biases in the representation of these processes can influence a model's carbon cycle response.

6 Climate-carbon cycle sensitivity

Due to the absence of large anthropogenic disturbances of the carbon cycle, the last millennium represents a testbed to estimate the climate-carbon cycle ~~feedback sensitivity γ (~~sensitivity, expressed as ppm °C $^{-1}$~~)~~, and can thus potentially help to constrain this quantity (e.g., Woodwell et al., 1998; Joos and Prentice, 2004; Scheffer et al., 2006; Cox and Jones, 2008; Frank et al., 2010). ~~Here, we use the experimental setup of CESM to estimate~~ Note, however, that there exist important differences between studies in how this sensitivity is calculated and what it implies. Studies using observations and fully-coupled simulations (Frank et al., 2010; Jungclaus et al., 2010) estimate the sensitivity from the ratio of changes in CO_2 over changes in temperature. This quantity folds in feebacks, as the initial sensitivity of the carbon cycle to climate change modifies itself via the climate change that arises from the changed carbon cycle. This is distinct from the climate-carbon cycle feedback sensitivity γ, ~~mimicking to some extent the methods by Frank et al. (2010) and Jungclaus et al. (2010) .~~ which uses idealized simulations with atmospheric CO_2 held constant, while the climate varies naturally, to isolate the feedback parameter (Friedlingstein et al., 2006) . The sensitivity that can be derived from our CESM transient simulation is subtly different again in that the carbon cycle will respond to changes in climate and this response will feed back on the carbon stocks through increased or decreased atmospheric CO_2 concentrations, yet these changes in CO_2 are not allowed to feed back on the climate. Such a sensitivity is expected to be lower than γ, which we can derive from CTRL.

Here, we estimate the climate-carbon cycle sensitivity for CESM as follows. We focus on the period before significant LULUC (850-1500 CE) and apply different low-pass filters of 20 to 120 years, taking 5-year increments, to the time series of NH SAT and global CO_2. The filtering aims at minimizing the influence of short-lived forcings such as volcanic eruptions that have a relatively direct impact on temperature and CO_2 (as seen above) and thus may hinder the detection of a low-frequency influence of temperature on CO_2. For each filter length we determine the highest lag correlation of the two time series, considering lags of up to 100 years. By design of our simulation we expect NH SAT to lead CO_2, which is confirmed by all lag correlations indicating positive lags for NH SAT (peak of lag correlation at 80.5±3.4 years). We regress the lagged time series and find a median estimate of 1.3 ppm $°C^{-1}$ with a range from 1.0 to 1.8 ppm $°C^{-1}$, depending on the filter length. ~~This~~ About -1 ppm $°C^{-1}$ is explained by the land carbon cycle, while the ocean shows smaller sensitivities of about -0.4 ppm $°C^{-1}$. Note, that we use NH SAT in order to be comparable with existing studies (Frank et al., 2010; Jungclaus et al., 2010) . Using global SAT instead of NH SAT can influence the sensitivity estimate, especially for the forced simulation: including the wast ocean area of the SH tends to dampen temperature variability induced by volcanoes and TSI variations. With temperature variability dampened, the sensitivity increases to 1.7 ppm $°C^{-1}$ (1.4-2.1).

This estimate is barely within the reconstruction-constrained range of 1.7-21.4 ppm $°C^{-1}$ (Frank et al., 2010) and suggests a comparably low sensitivity of the carbon cycle in CESM. This low sensitivity is in agreement with, e.g., Arora et al. (2013). Note that Frank et al. (2010) found different ~~γ~~ sensitivities for the early and late part of the last millennium with the mean for the period 1050-1549 CE being 4.3 ppm $°C^{-1}$. Indeed, a strong temporal dependence of ~~γ~~ the climate-carbon cycle sensitivity is also found in CESM when looking at individual 200-year windows (Fig. 13). The period 1300-1500 CE even shows negative ~~γ~~ sensitivity, which seems to be related to the different time scales with which SAT and CO_2 relax back to the pre-eruption conditions after perturbations from large volcanic eruptions (Fig. 9a and c): atmospheric CO_2 decreases from having overshoot while SAT increases after the initial cooling, leading to a negative correlation of the two quantities.

This illustrates the time-variant character of ~~γ~~ the climate-carbon cycle sensitivity, which substantially complicates any attempt to constrain it by last millennium data and warrants caution when making inferences from past to future sensitivities. Besides Frank et al. (2010), Frölicher et al. (2011) found ~~γ~~ the sensitivity to vary greatly in a coupled model with the time scale and magnitude of volcanic forcing considered. This issue is further highlighted by the larger ~~γ~~ sensitivity derived for idealized +1%-CO_2 $year^{-1}$ simulations with CESM (11.9 ppm $°C^{-1}$), for which a dependence on the background state, the scenario, and even the method is reported (Plattner et al., 2008; Arora et al., 2013). Further, it is worth stressing that such sensitivity estimates cannot be extrapolated easily across time scales, as different processes might be at play (Ciais et al., 2013) .

Applying the identical analysis to CTRL reveals other time scales of climate-carbon cycle feedback, suggesting maximum lags of less than 10 years and a ~~γ~~ sensitivity of 2.3 (1.4-2.9) ppm $°C^{-1}$.

Using global SAT instead of NH SAT has no discernible effect (2.3 ppm °C^{-1}), as the CTRL does not see volcanoes or TSI variations. A later peak in the lag correlation of ~~CTRL~~ NH SAT and CO_2 clusters at 73.3±1.1 years in CTRL, i.e., close to where the forced simulation shows its highest lag correlation, but these lag correlations are much weaker ($r \sim 0.4$ compared to $r \sim 0.7$ in the forced simulation). This is generally consistent with the finding by Jungclaus et al. (2010) that a forced simulation exhibits increased power on lower frequencies compared to a control simulation.

7 Discussion and conclusions

This study presents a simulation from 850 to 2100 CE with the fully-coupled CESM, including carbon cycle, and ~~investigates~~ provides an overview on the imprint of external forcing on different climate and carbon cycle diagnostics in the simulation. For comparison we draw on a number of PMIP3 simulations, particularly, comparable simulations with CCSM4 and MPI-ESM. The evolution of NH SAT during the preindustrial era in CESM is in reasonable agreement with both reconstructions and other models, albeit the uncertainties in reconstructions and forcing ~~are still~~ still being considerable. Comparing to more reliable data in the twentieth century, the anthropogenic warming in CESM is overestimated due to a lack of negative forcing from indirect aerosol effects. On the SH, CESM and most other models do not capture the evolution of the mean SAT as well. The discrepancies could be explained by (i) significant model biases in SH and also interhemispheric SAT variability (Neukom et al., 2014), (ii) spectral biases in proxies used in the reconstructions (Franke et al., 2013), (iii) uncertainties in the external forcing (Masson-Delmotte et al., 2013), or (iv) natural internal variability (Bothe et al., 2013). Unfortunately, these potential explanations are neither exclusive nor independent. Arguments for model bias come from the fact that reconstructed interhemispheric SAT variability lies outside the models' range over 40% of the time (Neukom et al., 2014); but these arguments are weakened by the uncertainty in external forcing. We show here that implementing the same TSI forcing in two different models results in a larger difference in simulated SAT than implementing two different TSI forcings in the same model. Hence, model structural uncertainty remains an issue in determining the role of external forcing over the last millennium.

Albeit beyond the scope of this study, detecting structural and spatial dependencies such as illustrated here offers an opportunity to reconcile the discrepancies (e.g., regarding SH volcanic signals) between reconstructions and simulations, which might originate from sampling bias, model deficiencies, a combination of these, or the fact that reality may be one realization by chance not encompassed by a multi-model ensemble (Deser et al., 2012; Lehner et al., 2012a; Bothe et al., 2013; Neukom et al., 2014).

Further, we compare simulations with and without orbital forcing and fail to attribute northern high latitude SAT trends over the last millennium to orbital forcing. This hampers, if not challenges, the validation of recent findings based on proxy archives that claim a distinct low-frequency orbital component in millennial trends (Kaufman et al., 2009; Esper et al., 2012). Instead, the decreasing trend in annual TSI – as opposed to seasonal and regional insolation – together with local feedbacks are able to account for a similar magnitude of trend.

When forced with emissions from LULUC, TSI variations, and volcanic eruptions over the last millennium, both CESM and MPI-ESM do not reproduce atmospheric CO_2 variability as suggested by ice cores. Notably, the large drop of CO_2 in the seventeenth century is not reproduced, similar as in earlier studies (Gerber et al., 2003; Stocker et al., 2011; Jungclaus et al., 2010). Neukom et al. (2014) hypothesized that the unique, globally synchronous cooling during the LIA (which might be related to ocean dynamics) can serve as an explanation for this drop. While both CESM and MPI-ESM show a global cooling during the LIA, they develop no apparent phasing of ocean dynamics or carbon uptake and do not show any marked CO_2 reduction around that time, leaving this issue unresolved. The strong volcanic forcing during the thirteenth century, on the other hand, is able to synchronize the AMOC on decadal scales, confirming similar results from the Bergen Climate Model and IPSL-CM5A-LR (Otterå et al., 2010; Swingedouw et al., 2013). Under anthropogenic emissions, land and ocean carbon uptake rates emerge from the envelope of natural variability as simulated for the last millennium by about 1947 CE and 1877 CE, respectively. Atmospheric CO_2 and global temperature emerge by 1755 CE and 1966 CE, suggesting that changes in carbon-cycle related variables would be easier to detect than temperature given sufficient observational data (Keller et al., 2015).

We find forced decadal-scale variability in CESM and MPI-ESM in response to major volcanic eruptions in both SAT and upper-ocean temperature, while the response in carbon cycle quantities is less coherent among models (see also Resplandy et al., 2015) . Outside volcanically active periods large parts of the decadal-scale variations cannot be attributed to external forcing, suggesting that internal variability masks external forcing influence. Note, however, that recent work suggest that small volcanic eruptions, which are typically not well-resolved in reconstructions of volcanic activity, ~~exeret~~ exert a significant cumulative effect on global temperature and climate (Ridley et al., 2014).

Volcanoes trigger a coherent global response in SAT and precipitation that is qualitatively in line with earlier studies on the volcanic influence on climate and carbon cycle (e.g., Jones and Cox, 2001; Brovkin et al., 2010; Frölicher et al., 2011, 2013). However, the carbon cycle response, in particular on land, shows fundamental model differences in terms of perturbation amplitude and persistence after volcanic eruptions. These differences arise from a differing land vegetation responses in the two models. The extent to which such structural uncertainties matter is illustrated by the large spread in the airborne fraction of CO_2 between these two (and other) models in the twenty-first century (see also Friedlingstein et al., 2014). In particular, known biases in CESM's carbon uptake in response to anthropogenic emissions in the twentieth and twenty-first century lead to a 20% overestimation of the atmospheric CO_2 concentration and the corresponding prognostic radiative forcing as compared to the prescribed RCP8.5 at year 2100 CE.

The climate-carbon cycle sensitivity of CESM as estimated from the anthropogenically unperturbed first part of the last millennium is ~~about 1.3~~between 1.0 and 2.1 ppm °C $^{-1}$, with a dependency on the filtering and the exact time period considered. Generally, the sensitivity of the carbon cycle to temperature variations in CESM is comparably small (Frank et al., 2010) and reveals a strong component of unforced natural variability. In a transient last millennium simulation with small temperature variations, the proper detection of a lead-lag relation between temperature and the carbon cycle is complicated by the superposition of perturbations and responses. In addition to the classic climate-carbon cycle sensitivity experiments (e.g., Arora et al., 2013) it is therefore desirable to conduct step function-like sensitivity experiments in order to isolate the response of the carbon cycle to a particular external forcing (Gerber et al., 2003).

Despite the challenges that paleoclimate modelling faces, a number of lessons regarding forcing and structural uncertainties can be learned from these experiments. In order to better understand the role of internal versus externally-forced variability – which remains particularly critical for a period of relatively weak external forcing, such as the last millennium – larger simulation ensembles ~~and~~ as well as ensembles with decomposed forcing should become a standard procedure in paleoclimate modelling. Since these are computationally expensive simulations, this calls for an informed discussion on the optimal usage of computing resources, to which studies like the one here can contribute valuable information. At the same time, uncertainties in forcings and ~~reconstruction~~ reconstructions need to be further reduced to be able to better validate models in the past with the goal of constraining their future response. Key targets for such constrains ~~remain~~ are the sensitivity of temperature to solar and volcanic forcing and the climate-carbon cycle sensitivity.

Appendix A

A1

Acknowledgements. We gratefully acknowledge Axel Timmermann, Bette Otto-Bliesner, Peter Lawrence, and Rosie Fisher for valuable discussions as well as four anonymous reviewers for very helpful comments. We are grateful to the NCAR in Boulder, USA, for providing the code of the CESM, to the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, to the climate modeling groups for producing and making available their model output. This study is supported by the Swiss National Science Foundation (grant no. 200020 147174), the European Commission through Seventh Framework Program (FP7) projects CARBOCHANGE (grant no. 264879) and Past4Future (grant no. 243908). J.M. has benefited from the support of the French Agence Nationale de la Recherche (HAMOC : ANR 13-BLAN-06-0003). The simulations for this study were performed on a CRAY XT5 and XE6 at the Swiss National Supercomputing Centre (CSCS) in Lugano.

References

Abram, N. J., Mulvaney, R., Vimeux, F., Phipps, S. J., Turner, J., and England, M. H.: Evolution of the Southern Annular Mode during the past millennium, Nature Climate Change, 4, 564–569, 2014.

Ammann, C. M., Joos, F., Schimel, D., Otto-Bliesner, B. L., and Tomas, R.: Solar influence on climate during the past millennium: Results from transient simulations with the NCAR Climate System Model, Proc. Natl. Acad. Sci., 104, 3713–3718, 2007.

Anchukaitis, K. J., Breitenmoser, P., Briffa, K. R., Buchwal, A., Buentgen, U., Cook, E. R., D'Arrigo, R. D., Esper, J., Evans, M. N., Frank, D., Grudd, H., Gunnarson, B. E., Hughes, M. K., Kirdyanov, A. V., Koerner, C., Krusic, P. J., Luckman, B., Melvin, T. M., Salzer, M. W., Shashkin, A. V., Timmreck, C., Vaganov, E. A., and Wilson, R. J. S.: Tree rings and volcanic cooling, Nature Geoscience, 5, 836–837, doi:10.1038/ngeo1645, 2012.

Andres, R. J., Boden, T. A., Breon, F. M., Ciais, P., Davis, S., Erickson, D., Gregg, J. S., Jacobson, A., Marland, G., Miller, J., Oda, T., Olivier, J. G. J., Raupach, M. R., Rayner, P., and Treanton, K.: A synthesis of carbon dioxide emissions from fossil-fuel combustion, Biogeosciences, 9, 1845–1871, doi:10.5194/bg-9-1845-2012, 2012.

Armstrong, R., Lee, C., Hedges, J., Honjo, S., and Wakeham, S.: A new, mechanistic model for organic carbon fluxes in the ocean based on the quantitative association of POC with ballast minerals, Deep Sea Res. Part II, 49, 219–236, 2002.

Arora, V. K., Boer, G. J., Friedlingstein, P., Eby, M., Jones, C. D., Christian, J. R., Bonan, G., Bopp, L., Brovkin, V., Cadule, P., Hajima, T., Ilyina, T., Lindsay, K., Tjiputra, J. F., and Wu, T.: Carbon-Concentration and Carbon-Climate Feedbacks in CMIP5 Earth System Models, J. Clim., 26, 5289–5314, doi:10.1175/JCLI-D-12-00494.1, 2013.

Berger, A. L.: Long-term variations of caloric insolation resulting from the Earth's orbital elements, J. Atm. Sci., 35, 2362–2367, doi:10.1016/0033-5894(78)90064-9, 1978.

Bothe, O., Jungclaus, J. H., and Zanchettin, D.: Consistency of the multi-model CMIP5/PMIP3-past1000 ensemble, Clim. Past, 9, 2471–2487, doi:10.5194/cp-9-2471-2013, 2013.

Bretagnon, P. and Francou, G.: Planetary theories in rectangular and spherical variables - VSOP-87 solutions, Astronomy & Astrophysics, 202, 309–315, 1988.

Brovkin, V., Lorenz, S. J., Jungclaus, J., Raddatz, T., Timmreck, C., Reick, C. H., Segschneider, J., and Six, K.: Sensitivity of a coupled climate-carbon cycle model to large volcanic eruptions during the last millennium, 62, 674–681, doi:10.1111/j.1600-0889.2010.00471.x, 2010.

Chikamoto, M. O., Timmermann, A., Yoshimori, M., Lehner, F., Laurian, A., Abe-Ouchi, A., Mouchet, A., Joos, F., and Cobb, K.: Delayed tropical Pacific biological productivity response to strong volcanic forcing, Geophys. Res. Let., submitted.

Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., Chhabra, A., DeFries, R., Galloway, J., Heimann, M., Jones, C., Le Quere, C., Myneni, R.B. Piao, S., and Thornton, P.: Carbon and Other Biogeochemical Cycles, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Cowtan, K. and Way, R. G.: Coverage bias in the HadCRUT4 temperature series and its impact on recent temperature trends, Quarterly J. Royal Meteor. Soc., 140, 1935–1944, doi:10.1002/qj.2297, 2014.

Cox, P. and Jones, C.: Climate change - illuminating the modern dance of climate and CO2, Science, 321, 1642–1644, doi:10.1126/science.1158907, 2008.

Cox, P. M., Pearson, D., Booth, B. B., Friedlingstein, P., Huntingford, C., Jones, C. D., and Luke, C. M.: Sensitivity of tropical carbon to climate change constrained by carbon dioxide variability, Nature, 494, 341–344, doi:10.1038/nature11882, 2013.

Crowley, T. J.: Causes of climate change over the past 1000 years, Science, 289, 270–277, 2000.

Crowley, T. J., Zielinski, G., Vinther, B., Udisti, R., Kreutzs, K., Cole-Dai, J., and Castellano, E.: Volcanism and the Little Ice Age, PAGES newsletter, 16, 22–23, 2008.

Danabasoglu, G., Bates, S. C., Briegleb, B. P., Jayne, S. R., Jochum, M., Large, W. G., Peacock, S., and Yeager, S. G.: The CCSM4 ocean component, J. Clim., 25, 1361–1389, doi:10.1175/JCLI-D-11-00091.1, 2012.

Deser, C., Knutti, R., Solomon, S., and Phillips, A. S.: Communication of the role of natural variability in future North American climate, Nature Climate Change, 2, 775–779, doi:10.1038/NCLIMATE1562, 2012.

Doney, S. C., Lindsay, K., Fung, I., and John, J.: Natural variability in a stable, 1000-yr global coupled climate-carbon cycle simulation, J. Clim., 19, 3033–3054, doi:10.1175/JCLI3783.1, 2006.

Esper, J., Frank, D. C., Timonen, M., Zorita, E., Wilson, R. J. S., Luterbacher, J., Holzkaemper, S., Fischer, N., Wagner, S., Nievergelt, D., Verstege, A., and Buentgen, U.: Orbital forcing of tree-ring data, Nature Climate Change, 2, 862–866, doi:10.1038/NCLIMATE1589, 2012.

Fernandez-Donado, L., Gonzalez-Rouco, J. F., Raible, C. C., Ammann, C. M., Barriopedro, D., Garcia-Bustamante, E., Jungclaus, J. H., Lorenz, S. J., Luterbacher, J., Phipps, S. J., Servonnat, J., Swingedouw, D., Tett, S. F. B., Wagner, S., Yiou, P., and Zorita, E.: Large-scale temperature response to external forcing in simulations and reconstructions of the last millennium, Clim. Past, 9, 393–421, doi:10.5194/cp-9-393-2013, 2013.

Frank, D. C., Esper, J., Raible, C. C., Buentgen, U., Trouet, V., Stocker, B., and Joos, F.: Ensemble reconstruction constraints on the global carbon cycle sensitivity to climate, Nature, 463, 527–532, doi:10.1038/nature08769, 2010.

Franke, J., Frank, D., Raible, C. C., Esper, J., and Broennimann, S.: Spectral biases in tree-ring climate proxies, Nature Climate Change, 3, 360–364, doi:10.1038/NCLIMATE1816, 2013.

Friedlingstein, P., Meinshausen, M., Arora, V. K., Jones, C. D., Anav, A., Liddicoat, S. K., and Knutti, R.: Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks, J. Clim., 27, 511–526, doi:10.1175/JCLI-D-12-00579.1, 2014.

Friedlingstein, P., Cox, P., Betts, R., Bopp, L., Von Bloh, W., Brovkin, V., Cadule, P., Doney, S., Eby, M., Fung, I., Bala, G., John, J., Jones, C., Joos, F., Kato, T., Kawamiya, M., Knorr, W., Lindsay, K., Matthews, H. D., Raddatz, T., Rayner, P., Reick, C., Roeckner, E., Schnitzler, K. G., Schnur, R., Strassmann, K., Weaver, A. J., Yoshikawa, C., and Zeng, N.: Climate-carbon cycle feedback analysis: Results from the C^4MIP model intercomparison, J. Clim., 19, 3337–3353, doi:10.1175/JCLI3800.1, 2006.

Friedrich, T., Timmermann, A., Abe-Ouchi, A., Bates, N. R., Chikamoto, M. O., Church, M. J., Dore, J. E., Gledhill, D. K., Gonzalez-Davila, M., Heinemann, M., Ilyina, T., Jungclaus, J. H., McLeod, E., Mouchet, A., and Santana-Casiano, J. M.: Detecting regional anthropogenic trends in ocean acidification against natural variability, Nature Climate Change, 2, 167–171, doi:10.1038/NCLIMATE1372, 2012.

Frölicher, T. L., Joos, F., and Raible, C. C.: Sensitivity of atmospheric CO2 and climate to explosive volcanic eruptions, Biogeosciences, 8, 2317–2339, doi:10.5194/bg-8-2317-2011, 2011.

Frölicher, T. L., Joos, F., Raible, C. C., and Sarmiento, J. L.: Atmospheric CO2 response to volcanic eruptions: The role of ENSO, season, and variability, Global Biogeochem. Cyc., 27, 239–251, doi:10.1002/gbc.20028, 2013.

Gao, C., Robock, A., and Ammann, C.: Volcanic forcing of climate over the past 1500 years: An improved ice core-based index for climate models, J. Geophys. Res., 113, doi:10.1029/2008JD010239, 2008.

Gent, P. R., Danabasoglu, G., Donner, L. J., Holland, M. M., Hunke, E. C., Jayne, S. R., Lawrence, D. M., Neale, R. B., Rasch, P. J., Vertenstein, M., Worley, P. H., Yang, Z.-L., and Zhang, M.: The Community Climate System Model Version 4, J. Clim., 24, 4973–4991, doi:10.1175/2011JCLI4083.1, 2011.

Geoffroy, O., Saint-Martin, D., and Voldoire, A.: Land-sea warming contrast: the role of the horizontal energy transport, Clim. Dyn., pp. 1–19, doi:10.1007/s00382-015-2552-y, 2015.

Gerber, S., Joos, F., Brugger, P., Stocker, T., Mann, M., Sitch, S., and Scholze, M.: Constraining temperature variations over the last millennium by comparing simulated and observed atmospheric CO2, Clim. Dyn., 20, 281–299, doi:10.1007/s00382-002-0270-8, 2003.

Hawkins, E. and Sutton, R.: Time of emergence of climate signals, Geophys. Res. Let., 39, doi:10.1029/2011GL050087, 2012.

Hoffman, F. M., Randerson, J. T., Arora, V. K., Bao, Q., Cadule, P., Ji, D., Jones, C. D., Kawamiya, M., Khatiwala, S., Lindsay, K., Obata, A., Shevliakova, E., Six, K. D., Tjiputra, J. F., Volodin, E. M., and Wu, T.: Causes and implications of persistent atmospheric carbon dioxide biases in Earth System Models, J. Geophys. Res., 119, 141–162, doi:10.1002/2013JG002381, 2014.

Holland, M. M. and Bitz, C. M.: Polar amplification of climate change in coupled models, Clim. Dyn., 21, 221–232, 2003.

Houghton, R. A.: How well do we know the flux of CO2 from land-use change?, Tellus, 62, 337–351, doi:10.1111/j.1600-0889.2010.00473.x, 2010.

Hunke, E. C. and Lipscomb, W. H.: CICE: the Los Alamos sea ice model documentation and software user's manual version 4.1, Tech. rep., Los Alamos National Laboratory (LANL), 76 pp., 2010.

Hurrell, J. W., Holland, M. M., Gent, P. R., Ghan, S., Kay, J. E., Kushner, P. J., Lamarque, J. F., Large, W. G., Lawrence, D., Lindsay, K., Lipscomb, W. H., Long, M. C., Mahowald, N., Marsh, D. R., Neale, R. B., Rasch, P., Vavrus, S., Vertenstein, M., Bader, D., Collins, W. D., Hack, J. J., Kiehl, J., and Marshall, S.: The Community Earth System Model: A Framework for Collaborative Research, Bull. Am. Meteorol. Soc., 94, 1339–1360, doi:10.1175/BAMS-D-12-00121.1, 2013.

Hurtt, G. C., Chini, L. P., Frolking, S., Betts, R. A., Feddema, J., Fischer, G., Fisk, J. P., Hibbard, K., Houghton, R. A., Janetos, A., Jones, C. D., Kindermann, G., Kinoshita, T., Goldewijk, K. K., Riahi, K., Shevliakova, E., Smith, S., Stehfest, E., Thomson, A., Thornton, P., van Vuuren, D. P., and Wang, Y. P.: Harmonization of land-use scenarios for the period 1500-2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands, Clim. Change, 109, 117–161, doi:10.1007/s10584-011-0153-2, 2011.

IPCC: Special Report on Emissions Scenarios (SRES), Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2000.

IPCC: Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2001.

IPCC: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Jones, C. and Cox, P.: Modeling the volcanic signal in the atmospheric CO2 record, Global Biogeochem. Cyc., 15, 453–465, doi:10.1029/2000GB001281, 2001.

Joos, F. and Prentice, I. C.: A paleo-perspective on changes in atmospheric CO_2 and climate, Island Press, 2004.

Joos, F., Roth, R., Fuglestvedt, J. S., Peters, G. P., Enting, I. G., von Bloh, W., Brovkin, V., Burke, E. J., Eby, M., Edwards, N. R., Friedrich, T., Froelicher, T. L., Halloran, P. R., Holden, P. B., Jones, C., Kleinen, T., Mackenzie, F. T., Matsumoto, K., Meinshausen, M., Plattner, G. K., Reisinger, A., Segschneider, J., Shaffer, G., Steinacher, M., Strassmann, K., Tanaka, K., Timmermann, A., and Weaver, A. J.: Carbon dioxide and climate impulse response functions for the computation of greenhouse gas metrics: a multi-model analysis, Atmos. Chem. Phys., 13, 2793–2825, doi:10.5194/acp-13-2793-2013, 2013.

Jungclaus, J. H., Lorenz, S. J., Timmreck, C., Reick, C. H., Brovkin, V., Six, K., Segschneider, J., Giorgetta, M. A., Crowley, T. J., Pongratz, J., Krivova, N. A., Vieira, L. E., Solanki, S. K., Klocke, D., Botzet, M., Esch, M., Gayler, V., Haak, H., Raddatz, T. J., Roeckner, E., Schnur, R., Widmann, H., Claussen, M., Stevens, B., and Marotzke, J.: Climate and carbon-cycle variability over the last millennium, Clim. Past, 6, 723–737, doi:10.5194/cp-6-723-2010, 2010.

Kaplan, J. O., Krumhardt, K. M., Ellis, E. C., Ruddiman, W. F., Lemmen, C., and Goldewijk, K. K.: Holocene carbon emissions as a result of anthropogenic land cover change, The Holocene, 21, 775–791, doi:10.1177/0959683610386983, 2011.

Kaufman, D. S., Schneider, D. P., McKay, N. P., Ammann, C. M., Bradley, R. S., Briffa, K. R., Miller, G. H., Otto-Bliesner, B. L., Overpeck, J. T., Vinther, B. M., and 2k Project Members, A. L.: Recent warming reverses long-term Arctic cooling, Science, 325, 1236–1239, doi:10.1126/science.1173983, 2009.

Keller, K. M., Joos, F., Lehner, F., and Raible, C. C.: Detecting changes in marine responses to ENSO 850-2100 CE: Insights from the ocean carbon cycle, Geophys. Res. Let., doi:10.1002/2014GL062398, 2015.

Keller, K. M., Joos, F., and Raible, C. C.: Time of emergence of trends in ocean biogeochemistry, Biogeosciences, 11, 3647–3659, doi:10.5194/bg-11-3647-2014, 2014.

Keppel-Aleks, G., Randerson, J. T., Lindsay, K., Stephens, B. B., Keith Moore, J., Doney, S. C., Thornton, P. E., Mahowald, N. M., Hoffman, F. M., Sweeney, C., Tans, P. P., Wennberg, P. O., and Wofsy, S. C.: Atmospheric carbon dioxide variability in the Community Earth System Model: Evaluation and transient dynamics during the twentieth and twenty-first centuries, J. Clim., 26, 4447–4475, doi:10.1175/JCLI-D-12-00589.1, 2013.

Key, R., Kozyr, A., Sabine, C., Lee, K., Wanninkhof, R., Bullister, J., Feely, R., Millero, F., Mordy, C., and Peng, T.: A global ocean carbon climatology: Results from Global Data Analysis Project (GLODAP), Global Biogeochem. Cyc., 18, doi:10.1029/2004GB002247, 2004.

Krivova, N. A., Balmaceda, L., and Solanki, S. K.: Reconstruction of solar total irradiance since 1700 from the surface magnetic flux, Astronomy & Astrophysics, 467, 335–346, doi:10.1051/0004-6361:20066725, 2007.

Lamarque, J. F., Bond, T. C., Eyring, V., Granier, C., Heil, A., Klimont, Z., Lee, D., Liousse, C., Mieville, A., Owen, B., Schultz, M. G., Shindell, D., Smith, S. J., Stehfest, E., Van Aardenne, J., Cooper, O. R., Kainuma, M., Mahowald, N., McConnell, J. R., Naik, V., Riahi, K., and van Vuuren, D. P.: Historical (1850-2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: Methodology and application, Atmospheric Chemistry and Physics, 10, 7017–7039, doi:10.5194/acp-10-7017-2010, 2010.

Lamarque, J.-F., Kyle, G. P., Meinshausen, M., Riahi, K., Smith, S. J., van Vuuren, D. P., Conley, A. J., and Vitt, F.: Global and regional evolution of short-lived radiatively-active gases and aerosols in the Representative Concentration Pathways, Clim. Change, 109, 191–212, doi:10.1007/s10584-011-0155-0, 2011.

Landrum, L., Otto-Bliesner, B. L., Wahl, E. R., Conley, A., Lawrence, P. J., Rosenbloom, N., and Teng, H.: Last millennium climate and its variability in CCSM4, J. Clim., 26, 1085–1111, doi:10.1175/JCLI-D-11-00326.1, 2013.

Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., Zeng, X., Yang, Z.-L., Levis, S., Sakaguchi, K., Bonan, G. B., and Slater, A. G.: Parameterization improvements and functional and structural advances in version 4 of the Community Land Model, J. of Adv. in Modeling Earth Systems, 3, doi:10.1029/2011MS000045, 2011.

Lawrence, P. J., Feddema, J. J., Bonan, G. B., Meehl, G. A., O'Neill, B. C., Oleson, K. W., Levis, S., Lawrence, D. M., Kluzek, E., Lindsay, K., and Thornton, P. E.: Simulating the biogeochemical and biogeophysical impacts of transient land cover change and wood harvest in the Community Climate System Model (CCSM4) from 1850 to 2100, J. Clim., 25, 3071–3095, doi:10.1175/JCLI-D-11-00256.1, 2012.

Le Quéré, C., Andres, R. J., Boden, T., Conway, T., Houghton, R., House, J. I., Marland, G., Peters, G. P., van der Werf, G., Ahlström, A., et al.: The global carbon budget 1959–2011, Earth System Science Data, 5, 165–185, doi:10.5194/essd-5-165-2013, 2013.

Lean, J., Rottman, G., Harder, J., and Kopp, G.: SORCE contributions to new understanding of global change and solar variability, Solar Physics, 230, 27–53, doi:10.1007/s11207-005-1527-2, 2005.

Lehner, F., Raible, C. C., and Stocker, T. F.: Testing the robustness of a precipitation proxy-based North Atlantic Oscillation reconstruction, Quat. Sci. Rev., 45, 85–94, doi:10.1016/j.quascirev.2012.04.025, 2012a.

Lehner, F., Raible, C. C., Stocker, T. F., and Hofer, D.: The freshwater balance of polar regions in transient simulations from 1500 to 2100 AD using a comprehensive coupled climate model, Clim. Dyn., 39, 347–363, doi:10.1007/s00382-011-1199-6, 2012b.

Lehner, F., Born, A., Raible, C. C., and Stocker, T. F.: Amplified inception of European Little Ice Age by sea ice-ocean-atmosphere feedbacks, J. Clim., 26, 7586–7602, doi:10.1175/JCLI-D-12-00690.1, 2013.

Lindsay, K., Bonan, G. B., Doney, S. C., Hoffman, F. M., Lawrence, D. M., Long, M. C., Mahowald, N. M., Moore, J. K., Randerson, J. T., and Thornton, P. E.: Preindustrial-Control and Twentieth-Century Carbon Cycle Experiments with the Earth System Model CESM1(BGC), J. Clim., 27, 8981–9005, doi:10.1175/JCLI-D-12-00565.1, 2014.

Long, M. C., Lindsay, K., Peacock, S., Moore, J. K., and Doney, S. C.: Twentieth-century oceanic carbon uptake and storage in CESM1(BGC), J. Clim., 26, 6775–6800, doi:10.1175/JCLI-D-12-00184.1, 2013.

Maher, N., Sen Gupta, A., and England, M. H.: Drivers of decadal hiatus periods in the 20th and 21st centuries, Geophys. Res. Let., 41, 5978–5986, doi:10.1002/2014GL060527, 2014.

Mann, E. M., Zhang, Z., Rutherford, S., Bradley, R. S., Hughes, M. K., Shindell, D., Ammann, C., Faluvegi, G., and Ni, F.: Global signatures and dynamical origins of the Little Ice Age and Medieval Climate Anomaly, Science, 326, 1256–1260, doi:10.1126/science.1177303, 2009.

Mann, M. E., Fuentes, J. D., and Rutherford, S.: Underestimation of volcanic cooling in tree-ring-based reconstructions of hemispheric temperatures, Nature Geoscience, 5, 202–205, doi:10.1038/NGEO1394, 2012.

Marsland, S., Haak, H., Jungclaus, J., Latif, M., and Roske, F.: The Max-Planck-Institute global ocean/sea ice model with orthogonal curvilinear coordinates, Ocean Modelling, 5, 91–127, doi:10.1016/S1463-5003(02)00015-X, 2003.

Masson-Delmotte, V., Schulz, M., Abe-Ouchi, A., Beer, J., Ganopolski, A., González Rouco, J., Jansen, E., Lambeck, K., Luterbacher, J., Naish, T., Osborn, T., Otto-Bliesner, B., Quinn, T., Ramesh, R., Rojas, M., Shao, X., and Timmermann, A.: Information from Paleoclimate Archives, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Matsumoto, K., Sarmiento, J., Key, R., Aumont, O., Bullister, J., Caldeira, K., Campin, J., Doney, S., Drange, H., Dutay, J., Follows, M., Gao, Y., Gnanadesikan, A., Gruber, N., Ishida, A., Joos, F., Lindsay, K., Maier-Reimer, E., Marshall, J., Matear, R., Monfray, P., Mouchet, A., Najjar, R., Plattner, G., Schlitzer, R., Slater, R., Swathi, P., Totterdell, I., Weirig, M., Yamanaka, Y., Yool, A., and Orr, J.: Evaluation of ocean carbon cycle models with data-based metrics, Geophys. Res. Let., 31, doi:10.1029/2003GL018970, 2004.

Meehl, G. A., Washington, W. M., Arblaster, J. M., Hu, A., Teng, H., Tebaldi, C., Sanderson, B. N., Lamarque, J.-F., Conley, A., Strand, W. G., and White, III, J. B.: Climate system response to external forcings and climate change projections in CCSM4, J. Clim., 25, 3661–3683, doi:10.1175/JCLI-D-11-00240.1, 2012.

Moffa-Sanchez, P., Born, A., Hall, I. R., Thornalley, D. J. R., and Barker, S.: Solar forcing of North Atlantic surface temperature and salinity over the past millennium, Nature Geoscience, 7, 275–278, doi:10.1038/ngeo2094, 2014.

Moore, J., Doney, S., and Lindsay, K.: Upper ocean ecosystem dynamics and iron cycling in a global three-dimensional model, Global Biogeochem. Cyc., 18, doi:10.1029/2004GB002220, 2004.

Moore, J. K. and Braucher, O.: Sedimentary and mineral dust sources of dissolved iron to the world ocean, Biogeosciences, 5, 631–656, 2008.

Moss, R. H., Edmonds, J. A., Hibbard, K. A., Manning, M. R., Rose, S. K., van Vuuren, D. P., Carter, T. R., Emori, S., Kainuma, M., Kram, T., Meehl, G. A., Mitchell, J. F. B., Nakicenovic, N., Riahi, K., Smith, S. J., Stouffer, R. J., Thomson, A. M., Weyant, J. P., and Wilbanks, T. J.: The next generation of scenarios for climate change research and assessment, Nature, 463, 747–756, doi:10.1038/nature08823, 2010.

Neale, R. B., Richter, J. H., Conley, A. J., Park, S., Lauritzen, P. H., Gettelman, A., Williamson, D. L., Rasch, P. J., Vavrus, S. J., Taylor, M. A., Collins, W. D., Zhang, M., and Lin, S.-J.: Description of the NCAR Community Atmosphere Model (CAM 4.0), Tech. rep., National Center for Atmospheric Research (NCAR), 212 pp., 2010.

Neukom, R., Gergis, J., Karoly, D. J., Wanner, H., Curran, M., Elbert, J., Gonzalez-Rouco, F., Linsley, B. K., Moy, A. D., Mundo, I., Raible, C. C., Steig, E. J., van Ommen, T., Vance, T., Villalba, R., Zinke, J., and Frank, D.: Inter-hemispheric temperature variability over the past millennium, Nature Climate Change, 4, 362–367, doi:10.1038/NCLIMATE2174, 2014.

Otterå, O. H., Bentsen, M., Drange, H., and Suo, L.: External forcing as a metronome for Atlantic multidecadal variability, Nature Geoscience, 3, 688–694, doi:10.1038/NGEO955, 2010.

PAGES 2k network: Continental-scale temperature variability during the past two millennia, Nature Geoscience, 6, 339–346, 2013.

Plattner, G. K., Knutti, R., Joos, F., Stocker, T. F., von Bloh, W., Brovkin, V., Cameron, D., Driesschaert, E., Dutkiewicz, S., Eby, M., Edwards, N. R., Fichefet, T., Hargreaves, J. C., Jones, C. D., Loutre, M. F., Matthews, H. D., Mouchet, A., Mueller, S. A., Nawrath, S., Price, A., Sokolov, A., Strassmann, K. M., and Weaver, A. J.: Long-term climate commitments projected with climate-carbon cycle models, J. Clim., 21, 2721–2751, doi:10.1175/2007JCLI1905.1, 2008.

Pongratz, J., Reick, C., Raddatz, T., and Claussen, M.: A reconstruction of global agricultural areas and land cover for the last millennium, Global Biogeochem. Cyc., 22, doi:10.1029/2007GB003153, 2008.

Pongratz, J., Caldeira, K., Reick, C. H., and Claussen, M.: Coupled climate-carbon simulations indicate minor global effects of wars and epidemics on atmospheric CO2 between ad 800 and 1850, The Holocene, 21, 843–851, doi:10.1177/0959683610386981, 2011.

Resplandy, L., Séférian, R., and Bopp, L.: Natural variability of CO2 and O2 fluxes: What can we learn from centuries-long climate models simulations?, J. Geophys. Res., doi:10.1002/2014JC010463, 2015.

Ridley, D. A., Solomon, S., Barnes, J. E., Burlakov, V. D., Deshler, T., Dolgii, S. I., Herber, A. B., Nagai, T., Neely, III, R. R., Nevzorov, A. V., Ritter, C., Sakai, T., Santer, B. D., Sato, M., Schmidt, A., Uchino, O., and Vernier, J. P.: Total volcanic stratospheric aerosol optical depths and implications for global climate change, Geophys. Res. Let., 41, 7763–7769, doi:10.1002/2014GL061541, 2014.

Rothenberg, D., Mahowald, N., Lindsay, K., Doney, S. C., Moore, J. K., and Thornton, P.: Volcano impacts on climate and biogeochemistry in a coupled carbon-climate model, Earth Sys. Dyn., 3, 121–136, doi:10.5194/esd-3-121-2012, 2012.

Scheffer, M., Brovkin, V., and Cox, P.: Positive feedback between global warming and atmospheric CO2 concentration inferred from past climate change, Geophys. Res. Let., 33, doi:10.1029/2005GL025044, 2006.

Schmidt, G. A., Jungclaus, J. H., Ammann, C. M., Bard, E., Braconnot, P., Crowley, T. J., Delaygue, G., Joos, F., Krivova, N. A., Muscheler, R., Otto-Bliesner, B. L., Pongratz, J., Shindell, D. T., Solanki, S. K., Steinhilber, F., and Vieira, L. E. A.: Climate forcing reconstructions for use in PMIP simulations of the last millennium (v1.0), Geoscientific Model Development, 4, 33–45, doi:10.5194/gmd-4-33-2011, 2011.

Schmidt, G. A., Jungclaus, J. H., Ammann, C. M., Bard, E., Braconnot, P., Crowley, T. J., Delaygue, G., Joos, F., Krivova, N. A., Muscheler, R., Otto-Bliesner, B. L., Pongratz, J., Shindell, D. T., Solanki, S. K., Steinhilber, F., and Vieira, L. E. A.: Climate forcing reconstructions for use in PMIP simulations of the Last Millennium (v1.1), Geoscientific Model Development, 5, 185–191, doi:10.5194/gmd-5-185-2012, 2012.

Schmidt, G. A., Annan, J. D., Bartlein, P. J., Cook, B. I., Guilyardi, E., Hargreaves, J. C., Harrison, S. P., Kageyama, M., LeGrande, A. N., Konecky, B., Lovejoy, S., Mann, M. E., Masson-Delmotte, V., Risi, C., Thompson, D., Timmermann, A., Tremblay, L. B., and Yiou, P.: Using palaeo-climate comparisons to constrain future projections in CMIP5, Clim. Past, 10, 221–250, doi:10.5194/cp-10-221-2014, 2014.

Schneider, D. P., Ammann, C. M., Otto-Bliesner, B. L., and Kaufman, D. S.: Climate response to large, high-latitude and low-latitude volcanic eruptions in the Community Climate System Model, J. Geophys. Res., 114, doi:10.1029/2008JD011222, 2009.

Schurer, A. P., Tett, S. F., and Hegerl, G. C.: Small influence of solar variability on climate over the past millennium, Nature Geoscience, 7, 104–108, 2014.

Schurer, A. P., Hegerl, G. C., Mann, M. E., Tett, S. F. B., and Phipps, S. J.: Separating forced from chaotic climate variability over the past millennium, J. Clim., 26, 6954–6973, doi:10.1175/JCLI-D-12-00826.1, 2013.

Schwinger, J., Tjiputra, J. F., Heinze, C., Bopp, L., Christian, J. R., Gehlen, M., Ilyina, T., Jones, C. D., Salas-Melia, D., Segschneider, J., Seferian, R., and Totterdell, I.: Nonlinearity of Ocean Carbon Cycle Feedbacks in CMIP5 Earth System Models, jc, 27, 3869–3888, doi:10.1175/JCLI-D-13-00452.1, 2014.

Servonnat, J., Yiou, P., Khodri, M., Swingedouw, D., and Denvil, S.: Influence of solar variability, CO2 and orbital forcing between 1000 and 1850 AD in the IPSLCM4 model, Clim. Past, 6, 445–460, doi:10.5194/cp-6-445-2010, 2010.

Shapiro, A. I., Schmutz, W., Rozanov, E., Schoell, M., Haberreiter, M., Shapiro, A. V., and Nyeki, S.: A new approach to the long-term reconstruction of the solar irradiance leads to large historical solar forcing, Astronomy & Astrophysics, 529, doi:10.1051/0004-6361/201016173, 2011.

Shevliakova, E., Pacala, S. W., Malyshev, S., Hurtt, G. C., Milly, P. C. D., Caspersen, J. P., Sentman, L. T., Fisk, J. P., Wirth, C., and Crevoisier, C.: Carbon cycling under 300 years of land use change: Importance of the secondary vegetation sink, Global Biogeochem. Cyc., 23, doi:10.1029/2007GB003176, 2009.

Sicre, M. A., Khodri, M., Mignot, J., Eiriksson, J., Knudsen, K. L., Ezat, U., Closset, I., Nogues, P., and Masse, G.: Sea surface temperature and sea ice variability in the subpolar North Atlantic from explosive volcanism of the late thirteenth century, Geophys. Res. Let., 40, 5526–5530, doi:10.1002/2013GL057282, 2013.

Smith, R., Jones, P., Briegleb, B., , Bryan, F., Danabasoglu, G., Dennis, J., Dukowicz, J., Eden, C., Fox-Kemper, B., Gent, P., Hecht, M., Jayne, S., Large, M. J. W., Lindsay, K., Maltrud, M., Norton, N., Peacock, S., Vertenstein, M., and Yeager, S.: The Parallel Ocean Program (POP) Reference Manual, Tech. rep., Los Alamos National Laboratory (LANL), 140 pp., 2010.

Steinacher, M., Joos, F., and Stocker, T. F.: Allowable carbon emissions lowered by multiple climate targets, Nature, 499, 197–203, doi:10.1038/nature12269, 2013.

Stocker, B. D., Strassmann, K., and Joos, F.: Sensitivity of Holocene atmospheric CO2 and the modern carbon budget to early human land use: Analyses with a process-based model, Biogeosciences, 8, 69–88, doi:10.5194/bg-8-69-2011, 2011.

Stocker, B. D., Feissli, F., Strassmann, K. M., Spahni, R., and Joos, F.: Past and future carbon fluxes from land use change, shifting cultivation and wood harvest, Tellus, 66, doi:10.3402/tellusb.v66.23188, 2014.

Swingedouw, D., Mignot, J., Labetoulle, S., Guilyardi, E., and Madec, G.: Initialisation and predictability of the AMOC over the last 50 years in a climate model, Clim. Dyn., 40, 2381–2399, doi:10.1007/s00382-012-1516-8, 2013.

Swingedouw, D., Ortega, P., Mignot, J., Guilyardi, E., Masson-Delmotte, V., Butler, P. G., Khodri, M., and Seferian, R.: Bidecadal North Atlantic ocean circulation variability controlled by timing of volcanic eruptions, Nature Communications, 6, doi:10.1038/ncomms7545, 2015.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An overview of CMIP5 and the experiment design, Bull. Am. Meteorol. Soc., 93, 485–498, doi:10.1175/BAMS-D-11-00094.1, 2012.

Thompson, D. W. J., Wallace, J. M., Jones, P. D., and Kennedy, J. J.: Identifying signatures of natural climate variability in time series of global-mean surface temperature: Methodology and insights, J. Clim., 22, 6120–6141, doi:10.1175/2009JCLI3089.1, 2009.

Timmreck, C., Graf, H.-F., Lorenz, S. J., Niemeier, U., Zanchettin, D., Matei, D., Jungclaus, J. H., and Crowley, T. J.: Aerosol size confines climate response to volcanic super-eruptions, Geophys. Res. Let., 37, doi:10.1029/2010GL045464, 2010.

Tingley, M. P., Stine, A. R., and Huybers, P.: Temperature reconstructions from tree-ring densities overestimate volcanic cooling, Geophys. Res. Let., 41, 7838–7845, doi:10.1002/2014GL061268, 2014.

Tjiputra, J. F. and Otter, O. H.: Role of volcanic forcing on future global carbon cycle, Earth Sys. Dyn., 2, 53–67, doi:10.5194/esd-2-53-2011, 2011.

Trenberth, K. E. and Dai, A.: Effects of Mount Pinatubo volcanic eruption on the hydrological cycle as an analog of geoengineering, Geophys. Res. Let., 34, doi:10.1029/2007GL030524, 2007.

Tschumi, J. and Stauffer, B.: Reconstructing past atmospheric CO2 concentration based on ice-core analyses: open questions due to in situ production of CO2 in the ice, J. Glaciology, 46, 45–53, doi:10.3189/172756500781833359, 2000.

Vieira, L. E. A. and Solanki, S. K.: Evolution of the solar magnetic flux on time scales of years to millenia, Astronomy & Astrophysics, 509, doi:10.1051/0004-6361/200913276, 2010.

Wang, J., Zeng, N., Liu, Y., and Bao, Q.: To what extent can interannual CO2 variability constrain carbon cycle sensitivity to climate change in CMIP5 Earth System Models?, Geophys. Res. Let., 41, 3535–3544, doi:10.1002/2014GL060004, 2014.

Wanner, H., Beer, J., Buetikofer, J., Crowley, T. J., Cubasch, U., Flueckiger, J., Goosse, H., Grosjean, M., Joos, F., Kaplan, J. O., Kuettel, M., Mueller, S. A., Prentice, I. C., Solomina, O., Stocker, T. F., Tarasov, P., Wagner, M., and Widmann, M.: Mid- to Late Holocene climate change: An overview, Quat. Sci. Rev., 27, 1791–1828, doi:10.1016/j.quascirev.2008.06.013, 2008.

Wenzel, S., Cox, P. M., Eyring, V., and Friedlingstein, P.: Emergent constraints on climate-carbon cycle feedbacks in the CMIP5 Earth System Models, J. Geophys. Res., 119, 794–807, doi:10.1002/2013JG002591, 2014.

Woodwell, G., Mackenzie, F., Houghton, R., Apps, M., Gorham, E., and Davidson, E.: Biotic feedbacks in the warming of the Earth, Clim. Change, 40, 495–518, doi:10.1023/A:1005345429236, 1998.

Wunsch, C. and Heimbach, P.: Practical global oceanic state estimation, Physica D – nonlinear phenomena, 230, 197–208, doi:10.1016/j.physd.2006.09.040, Data Assimilation for Geophysical Systems Program, Stat & Appl Math Sci Inst, Research Triangle Pk, NC, 2005, 2007.

Wunsch, C. and Heimbach, P.: How long to oceanic tracer and proxy equilibrium?, Quat. Sci. Rev., 27, 637–651, doi:10.1016/j.quascirev.2008.01.006, 2008.

Zanchettin, D., Timmreck, C., Graf, H.-F., Rubino, A., Lorenz, S., Lohmann, K., Krueger, K., and Jungclaus, J. H.: Bi-decadal variability excited in the coupled ocean-atmosphere system by strong tropical volcanic eruptions, Clim. Dyn., 39, 419–444, doi:10.1007/s00382-011-1167-1, 2012.

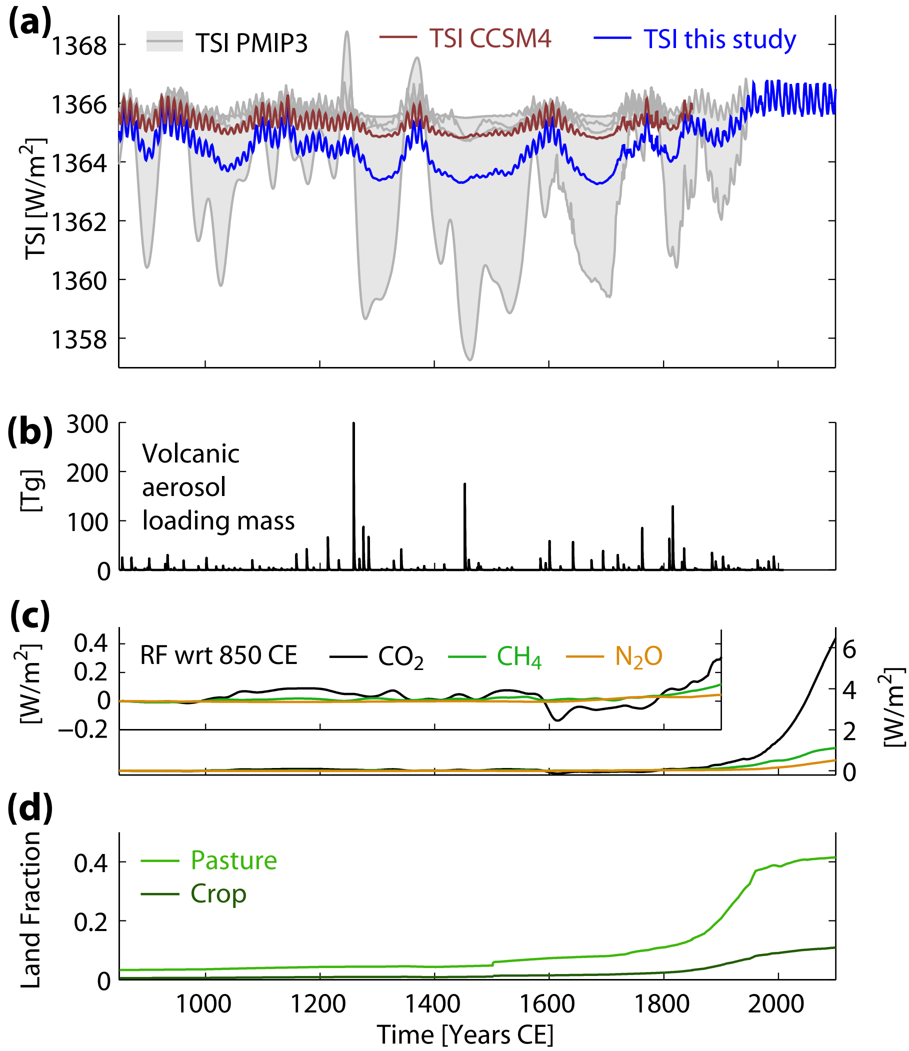


Figure 1. Forcings used in the last millennium simulation with CESM. (a) TSI in comparison with the different TSI reconstructions proposed by PMIP3. (b) Volcanic forcing as total volcanic aerosol mass. (c) Radiative forcing (RF, calculated according to IPCC, 2001) from the greenhouse gases CO_2, CH_4, and N_2O. (d) Major changes in land cover (as fraction of global land area). See text for details.

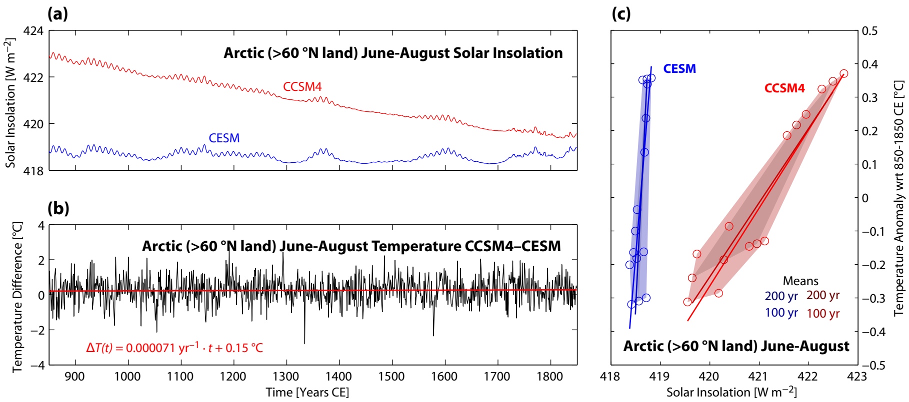


Figure 2. (a) Mean June-August (JJA) Arctic (>60° N land) solar insolation in CCSM4 with time-varying orbital parameters and CESM with fixed orbital parameters. (b) Arctic JJA temperature difference between CCSM4 and CESM. The least-squares linear trend of this temperature difference is given in red. (c) Arctic JJA temperature anomalies (from their 850-1850 AD mean) versus solar insolation as 100-year and 200-year averages (10 and 5 circles, respectively) from CCSM4 and CESM (red and blue, respectively). The least-squares linear trend for each cloud of 100-year and 200-year averages is given in the respective color. The shading envelops the range of temperature versus solar insolation for each cloud of means.

Table 1. List of simulations conducted for this study. See text for details regarding the forcing. TSI = total solar irradiance, GHGs = greenhouse gases, E_{CO_2} = anthropogenic CO_2 emissions from fossil fuel burning and cement production. LULUC = land use and land use change.

	Control simulation (CTRL)	Transient simulation (CESM)
Forcing	850 CE (500 years)	850-2099 CE
TSI	1360.228 W m^{-2}	adjusted Vieira and Solanki (2010) and Lean et al. (2005)
Volcanic	none	Gao et al. (2008)
GHGs	CO_2 (279.3 ppm) CH_4 (674.5 ppb) N_2O (266.9 ppb)	Schmidt et al. (2011)
E_{CO_2}	none	Andres et al. (2012) and Moss et al. (2010)
Aerosol	1850 CE from Lamarque et al. (2010)	Lamarque et al. (2010, 2011)
Orbital	1990 CE after Berger (1978)	1990 CE after Berger (1978)
LULUC	850 CE from Pongratz et al. (2008)	Pongratz et al. (2008) and Hurtt et al. (2011)

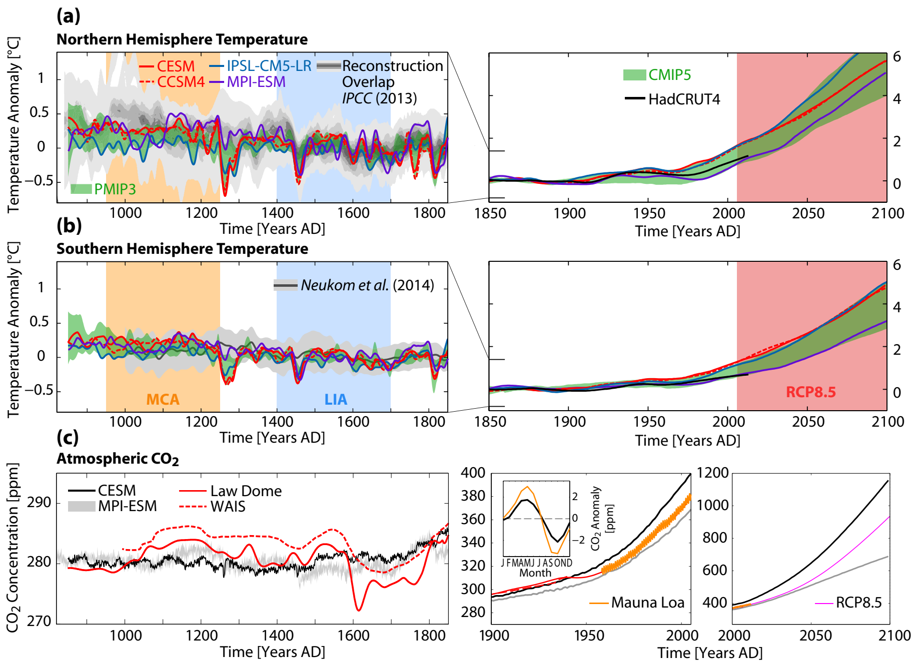


Figure 3. (a) Northern Hemisphere and (b) Southern Hemisphere temperature anomalies in model simulations and reconstructions. The anomalies are with reference to 1500-1899 CE (left panels) and 1850-1899 CE (right panels). Gray shading in (a) indicates the reconstruction overlap (IPCC, 2013), in (b) the reconstruction by Neukom et al. (2014). The 5-95% range of the simulations from the third Paleoclimate Modelling Intercomparison Project (PMIP3) and the fifth Coupled Model Intercomparison Project (CMIP5; applying the RCP 8.5) are given in green and red shading, respectively. Note that MPI-ESM applies the A1B scenario (IPCC, 2000), which has a weaker forcing than RCP 8.5. Hemispheric means from observations are shown as thick black line (Cowtan and Way, 2014). All time series have been smoothed by a local regression filter which suppresses variability higher than 30 years. The Medieval Climate Anomaly (MCA) and the Little Ice Age (LIA) are indicated as defined in Mann et al. (2009). (c) Evolution of atmospheric CO_2 in CESM (black), MPI-ESM (grey; ensemble range), from ice cores (red), from measurements (orange), and from RCP8.5 used to force the radiative code in CESM (magenta). The small inset in the middle panel shows the observed annual cycle at Mauna Loa, Hawaii, and a $2^\circ \times 2^\circ$ average over Hawaii from CESM, both derived from the period 1958-2012.

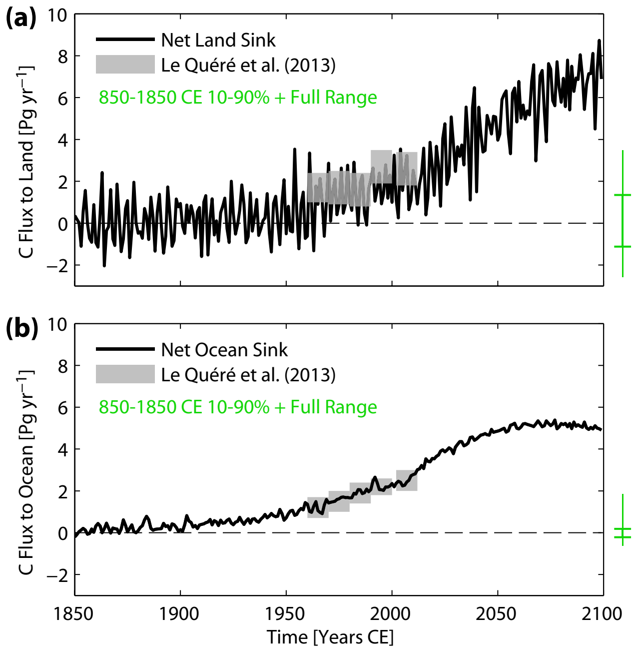


Figure 4. Annual mean net carbon flux from the atmosphere to (a) land and (b) ocean. Green bars given the full and 10-90% range from the preindustrial part of the simulation. Observational estimates are from Le Quéré et al. (2013)

Table 2. Selected forcing details ofsimulations used in comparisons with CESM. TSI = total solar irradiance, LULUC = land use and land use change.

Forcing	CESM	CCSM4	IPSL-CM5A-LR	MPI-ESM
TSI	adjusted Vieira and Solanki (2010) and Lean et al. (2005)	Vieira and Solanki (2010) and Lean et al. (2005)	Vieira and Solanki (2010) and Lean et al. (2005)	Krivova et al. (2007)
Volcanic	Gao et al. (2008)	Gao et al. (2008)	Gao et al. (2008)	Crowley et al. (2008)
Orbital	1990 CE, Berger (1978)	transient, Berger (1978)	transient, Berger (1978)	transient, Bretagnon and Francou
LULUC	Pongratz et al. (2008) and Hurtt et al. (2011)	Pongratz et al. (2008) and Hurtt et al. (2011)	non-transient	Pongratz et al. (2008)

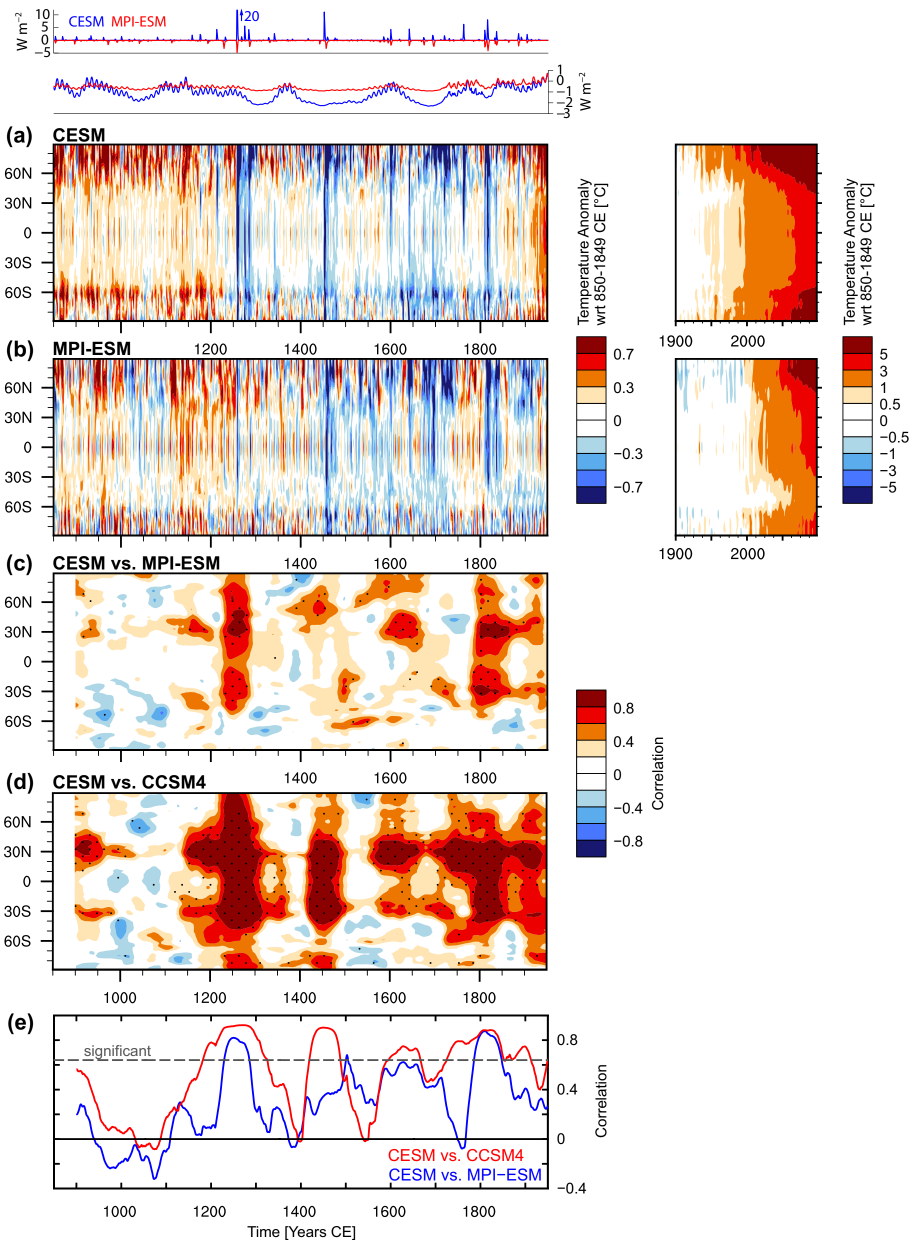


Figure 5. 5-year filtered zonal mean anomalies of surface air temperature (SAT), relative to 850-1849 CE from (a) CESM and (b) MPI-ESM. (c) 100-year running-window correlation of zonal mean SAT from CESM and MPI-ESM. 0.75 Tukey window has been applied to the data before correlation to weaken sharp transitions. Stippling indicates significance (5% level), taking into account autocorrelation estimated from the entire time period. (d) As (c) but for the correlation of CESM with CCSM4. (e) As (d) but for global mean SAT. Small inset on top shows volcanic and solar forcing of CESM and MPI-ESM. Volcanic forcing of CESM scaled to have the same radiative forcing as MPI-ESM for Pinatubo in 1991 CE. Solar forcing relative to 1850 CE.

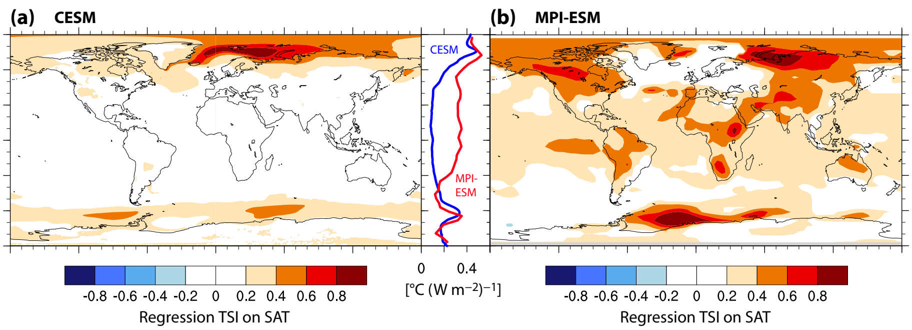


Figure 6. Regression of total solar irradiance (TSI) on surface air temperature (SAT) for the period 850-1850 CE in (a) CESM and (b) MPI-ESM. Time series at each gridpoint have been 5-year filtered. Only significant regression coefficients at the 5% level are shown. The small panel shows zonal means.

Table 3. Cumulative carbon emissions by different components over different time periods in CESM, in Pg C. Positive (negative) values indicate emission to (uptake from) the atmosphere.

	850-1500 CE	1501-1750 CE	1751-2011 CE	2012-2100 CE
Ocean	26.0	-4.0	-151.3	-413.0
Land	-15.0	10.3	82.5	139.4
Land (without LULUC)	-24.4	-9.3	-94.7	-436.3
Fossil Fuels	0.0	0.0	358.0	1,851.5

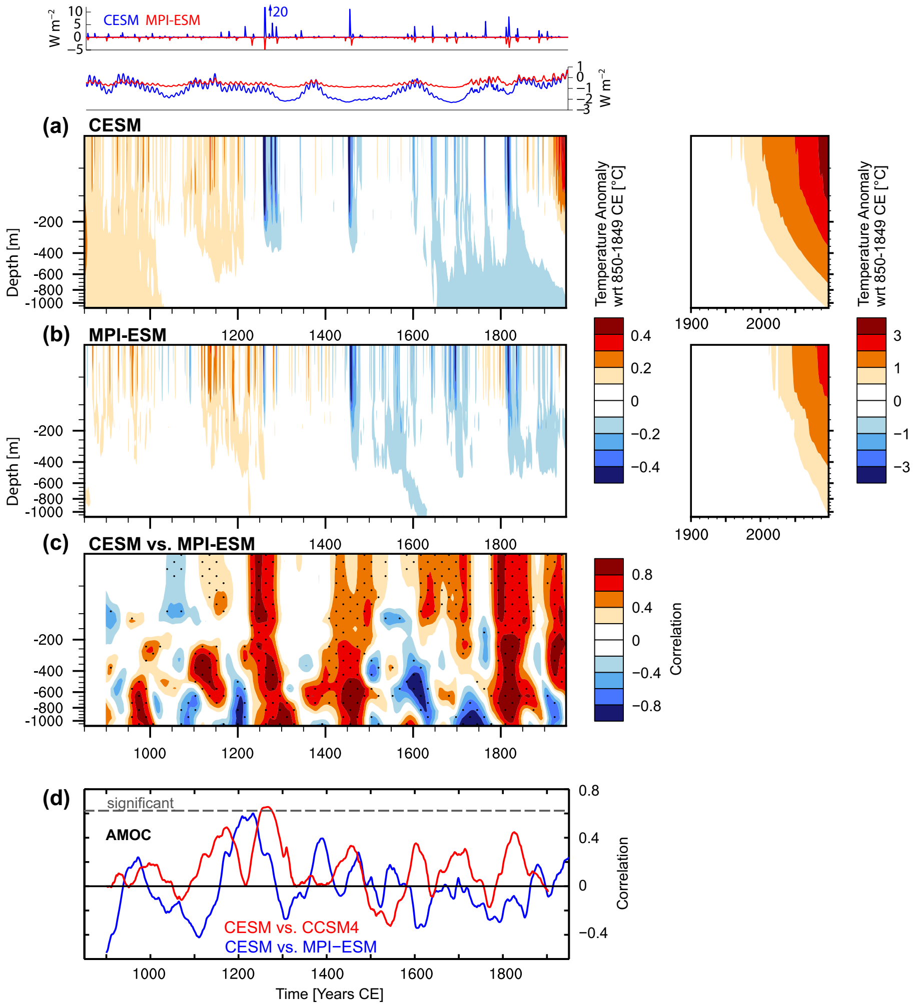


Figure 7. 5-year filtered zonal mean anomalies of horizontally averaged ocean temperature, relative to 850-1849 CE from (a) CESM and (b) MPI-ESM. (c) 100-year running-window correlation of zonal mean SAT from CESM and MPI-ESM. A 0.75 Tukey window has been applied to the data before correlation to weaken sharp transitions. Stippling indicates significance at the 5% level, taking into account autocorrelation estimated from the entire time period. (d) 100-year running-window correlation of the Atlantic Meridional Overturning Circulation (AMOC) in CESM and MPI-ESM.

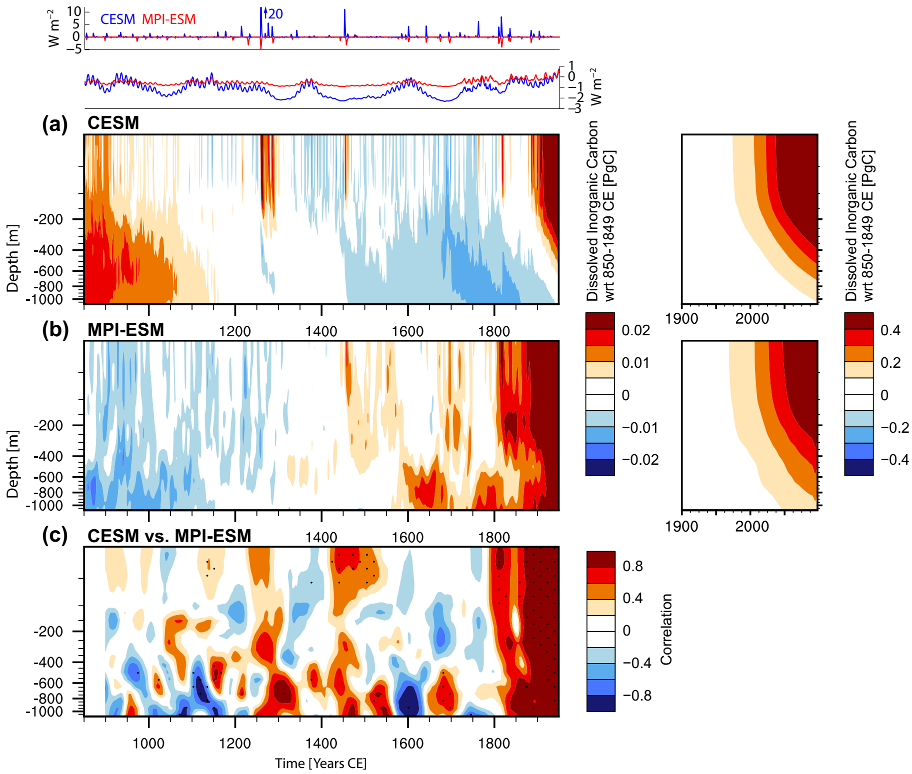


Figure 8. 5-year filtered zonal mean anomalies of horizontally integrated dissolved inorganic carbon (DIC), relative to 850-1849 CE from (a) CESM and (b) MPI-ESM. (c) 100-year running-window correlation of zonal mean SAT from CESM and MPI-ESM. A 0.75 Tukey window has been applied to the data before correlation to weaken sharp transitions. Stippling indicates significance at the 5% level, taking into account autocorrelation estimated from the entire time period.

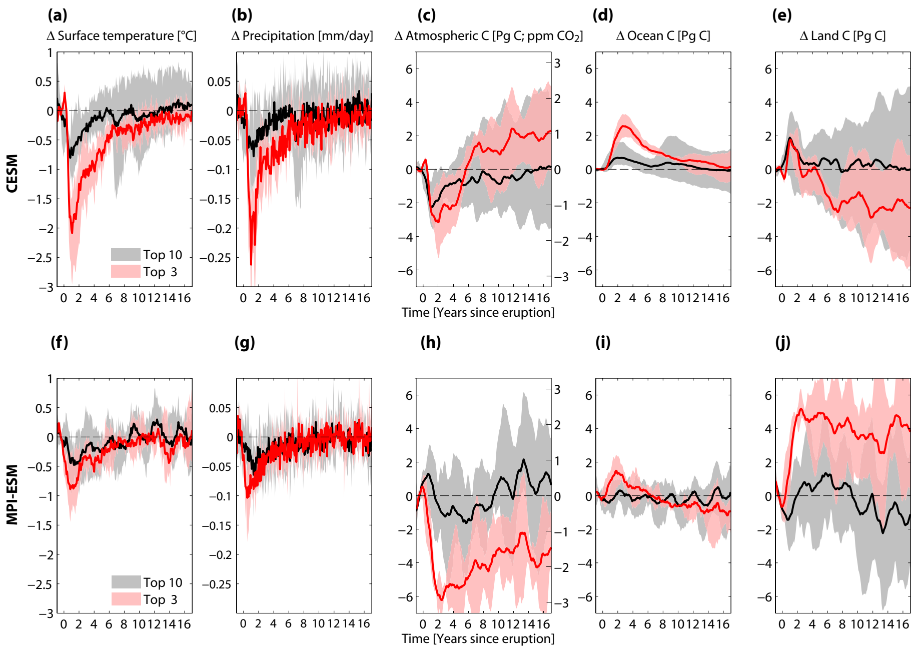


Figure 9. Superposed Epoch Analysis of the strongest three (top3) and following strongest seven eruptions (top10) of the period 850-1850 CE in (a-e) CESM and (f-j) MPI-ESM for (a, f) global mean surface air temperature, (b, g) global mean precipitation, (c, h) atmospheric carbon given in Pg C on the left y-axis and in ppm CO_2 on the right y-axis, (d, i) ocean carbon, and (e, j) land carbon. Time series are deseasonalized and calculated as anomalies to the mean of the preceding five years. The shading shows the 10-90% range.

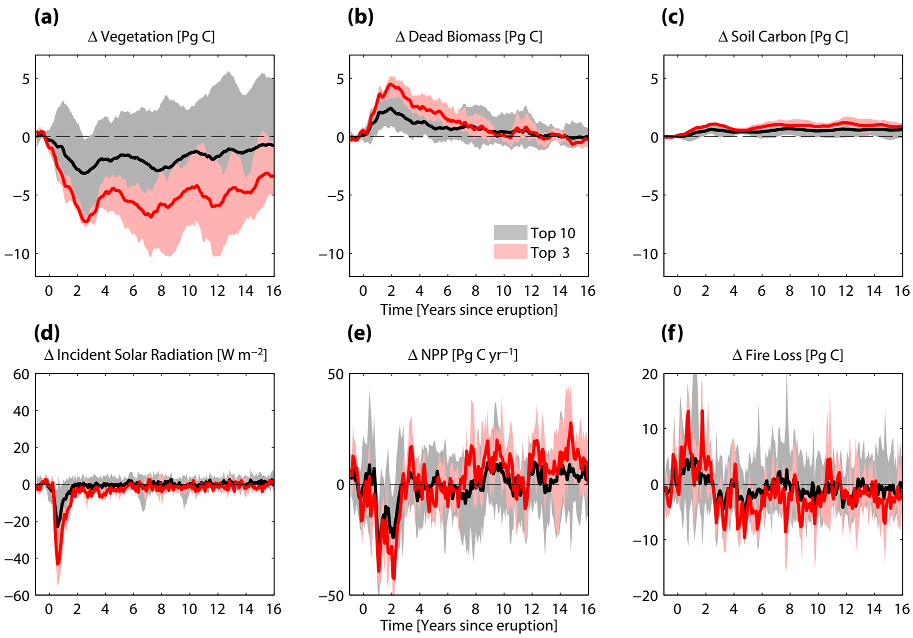


Figure 10. Superposed Epoch Analysis of the strongest three (top3) and following strongest seven eruptions (top10) for tropical land (25° S to 25° N) in CESM during the period 850-1850 CE. Land carbon inventory changes split up in (a) vegetation, (b) dead biomass (litter and wooden debris), and (c) soil. Further, changes in (d) solar radiation, (e) net primary production (NPP), and (e) loss of carbon through fire. Time series are deseasonalized and calculated as anomalies to the mean of the preceding five years. The shading shows the 10-90% range.

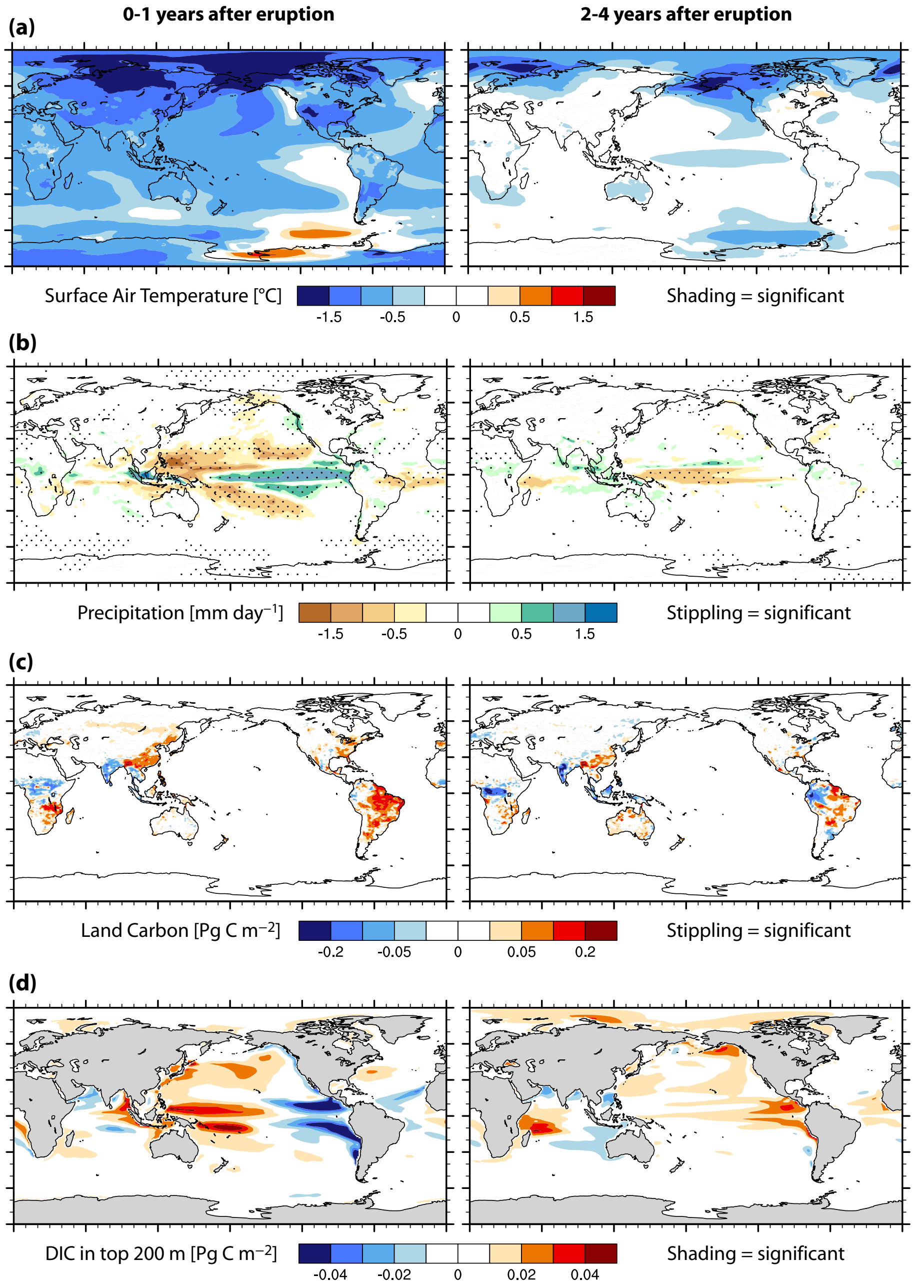


Figure 11. Composites of top10 post-volcanic eruption years as anomalies to the preceeding 5 years, averaged over (left) the first 2 years startting with the year of the eruption, and (right) the following three years. (a) Surface air temperature, (b) precipitation, (c) total land carbon, (d) dissolved inorganic carbon (DIC) integrated over the top 200 meters. Shading or stippling indicates significance at the 5% level. Note, that for land carbon at an individual grid cell hardly any significant changes are detected due to the large inter-annual variability.

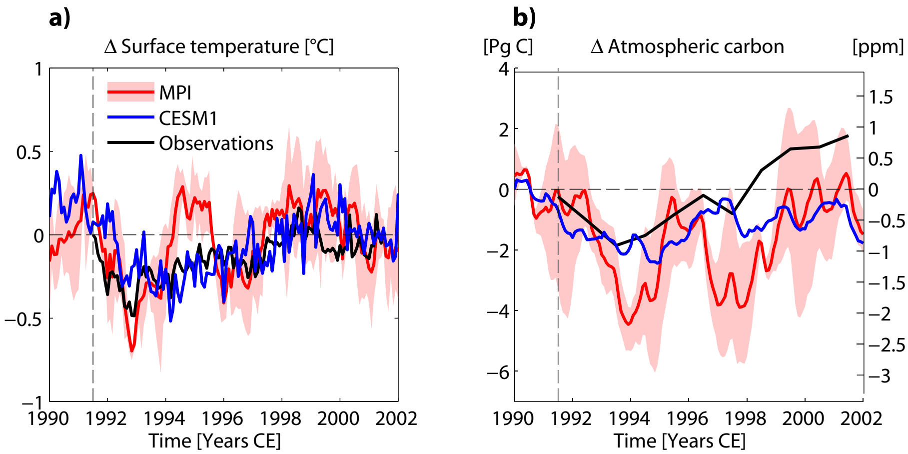


Figure 12. Global mean changes in response to Pinatubo. (a) Global mean surface air temperature and (b) atmospheric carbon, both deseasonalized and linearly detrended over 30 years centered on June 1991; temperature observations were corrected for El Niño-Southern Oscillation and other dynamical components (Thompson et al., 2009), CO_2 observations were corrected for El Niño-Southern Oscillation and anthropogenic emissions (Frölicher et al., 2013).

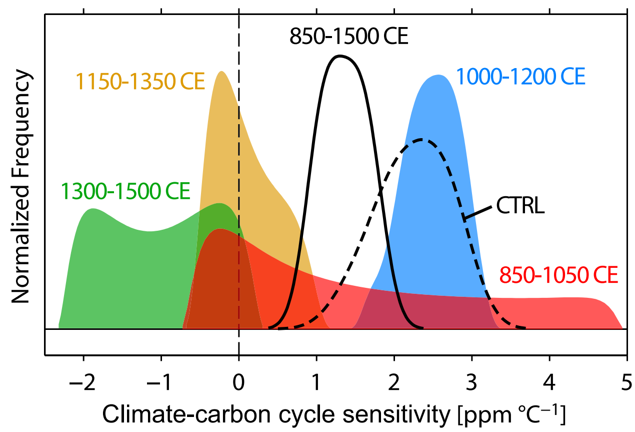


Figure 13. Temporal dependence of the ~~climate carbon~~ climate-carbon cycle sensitivity γ in CESM. Normalized probability density functions (PDF) of γ for 200-year windows overlapping by 50 years (color-filled), for the full period 850-1500 CE (black solid), and for the CTRL (black dashed). The spread of each PDF arises from the range of low-pass filters applied (20 to 120 years).