Response to Referee RC1 (Chris Jones):

We would like to thank the referee for the review of this manuscript and their constructive comments. Our response to each comment is below with the referee's comments highlighted in bold typeface. Where appropriate, we have also included relevant changes in the revised manuscript in italic typeface.

My main concern with this model is the length of time taken to achieve a spun-up state. Law et al document this nicely, but I think it requires more discussion in this paper too what the implications are. The drifts in carbon stores are still non-egligible even after 800 years of spin-up (your start point here). I think this should be laid out explicitly before the analysis starts. You do, in the land carbon section, acknowledge this and subtract the control run drift. But unless a reader has been through Law et al they would not know how big this drift is. For the ocean it is more important still, and the ocean section does not mention this at all. The drift of circa 0.7 GtC/yr (figure 11a in Law et al) is of similar magnitude to your historical fluxes (I assume these are corrected for the drift). If a reader hadn't seen that figure then they would not realise from this paper the size of drift being subtracted.

We will include a paragraph in the Simulations section 2.2 to acknowledge the drift in carbon stores more explicitly for land and ocean. We will also make it clear in the ocean section that the historical fluxes have been corrected for the drift.

As noted in Law et al. (2015) the net carbon fluxes for land and ocean did not equilibrate to zero. At the end of the control run (i.e. year 800 to 955), global NEE is 0.3 PgC/yr for PresLAI and 0.08 PgC for ProgLAI. The net autgassing from the ocean is about 0.6 PgC/yr at the end of the control run. We take this drift into account when we calculate the net uptake of carbon for land and ocean.

In CONCENTRATION-driven runs like this, you can of course force the correct CO2 and correct for the drift after the simulation. But that is not possible in an EMISSION-driven run, and such a drift would cause a massive drift in atmospheric CO2 render-ing an emissions-driven historical run meaningless. This would, at present preclude use of this model in C4MIP for example which would be a great shame. The latest C4MIP protocol (Jones et al 2016, GMDD - CMIP special issue) recommends a maximum acceptable drift of 10 GtC per century. I therefore throughly recommend that ACCESS modellers attempt to find accelerated means to derive a spun-up state in time for CMIP6. There are numerous options, such as running offline (for either land or ocean), or using reduced turn-over time techniques as per Koven et al for CLM (http://www.biogeosciences.net/10/7109/2013/bg-10-7109-2013.pdf). as a final word on this, lack of carbon conservation would also be more of an issue for E-driven runs.

We are currently exploring various spin-up options (including offline simulations) to improve the drift for ocean and land. If successful, this will then be implemented in future versions of ACCESS-ESM.

My second concern is the lack of land-use change as a forcing. You already know this, and acknowledge it in the paper, so no revisions to the manuscript are required, but I just take this opportunity to stress that simulations of contemporary and future climate/carbon cycle are very much reduced in usefulness if they lack the

very large land-use forcing of the land carbon cycle. Implementing this for CMIP6 is also therefore a priority I would say.

We agree with the referee and it is a high priority for us to include land use change in future versions of ACCESS-ESM.

Having quickly read Law et al before I reviewed this one I was struck that there was not an evaluation there of (land) carbon stocks. I do feel that the land carbon modelling community have become fixated on evaluation against fluxes to the detriment of stocks and residence times. This is beginning to change and I was pleased to see some discussion of carbon stocks in this paper. It would be nice to see the time changes in these though as well - could table 2 be extended to show pre-industrial and present day stocks? In the discussion on biomass you mention that your results are higher than observational based estimates - but of course you lack land-use change as a driver. So it could easily be expected that your biomass is of the order 100-150 GtC too high due to this. If you masked for present day agricultural regions you would probably get a much closer fit to expected global totals. So your simulation is actually not bad.

In a revised version we will extend Table 2 to show both, pre-industrial and present day stocks. We thank the referee for the comment about the size of our biomass pool and we will include this statement in a revised version.

	Pre-industrial						Present day						Historical change C	
	PresLAI			ProgLAI			PresLAI			ProgLAI			PresLAI	ProgLAI
Pool	C	N	P	C	N	P	C	N	P	C	N	P	ΔC	ΔC
Biomass	611	5.7	0.31	731	6.15	0.33	670	6.2	0.34	807	6.8	0.37	69.5	87.2
Litter	117	0.85	0.04	149	1.02	0.05	126	0.9	0.05	163	1.1	0.06	7.6	12.3
SOC	1034	82	9.6	1187	86.1	11.9	1050	83.4	10.1	1217	88.5	12.6	20.5	37
$\sum$	1762	88.6	10.0	2067	93.3	12.3	1846	90.5	10.5	2187	96.4	13.0	97.6	136.5

Abstract - you can say that aerosol forcing is large or larger than other models. But don't say "over-sensitive" as we simply don't know. Maybe this is correct ...

We will change this statement as suggested in a revised version.

section 2.3 - I like the comparison vs CMIP5 models. This is a nice way to put model-data discrepancies in context. But why do you only include 5 CMIP5 models (counting the 2 IPSLs as one model). Anav et al used more than twice that - any reason not to use the full set?

We agree that ideally we would have been able to use more Earth System Model (ESM) results like Anav et al., though unfortunately, at the time we were doing the analysis not all the output was readily available from the ESG servers. After careful examination of our final results and those of Anav et al., we are confident that our figures still capture the spread of CMIP5-ESM results and more models may only serve to broaden the spread of the models.

in a couple of places (e.g. start sec 3.1) you mention the variability of the land

sink and/or the atmospheric CO2. You could go further and use this as an evaluation metric. Both Cadule et al (2010, GBC) and Cox et al (2013, Nature) show the power of the C-cycle sensitivity to ENSO on inter-annual timescales as a really strong evaluation metric and constraint on the sensitivity of ESMs

We thank the referee for his suggestion and will consider this in future evaluation papers using the next version of ACCESS-ESM.

p.7, line 21. I don't understand why you attribute your slower warming to the initial warm bias - how do you know the warm bias causes this? Could just be under-sensitive SSTs not related to a bias.

The discussion was intended to be descriptive rather than mechanistic. We clarify now the text so as not to attribute the reduced warming to the warm bias early in the historical period.

p.8 line 3. I was amused to see that having an error different from other models was "encouraging"! not sure why! Can you say why it is better to have the opposite MLD bias from other models?

Encouraging in the sense that simulating well winter mixed layer depths are critical for setting interior ocean properties supplying nutrients to the upper ocean to fuel the biologically active growing season. We have now augemented the manuscript to say:

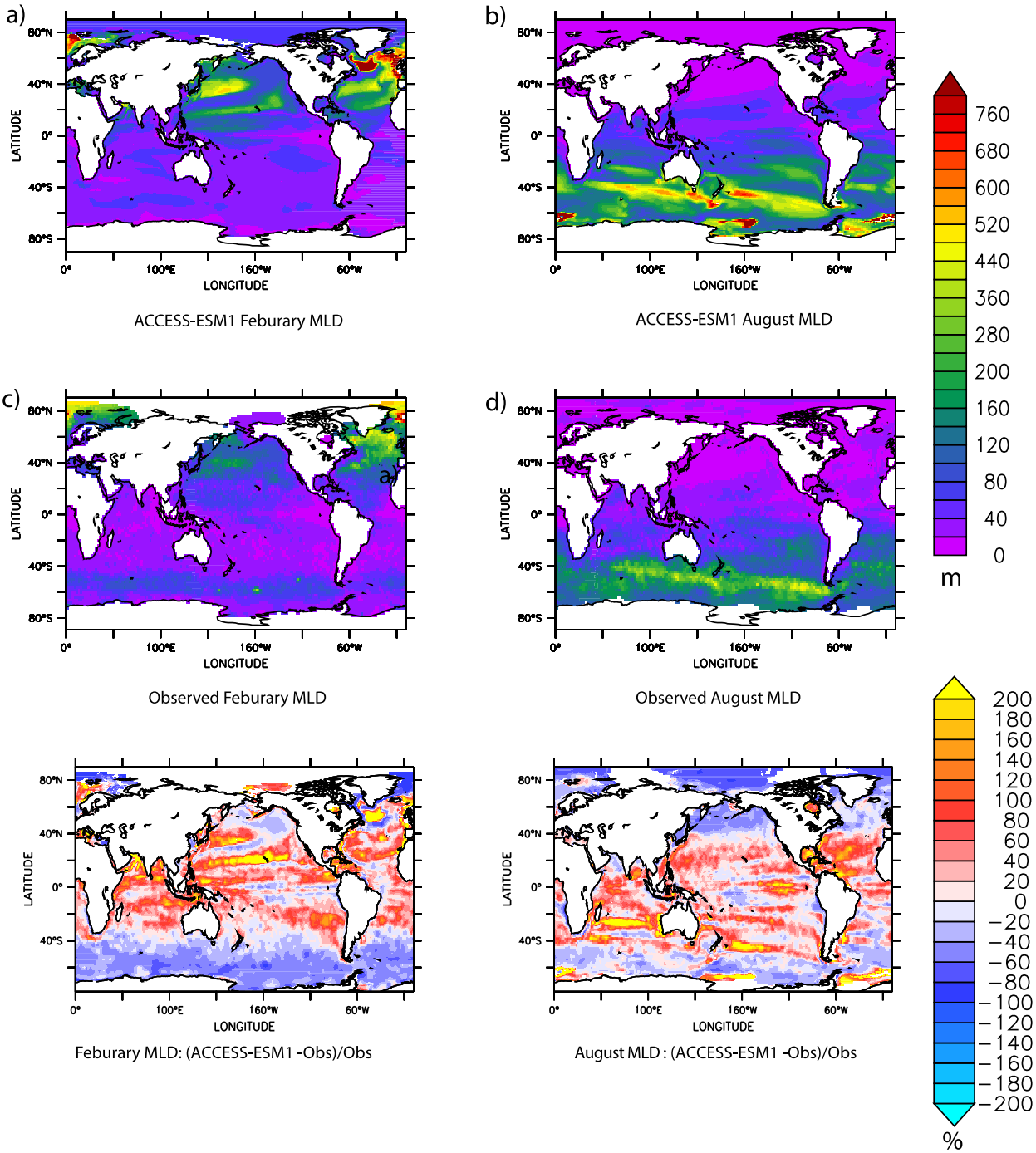
In the higher latitude the winter mixed layers are well captured by ACCESS-ESM1 Figure 4. This is encouraging given that many ocean models tend to underestimate winter mixed layer depths (Sallee et al (2013) and Downes (2015)). Simulating winter mixed layers correctly is critical for setting interior ocean properties supplying nutrients to the upper ocean to fuel the biologically active growing season (Rodgers et al, 2014). However, in contrast to the winter, ACCESS-ESM1 appears to systematically underestimate mixed layer depths in the high latitude ocean in summer, ~60% (or 30-40m) in the Southern, Pacific and Atlantic Oceans. In the Southern Ocean, in particular, the underestimation of summer mixed layer depths is consistent with Sallee et al (2013) and Haung et al (2013) who showed that most CMIP5 models underestimate summer mixed layer depths. Haung et al (2013) attributed this to a lack of vertical mixing in CMIP5 rather than sea surface forcing related to individual models, this is consistent with Downes et al (2015) who showed that these biases are also present in the ocean only simulations of ACCESS-ESM1 (ACCESS-O).

in general I thought this MLD section (3.2) was a bit superficial - can you be a bit more quantitative in your comparison and description? the figure shows the data but it can be hard to tell from there if the differences you describe are of the order of a few % or 10s of % or factors of 2 or more or what? And can you mention your confidence in the ons? presumably global maps of MLD are not directly observed but must have certain extrapolation uncertainties and so on. Are some areas/seasons of the oceans better sampled than others etc …

To better illustrate the chages we have added a lower panel to Figure 4 that is the percentage changes of the difference between the (ACCESS-ESM1 – obs/obs) *100. This illustrates the relative changes as suggested by the reviewer. Also please see above for an enhanced discussion.

Regading the number of samples and coverage of the obs we have added the following paragraph:

Ocean mixed layer depths are compared with the observations following DeBoyer Montegut et al (2004) based on more than 880,000 depth profiles from research ships and ARGO profiles, and based on a 0.03kg/m^3 density change from the surface. Significant advances in autonomous measurement platforms have allowed the mixed layer to be increasingly well constrained in all seasons across the global ocean.



p. 8 line 21. Can you define what you mean by IAV. Interannual Variability I know, but how do you turn this into a number? is it the standard deviation of a time series of annual means? in which case is the time series de-trended first? Etc

The interannual variability (IAV) is calculated as the standard deviation for the de-trended annual mean values. This explanation will be included in a revised version.

p.9 When discussing historical changes in land carbon can you split into veg and soil changes (e.g. put in table 2). You could compare directly with the 2 models in Jones et al (2013) which also don't have land-use forcing (dashed lines in figure 2). You could probably also compare with model results from detection-and-attribution studies which ran with/without certain forcings. The are probably various no-land-use runs to look at.

We will included the change in land carbon for vegetation, litter and soil for both scenarios (prescribed LAI and prognostic LAI) in Table 2 and also compare these values against simulation of the two models without land use change from Jones et al. (2013) in a revised version.

	Pre-industrial						Present day						Historical change C	
	PresLAI			ProgLAI			PresLAI			ProgLAI			PresLAI	ProgLAI
Pool	C	N	P	C	N	P	C	N	P	C	N	P	ΔC	ΔC
Biomass	611	5.7	0.31	731	6.15	0.33	670	6.2	0.34	807	6.8	0.37	69.5	87.2
Litter	117	0.85	0.04	149	1.02	0.05	126	0.9	0.05	163	1.1	0.06	7.6	12.3
SOC	1034	82	9.6	1187	86.1	11.9	1050	83.4	10.1	1217	88.5	12.6	20.5	37
$\sum$	1762	88.6	10.0	2067	93.3	12.3	1846	90.5	10.5	2187	96.4	13.0	97.6	136.5

When we calculated the change in land carbon based on the pool sizes rather than using the net flux, we noticed that the total land carbon uptake over the historical period is actually much smaller than stated in the paper: 98 PgC instead of 128 PgC (PresLAI) and 137 PgC instead of 154 PgC. This is because we used an earlier section of the control run (years 325-480) to calculate the drift. However, the historical runs we describe in the paper were started at year 800 from the control run and this section (i.e. year 801-955) shows a much smaller drift for both scenarios. Using the correct drift applied to the net flux over the historical period provides now about the same total land carbon uptake as calculated via the pool sizes. The numbers for the total uptake of carbon by the land will be corrected in a revised version.

We have also updated the NEE time series plot (Fig. 5c) using the correct drift, although changes are hardly visible, particularly for the 5yr running mean.

Sec 4.2. Despite the large drift your historical ocean sink does look a very close match to the obs. Can you also quote a cumulative uptake here?

We have added the following text to the paper:

The cumulative uptake of carbon by air-sea CO_2 fluxes in the period 1959-2005 from ACCESS-ESM1 is 83 PgC which is good agreement with the GCP value of 82PgC over the same period.

sec 5.1.2. I was curious to see that your prescribed LAI didn't match your obser-

vations. Can you explain why not? You illy that this is because there are differences between observed datasets - which is of course true. But then why do you choose one dataset to prescribe LAI to the model, and then a different one to evaluate against? If one is better than the other can you use it for both?

The prescribed LAI we use in ACCESS-ESM1 is based on MODIS observations and has no interannual variability. We decided to compare this with a more recent LAI product (MODIS/AVHRR combination) to (a) investigate if there are significant differences in the mean seasonal cycle for present day (as shown in Fig. 10) and (b) to investigate what the interannual variability in the observations looks like (not show in the manuscript).

We don't think that one LAI product is better than another. Historically, CABLE uses a prescribed LAI with no interannual variability. However, in future CABLE/ACCESS versions we might update our prescribed LAI to a product with interannual variability.

As mentioned in the paper, LAI products differ from each other because different sensors and algorithms are used. In addition, the LAI products shown here are also based on different observing time periods.

p.13 discussion of different carbon stocks for the two configurations. Given both have very similar GPP, the large difference in biomass is presumably due to residence times? Having GPP further north in compared to the tropics means that for the same global GPP you have a higher biomass? You might consider next time an evaluation of turnover times - e.g. as per Carvalhais et al (2014, Nature)

We agree with the referee, that the difference in carbon stocks for the two scenarios (prescribed LAI and prognostic LAI) can probably be explained by a difference in residence time. We thank the referee for pointing this out and we aim to include an evaluation of turnover times in future work.

section 5.2.1 / 5.2.2. Can you swap the order of figures 12 and 13? you describe 13 first.

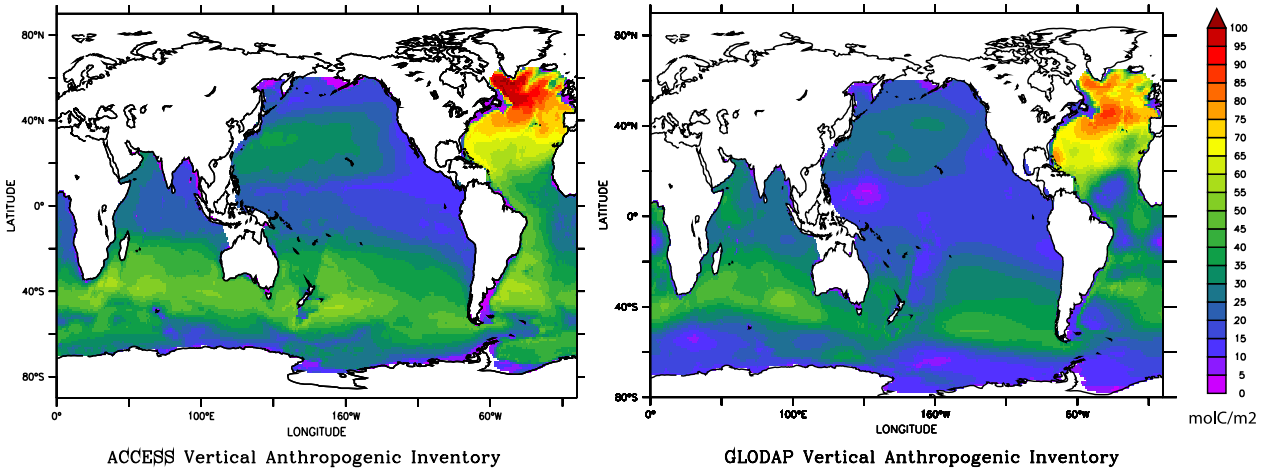
The order of these figures will be changed in a revised version.

sec 5.2.3. You compare global totals between model and GLODAP, but the map shows missing areas in the ons dataset. So should you first mask out the model to match the same area before you compare totals? (or quote a full AND a masked number for the model). The values in the Arctic for example are very high but missing in the obs (although of course the area is smaller than this projection makes it look).

We have redrawn the ACCESS-ESM1 map with only ares that observations exist. Furthermore we now quote the number from comparison with GLODAP and for the entire domain:

The inventory for the period 1850-1994 in ACCESS-ESM1 is 132 PgC, which is close to the estimated value from GLODAP of 118 +/- 19PgC (Sabine et al, 2004) over the same domain. This suggests that despite a somewhat limited representation of the seasonal cycle of sea-air CO_2 fluxes in key regions of anthropogenic uptake such as the Southern Ocean, ACCESS-ESM1 is doing a very good job, spatially and temporally, of capturing

and storing anthropogenic carbon. If the entire domain (including the Arctic Ocean) is integrated the anthropogenic uptake is 143 PgC over the same period.



sec 5.3. It seems reasonable that the land drives most of the seasonal cycle at your ons stations except the south pole. I like the way you have labelled the land and ocean CO2 separately to be able to diagnose this. But I couldn't work out why the contribution of the ocean at the south pole looks different for your two LAI configurations. If I read the figure correctly then the role of the ocean is given by the blue solid minus the blue dashed line and the red solid minus the red dashed line. These are quite different by eye - e.g. the blue lines are quite far apart in December and the red lines quite close. So why does your LAI treatment affect your December ocean fluxes so much?

The LAI scenario should not affect the ocean carbon fluxes and we thank the referee for pointing out this inconsistency. By mistake, we showed the total flux using two different ocean configurations in Figure 16 (mean seasonal cycle ot atmospheric CO2). Using the same ocean configuration for both scenarios (prescribed LAI and prognostic LAI) solves this issue and the dashed lines (land carbon flux) and solid lines (total flux) are much closer together (see figure below). Figure 16 will be updated in a revised version accordingly.

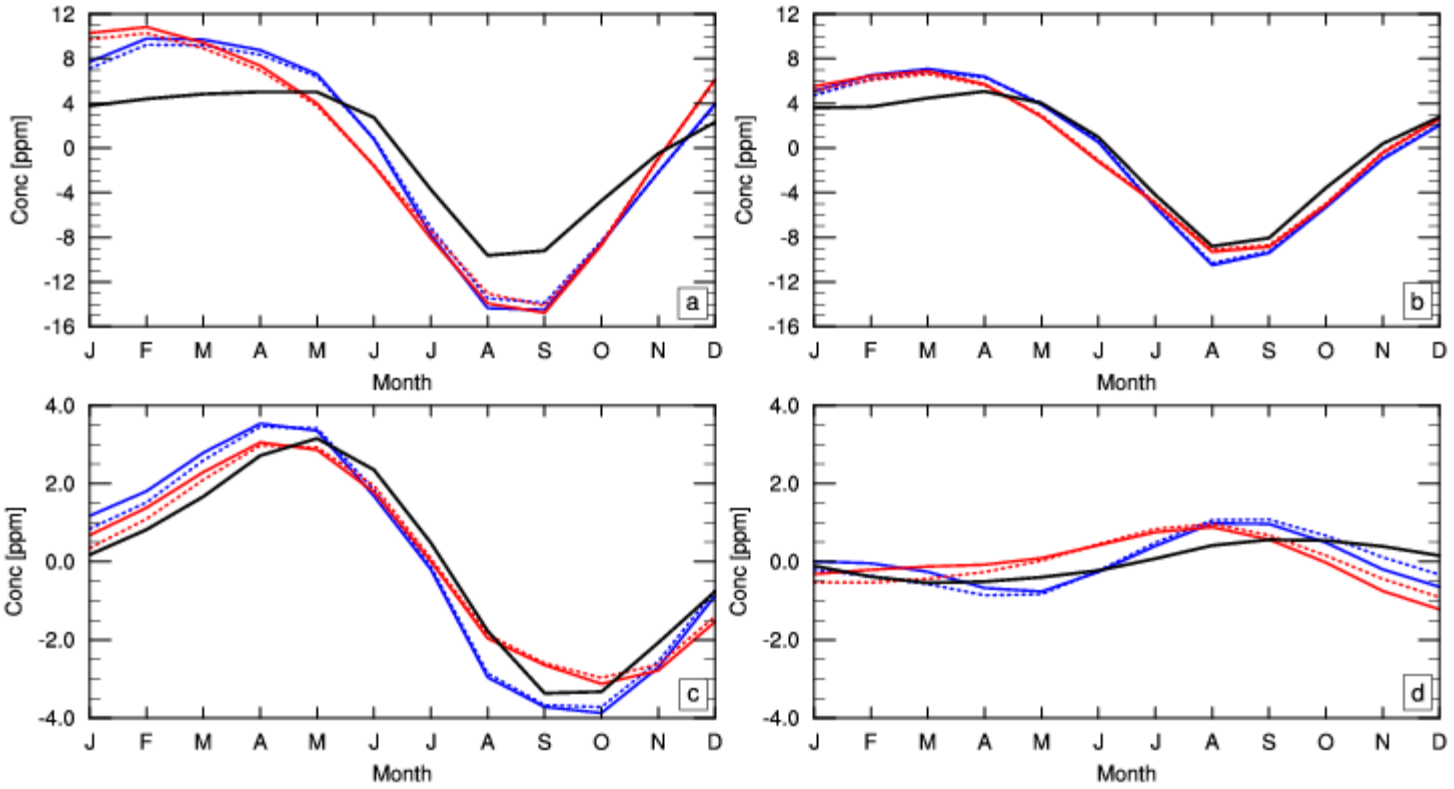


Table 2. Can you add a final row at the bottom of "total"

We will do this in a revised version.

	Pre-industrial						Present day						Historical change C	
	PresLAI			ProgLAI			PresLAI			ProgLAI			PresLAI	ProgLAI
Pool	C	N	P	C	N	P	C	N	P	C	N	P	ΔC	ΔC
Biomass	611	5.7	0.31	731	6.15	0.33	670	6.2	0.34	807	6.8	0.37	69.5	87.2
Litter	117	0.85	0.04	149	1.02	0.05	126	0.9	0.05	163	1.1	0.06	7.6	12.3
SOC	1034	82	9.6	1187	86.1	11.9	1050	83.4	10.1	1217	88.5	12.6	20.5	37
∑	1762	88.6	10.0	2067	93.3	12.3	1846	90.5	10.5	2187	96.4	13.0	97.6	136.5

general comment on figures - maybe a personal preference but please can you add a legend so that I can easily read which line is which. The text in the captions is very good - but I have to read a whole paragraph to spot what the red/blue lines are.

Would be great to simply see a legend with this in as well as the detail in the caption text.

A legend will be added to all relevant figures in a revised version.

Response to Referee RC2:

We would like to thank the referee for the review of this manuscript and their constructive comments. Our response to each comment is below with the referee's comments highlighted in bold typeface. Where appropriate, we have also included relevant changes in the revised manuscript in italic typeface.

The main focus of these specific comments is to highlight text that could benefit from being more quantitative rather than subjective. In addition, there are some specific suggestions for clarifications aimed at improving the overall flow and clarity of the manuscript. Before starting the specific suggestions, one other general remark I have is that the figure axes and color bar labels are very small and hard to read.

We will improve the readability (i.e. increase size of labels) of all figures in a revised version.

Abstract

The abstract provides an overall view of how well ACCESS-ESM1 performs; however, they provide this information using terms such as "good" and "performs well". How well? Good according to whom? Is your good the same as my good? First off, the authors state ACCESS-ESM1 overestimates the seasonal amplitude of LAI, but do not attach any quantities to this statement. Rather than providing a quantitative assessment of this statement, the reader is left to wonder if this is a substantial bias, perhaps even being prohibitive of using this model in the future, or merely a relatively minor difference that is offset by other positive features. The paper then continues with the statement that the oceanic and land fluxes "show good agreement with the observations", but again, no metric is used. How big are the differences, what is the error, or how closely correlated, both spatially and temporally, are they to observations and/or other CMIP-5 models? In the last overview comparison, the authors state that the seasonal cycle is "close to the observed seasonal cycle", but as the reader I have to wonder how close is close? I believe that putting quantifiable metrics on at least some of these statements will strengthen the concluding remark that ACCESS-ESM1 is indeed a useful tool.

We will include more quantitative statements in the abstract in a revised version.

Observations

The model evaluation (later in the text) proceeds through a straightforward succession of comparisons; however, the order of these comparisons is different than the order the data is listed in. For clarity and consistency, it would be an easy fix to reorganize this section to present the data in the order that they are used.

We will reorganise section 2.4 (Observations) in a revised version, so that data are presented in the order they are used later in the comparison.

Land Temperature and Precipitation

This section presents a time-series comparison of temperature and precipitation, but contains very little quantitative statements. For example, the temperature anomalies are "close to the observed anomalies through most of the period". Again I have to wonder, how close is close? Perhaps a difference plot showing the errors would be useful? Or perhaps a correlation coefficient that may or may not be

significant? Some sort of metric on this statement would be much more enlightening to the reader. For example, as I look at Fig. 1 myself, I see the author's point that ACCESS-ESM1 is lower than the observations 1965-2005; however, to my eye I also see the model looks quite a bit lower in the 1940s. I think that a difference plot would help identify these areas, rather then relying on the reader to have to assess the differences between the models and the observations on eyesight alone. Another example later in this section occurs in the precipitation anomaly discussion: the authors state the differences are "generally small," but provide no values to suggest what is meant by small. This is followed by the statement that "the simulations compare well with observed rainfall anomalies until about 1950," with no supporting metric such as minimal differences or significant correlation coefficients to back up this statement. Overall, I found that this entire section had very few quantitative comparisons, instead relying heavily on subjective terminology, making me believe that at least some measurable metrics would improve the paper.

We have attached an updated Fig. 1 showing differences for temperature and precipitation anomaly (simulated – observed). However, we might not include the difference plot in the manuscript, because we don't think it adds a lot more information. Instead, we will focus on decadel mean differences and refer to the values in the text as appropriate:

Both ACCESS-ESM1 simulation scenarios (PresLAI and ProgLAI) show similar temperature anomalies over most of the historical period, being close to the observed anomalies through most of the period (decadal mean difference smaller than 0.2 K), apart from the 1940s where the PresLAI scenario shows a larger negative anomaly (decadal mean difference of about 0.37 K), which will be discussed later. From about 1965-2005 anomalies are by up to 0.4 K (decadal mean difference) lower than observations for both scenarios.

ACCESS-ESM1 simulations compare well with observed rainfall anomalies until about 1960 (decadal mean difference smaller than 8 mm/yr), with the exeption of the period 1911-1920 for PresLAI (decadal mean difference of about 12 mm/yr) and the period 1951-1960 for ProgLAI (decadal mean difference of about 17 mm/yr). After that, observed anomalies are mostly higher than the simulation results (decadal mean difference of up to 41 mm/yr), a feature also seen in the ACCESS1.3 historical ensemble (Fig. 6a, Lewis and Karoly 2014).

The anomaly in the 1940s is already discussed in the manuscript on page 6, lines 26-30.

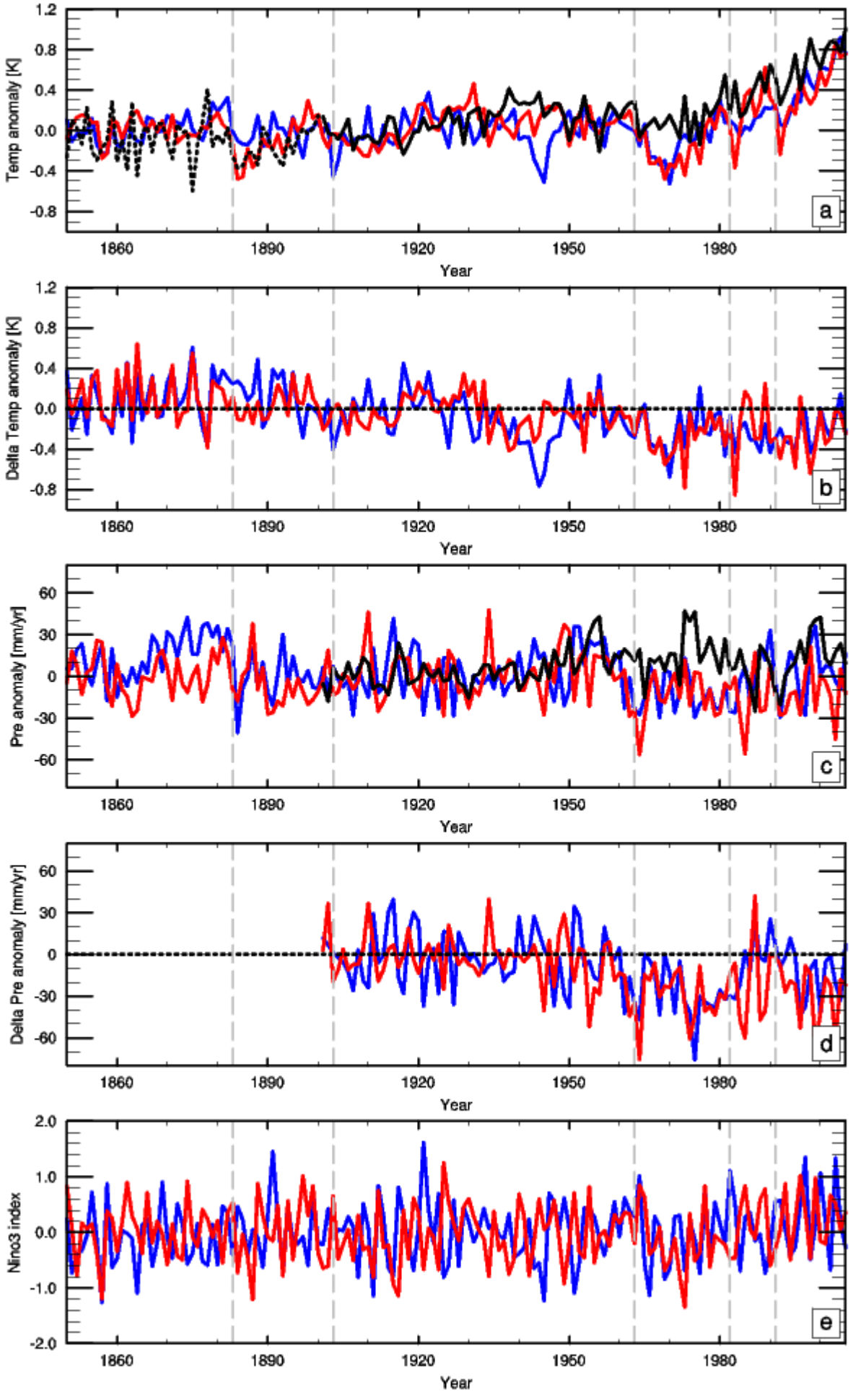
We do not think that a correlation coefficient would be a very meaningful metric to assess the errors between simulated and observed temperature and precipitation. According to Anav et al. (2013), there is no reason to expect models and observations to agree on the phasing of internal interannual variations. We therefore calculate the model variability index (MVI) to analyse the performance of ACCESS-ESM1 for temperature and precipitation. The MVI compares the models variability at every grid cell, which is then averaged for the globe. Perfect model – observation agreement would result in an MVI value of 0. For example, for temperature we calculate an MVI of 0.3 (PresLAI) and for precipitation an MVI of 1.7 (PresLAI) over the period 1901-2005. We will include this information in a revised version of the paper:

The interannual variability in temperature is well reproduced by both ACCESS-ESM1 scenarios, showing an MVI of 0.3 (PresLAI) and 0.4 (ProgLAI) for the period 1901-2005.

According to Anav et al. (2013) only a few CMIP5 models show an MVI of lower than 0.5 (although their calculation is based on present day, i.e. 1986-2005).

For precipitation we calculate an MVI of 1.7 (PresLAI) and 1.8 (ProgLAI) for the period 1901-2005, which suggests that the IAV is not well represented in ACCESS-ESM1. However, according to Anav et al. (2013) none of the CMIP5 models had an MVI close to the threshold of 0.5. Also note that for the calculation of the MVI for precipitation we had to exclude 60 land points (mainly coastal points) due to inconsistencies in the regridding.

Anav, A., Friedlingstein, P., Kidston, M., Bopp, L., Ciais, P., Cox, P., Jones, C., Jung, M., Myneni, R., and Zhu, Z.: Evaluating the Land and
Ocean Components of the Global Carbon Cycle in the CMIP5 Earth System Models, J. Climate, 26, 6801–6843, 2013.


Temp anomaly [K]
1.2
0.8
0.4
0.0
-0.4
-0.8
a
1860
1890
1920
1950
1980
Year
Delta Temp anomaly [K]
b
Pre anomaly [mm/yr]
60
30
0
-30
-60
c
Delta Pre anomaly [mm/yr]
d
Nino3 index
2.0
1.0
0.0
-1.0
-2.0
e

The last paragraph in this section discusses the timing of precipitation anomalies versus volcanic eruptions. The authors point out a reduction in precipitation following eruptions, with the one exception of El Chichon; however, when I looked at the time-series, I did not see a decrease after the Santa Maria eruption in addition to El Chichon. In fact, looking at the time-series, the decrease after volcanic eruptions did not stand out to me, especially when the authors note the reduction following the Mt. Pinatubo eruption, which does stand out, is too far away from the eruption date to be related. I think a more quantifiable analysis of the magnitude and the timing of this decrease, in days or weeks or some stated time-scale, would be helpful.

We agree with the referee that there is no decrease in precipitation visible following the eruption of Santa Maria. We will include this statement in a revised version of the manuscript.

However, we did not say that a reduction in rainfall following Mt. Pinatubo is too far away from the eruption date. In fact, along with Krakatao and Mt. Agung, the Mt Pinatubo event shows a significant reduction in precipitation anomalies immediately after the eruption.

<u>Sea Surface Temperature and Mixed Layer Depth</u>
The second full paragraph of this section discusses spatial patterns of sea surface temperatures, except no time period for the analysis is provided, not even in the caption for Figure 3. Are these differences over the entire simulation or a selected time period? Shifting to the text, the authors state that ACCESS-ESM1 "produces very heterogeneous differences from observations." Reading this, it was unclear to me what was meant. When I turned to look at the figure I expected to see random errors; however, in my opinion the differences are spatially coherent in latitude bands. Then in this same discussion they state "there do not appear to be strong seasonal biases," but with this terminology I have to wonder are there or aren't there? Then they state the exception of the North Atlantic, which has a coherent bias towards cooler temperatures, but to me it looks like this is a year-round bias more than a seasonal bias. It perhaps does vary in magnitude with season, being a larger bias in August, but it looks like the sea surface temperature here is underestimated year round. Further looking at Figure 3, I also see a flip-flop in errors in the southeastern Pacific Ocean, with positive differences in August and negative differences in February. To me, having differences that vary with time of year makes it seasonal, but this region is not mentioned. I think this section could benefit from more careful wording and analysis.

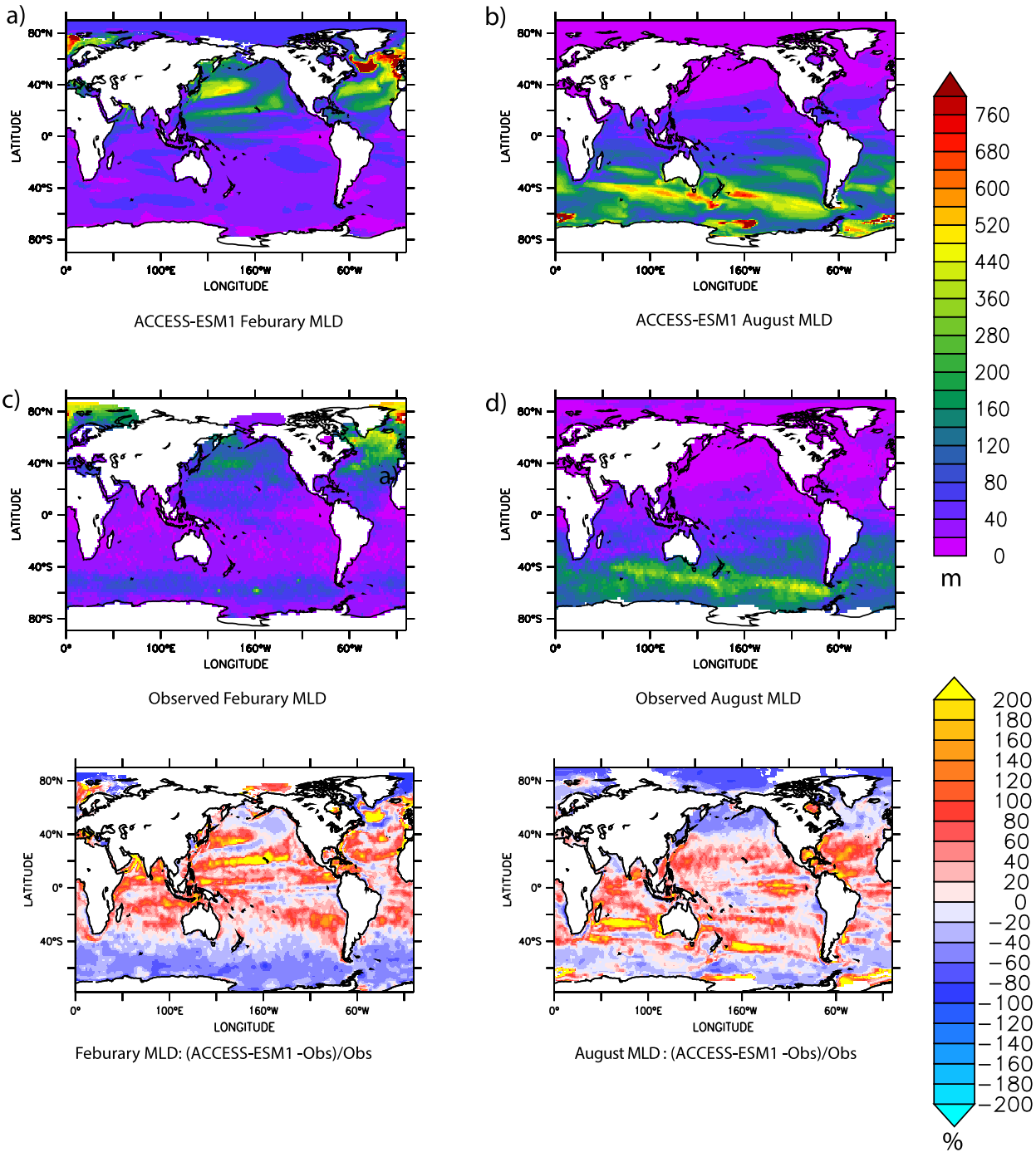
We apologies for the oversight and have now stated in the caption and text that the time period of the figures is the IPCC historical period 1986-2005. We have rewritten the text to better convey our intended meaning and analysis. Indeed the spatial patterns of the warming are perplexing better upon a much closer examination appear to be associated with the Met Office Unified Model (MetUM) which has known biases and which ACCESS-ESM1 utilizes as it atmospheric model, indeed similar biases are seen in HadGEM2 that also use the MetUM.

The ocean mixed layer depth discussion could also benefit from a more quantitative analysis, rather than using statements like "appears to slightly overestimate the

depth in winter" and "appears to underestimate the depths in summer," with no statistical support for these subjective comments. Also, in Figure 4 in this section, the caption states that differences are shown, suggesting a difference plot; however, the figures just show the results from both the model and the observations. I would suggest rewording the caption to avoid confusion.

To address the reviewer's comments we have rewritten the section on mixed layer depth comparison, and added an additional figure showing the percentage changes in mixed layer depth between the observations and ACCESS-ESM1. This allows the changes between the obs and model to be quantified relative to the total observed mixed layer depth. We have also updated the caption to reflect the reviewers concerns.

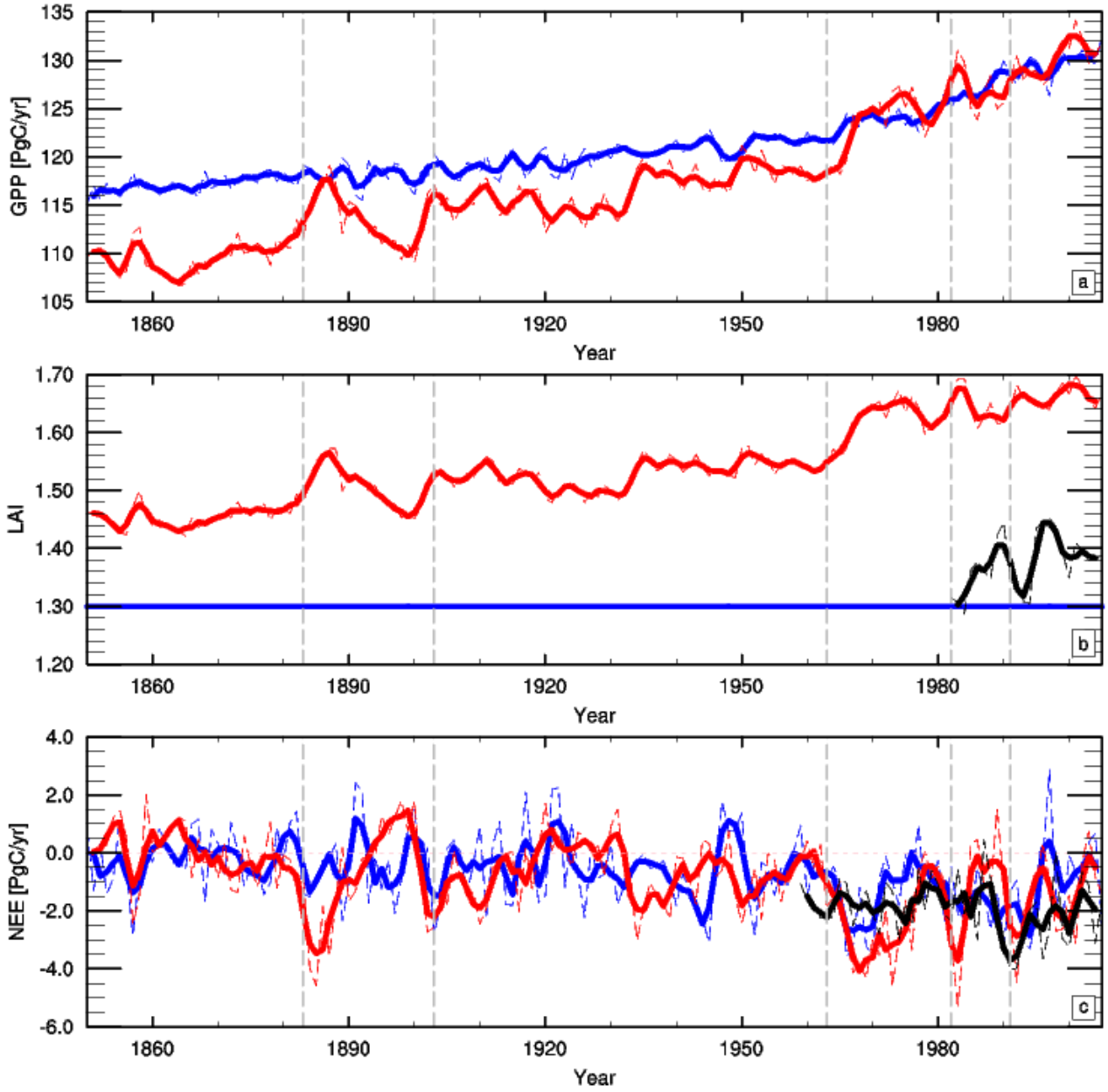
In the higher latitude that the winter mixed layers are well captured by ACCESS-ESM1 Figure 4. This is encouraging given that many ocean models tend to underestimate winter mixed layer depths (Sallee et al (2013) and Downes (2015)). Simulating winter mixed layers correctly is critical for setting interior ocean properties supplying nutrients to the upper ocean to fuel the biologically active growing season (Rodgers et al, 2014). However in contrast to the winter, ACCESS-ESM1 appears to systematically underestimate mixed layer depths in the high latitude ocean in summer, ~60% (or 30-40m) in the Southern, Pacific and Atlantic Oceans. In the Southern Ocean, in particular, the underestimation of summer mixed layer depths is consistent with Sallee et al (2013) and Haung et al (2013) who showed that most CMIP5 models underestimate summer mixed layer depths. Haung et al (2013) attributed this to a lack of vertical mixing in CMIP5 rather than sea surface forcing related to individual models, this is consistent with Downes et al (2015) who showed that these biases are also present in the ocean only simulations of ACCESS-ESM1 (ACCESS-O).



Land Carbon Response

I think it would be helpful to include the MODIS/AVHRR LAI data in Figure 5b. I realize this would only be for the last few years of the simulation, but the benefit is that it would provide a reference to the simulated LAI values. After reading it through, I discovered there is a section on LAI where it is discussed in more detail, but this figure comes first, and when I read it I was wondering how it compared.

Observation based LAI data (MODIS/AVHRR) have been included in Fig. 5b for the period 1982 to 2005. For comparison we have also included the prescribed LAI used in ACCESS-ESM1, which has no interannual variability. The updated figure will be included in a revised version of the paper



I found the discussion in the last two paragraphs of this section confusing. I loved to see values and uncertainties; however, it was unclear to me what values were comparable. The section starts with a discussion on carbon uptake, which at first I assumed represents total uptake, or GPP. But from what I could tell, the same values provided for ACCESS-ESM1 were then used in the following paragraph, which talks about NEP. I read the section several times, but only found the one value of 154 Pg C for prognostic ACCESS-ESM1. I was then confused when it was stated that this value is "at the low end of the CMIP5 range," when that range is estimated to be from -59 to 18 PgC according to Shao et al. (when outliers are not included) or from

-124 to 50 PgC from Jones et al. Based on these numbers, doesn't ACCESS-ESM1 take up and store much more carbon (+154 Pg C) than the CMIP5 models? It doesn't help that the signs make this analysis even more confusing. Since NEP and uptake were discussed, I assumed a negative value was a source of carbon. I know this confusion on signs is difficult to handle, but I just wanted to raise awareness that it contributed to making this section difficult to follow. I apologize if I got these comparisons incorrect, but that indicates that more careful discussion and use of terminology would be helpful.

The section "Land carbon response" discusses the impact of the historical forcing on some carbon related variables, i.e. gross primary production (GPP), leaf area index (LAI) and net ecosystem exchange (NEE) and their interactions.

We actually do not discuss absolute values for GPP in this section, we focus mainly on interannual variability (IAV) and trend. Absolute values of GPP (i.e. mean GPP for present day) are discussed and compared against observations in section 5.1.1. "GPP".

The last two paragraphs analyse the total land carbon uptake over the historical period, which is the sum of the net ecosystem production NEP (opposite sign to NEE, i.e. NEP = -1 x NEE) from 1850 to 2005. Throughout the paper we consistently analyse the flux to the atmosphere (i.e. land to air and sea to air) which is commonly used for analysing CMIP5 modelling results. However, in order to calculate the uptake by land and ocean we need to reverse the sign. We will clarify this in a revised version of the paper.

The value of 154 PgC represents the total land carbon uptake over the historical period for the scenario with prognostic LAI (i.e. cumulative NEP). Note, this value will be corrected in a revised version (see also reply to reviewer 1).

We currently do not consider disturbances such as land use and land cover change (LULCC) in our simulations, which means that our land carbon uptake is simply calculated based on NEP. The majority of CMIP5 models include LULCC in some form or the other, which makes it difficult to compare our calculated uptake against land carbon uptake from CMIP5 models. We tried to do this in two ways:

(a) we compare our results against cumulative NEP with values reported in Shao et al. (2013) with NEP ranging from 24 to 1730 PgC, which means ACCESS-ESM1 is at the lower end of this range.

(b) we compare our results against observational based estimates of land use emissions which are thought to be 108-188 PgC for the historical period. This means we get an almost neutral behaviour by accounting for LULCC in this way. CMIP5 models that include disturbances also estimate a neutral behaviour by providing an estimate of land carbon uptake of -59 to 18 PgC (Shao et al., 2013) and -124 to 50 PgC (Jones et al., 2013).

We will revise the whole section accordingly.

<u>Ocean Response</u>

I thought the second half of the discussion in this section was clear and informative; however, I had a two questions on Figure 6. First off, the caption states that it's "Integrated Primary Production," but doesn't define what that is. I assume that's the same as NPP? I'm unfamiliar with that terminology, so it might be worth

clarification. Second, the values in the text and the figures don't match up. In the text the global mean ACCESS-ESM1 NPP is 46 PgC/yr, but from what I can tell this must be for the entire simulation. This is then compared to SeaWIFS value of 52 PgC/yr for 1998- 2005. Upon reading this, I expect ACCESS- ESM1 to be lower than SeaWIFS in Figure 6; however, looking at it, ACCESS-ESM1 is higher than SeaWIFS for these years. I personally think it would be better if these comparisons were values representing the same time period, both to have a fair comparison and to match up the values with what is seen in Figure 6.

We apologise for any inconsistency, it should have been 51 PgC/yr, and we have now clarified the text to replace integrated primary production with net primary production and added a section explaining how NPP is calculated.

GPP

I would consider removing the first sentence in the second paragraph, as it is subjective and is not needed with the supporting text. I would then combine the first and second paragraphs to have a complete discussion. I would also remove or modify the final sentence in the second paragraph that states that containing nitrogen and phosphorous "ensures a more realistic simulation." While I think there is evidence that including nitrogen and phosphorous is beneficial in many circumstances, that alone does not ensure a model outperforms one where these are not included but instead has more realistic representations of other important processes.

As suggested by the reviewer we will remove the first sentence of the second paragraph and combine the first two paragraphs in a revised version. We will change the final sentence of the second paragraph to:

ACCESS-ESM1 contains both nitrogen and phosphorus limitation, which may provide a more realistic simulation of carbon cycle uptake by the terrestrial biosphere.

The second section in this section discussing the mean annual cycle of GPP is again quite subjective, and perhaps it wouldn't be too difficult to provide a few quantitative statements.

We will include more quantitative statements in the discussion of the mean annual cycle of GPP in a revised version.

For the final discussion in the section on IAV, first off I was wondering how IAV is calculated? I can think of a couple of different methods, and it is not defined how they actually calculate the values that are shown. For this section, I suggest a PDF of errors, therefore when you say there is good agreement in the spatial pattern, it can be backed up with "x% of the globe has errors less than x kg C/m2." I will also note that the labels in Figure 9 were particularly hard to read.

The interannual variability (IAV) is calculated as the standard deviation for the de-trended annual mean values. This explanation will be included in a revised version.

As suggested by the reviewer we have calculated the absolute error for present day mean GPP for each land grid point. For example, 95% of all land points have errors smaller than 0.5 kgC/m2/yr for the scenario with prescribed LAI (86% for the scenario with prognostic LAI). We will include those numbers in a revised version.

We will increase the size of labels in Fig. 9 in a revised version.

CNP Pool Sizes

The discussion on the HWSD and soil carbon I found to be quite subjective, again focusing on the comparisons being "good" or "generally good."

For the nitrogen comparison, first off please clarify what the value reported, 85 Pg N, represents (i.e. global over entire simulation?) And just a thought: it might be interesting to show or state how this evolves in time, similar to the time-series shown for carbon.

The 85 PgN represent the mean soil organic (SOC) pool size for the last 20 years of the historical period (1886-2005). In the manuscript on page 13, line 2 we stated that all pool sizes are calculated over the last 20 years of the historical period.

We will include initial pool sizes (i.e. spun up pools from pre-industrial simulation) for CNP for both scenarios in a revised version in Table 2 so that they can be compared against present day pool sizes:

	Pre-industrial						Present day						Historical change C	
	PresLAI			ProgLAI			PresLAI			ProgLAI			PresLAI	ProgLAI
Pool	C	N	P	C	N	P	C	N	P	C	N	P	ΔC	ΔC
Biomass	611	5.7	0.31	731	6.15	0.33	670	6.2	0.34	807	6.8	0.37	69.5	87.2
Litter	117	0.85	0.04	149	1.02	0.05	126	0.9	0.05	163	1.1	0.06	7.6	12.3
SOC	1034	82	9.6	1187	86.1	11.9	1050	83.4	10.1	1217	88.5	12.6	20.5	37
∑	1762	88.6	10.0	2067	93.3	12.3	1846	90.5	10.5	2187	96.4	13.0	97.6	136.5

For the phosphorus discussion, I would suggest either removing the "slightly" modifier used in the discussion on how the model results are lower than the estimated range, or give references on how this is smaller than previous modeling estimates to give a frame of reference for that adverb.

We will remove the word "slightly" as suggested by the reviewer in a revised version.

Ocean Carbon

The figures in this section are out of order. I would recommend swapping Figures 12 and 13, rather than discussing Figure 13 first and then going back to 12.

We have now swapped the figures consistent with the suggestion of the reviewer.

Ocean NPP

This section shows and discusses mean seasonal cycles; however, after discussing seasonal aspects such as shifts in timing, the section ends with the disclaimer that the timing cannot be compared. I'm wondering, if that's the case, might it be better to just look at amplitudes in a different format, such as a table or bar graph? If you leave the figures, I would recommend moving this statement to the beginning of the

discussion and making it clear the timing aspects are only being compared between ACCESS- ESM1 and CMIP5. With the figure as it is currently, I again found the accompanying text subjective, which might also support looking into a table or bar graph format for this section to provide more quantitative comparisons.

We regret any misunderstanding; the statement or caveat here refers to the challenge of comparing the response only in the tropical ocean given that ENSO cycles have a strong influence on ocean productivity. We have now reordered the text in this section to make it clearer for the reader and provided more insights into the mechanisms driving the differences between ACCESS-ESM1, observations and CMIP5. We are keen to stick with the plots to highlight the differences in both the magnitude and phase of the seasonal cycle. Additionally, we have also added that the plots and text refer to the observational period 1998-2005, consistent with Anav et al (2013). The manuscript now states:

To assess the seasonal anomaly of Net Primary Production (NPP), calculated as the anomaly of vertically integrated primary productivity through the water column, the global ocean is broken down into 5 regions, following Anav et al (2013). Figure 12 shows the NPP seasonal anomaly from ACCESS-ESM1, CMIP5 models and SeaWIFS over the (SeaWIFS) observational period 1998-2005. At the global ocean scale, seasonally we see that the magnitude of NPP from ACCESS-ESM1 is less than the amplitude of CMIP5 and SeaWIFS, with poor phasing. This likely reflects the biases in ACCESS-ESM1 toward lower latitudes, reflecting excess nutrient supply, and utilization, to the upper oligotrophic ocean (Law et al 2015) associated with deeper than observed mixed layers.

In the northern and southern subtropical gyres ACCESS-ESM1 (18N-49N and 19S-44S respectively) appears to overestimate the amplitude of the observed seasonal cycle when compared with SeaWIFS. Again this overestimate of NPP is associated with deeper than observed mixed layers which increase nutrient supply to the oligotrophic upper ocean. The phase of the NPP in these regions, where agreement between observations and CMIP5 is very good, is delayed by about three months. This delay may also be explained by a combination of higher (than observed) concentrations of nutrients and slower than expected biological productions associated with cool biases, particularly in the Atlantic Ocean allowing the bloom to occur later.

In the high latitude northern hemisphere, the magnitude of the seasonal cycle of NPP is not well captured in ACCESS-ESM1. While CMIP5 appears also to underestimate the magnitude of the seasonal cycle, ACCESS-ESM1 is lower again. In contrast in the Southern Ocean the amplitude of the seasonal cycle of NPP in ACCESS-ESM1 shows good agreement with observations. However in the high latitude oceans the phase of NPP is delayed by about 2 months. This delay may be attributed to the too shallow mixed layers that exist in these regions, which means that it is only when mixed layers start to deepen that biological productivity can start to occur. As a result the remaining growing season is shorter (than observed) leading to a reduced total productivity. This may in part explain why the total NPP northern hemisphere is much less than observed.

Interestingly, in the tropical ocean we see very good agreement in the amplitude of the seasonal cycle with CMIP5 and SeaWIFS. We note however, that comparing the phase of the seasonal cycle from ESMs (ACCESS-ESM1 and CMIP5) with SeaWIFS is not very meaningful in this region as they all simulate their own ENSO cycle with their own timing. Therefore any comparison over a 20 year period between models has the potential to be biased by the number of El Nino or La Nina events.

I also want to note on Figure 13 the Southern Ocean has different values on the x- axis, which was not mentioned. This was confusing because at first it appears the Southern Ocean does not have a seasonal cycle, but from what I can tell (which is difficult given the large range), the amplitude is comparable to the other oceanic regions.

Clearly it is desirable to have all of the plots on the same scale. However, in order to highlight all features, different scales cannot be avoided. We have now highlighted in the captions that we use different scales in Fig.13.

Ocean Sea-Air CO2 Fluxes

In this discussion, the authors state that the Southern subtropical gyres overestimate the observed sea- air flux; however, when I look at the figure, it looks to me the biggest uptake is occurring on the coastlines. I personally would like more of a mention of this, and maybe a discussion on why coastlines, particularly off the west coast of Australia, are taking up so much carbon compared to the broader ocean gyres.

I think that role of the coastal ocean in the global carbon cycle is very interesting and has been suggested by authors such as Borges et al (2005). However coastal oceans in ESMs is not well represented in terms of the key mechanisms in CMIP5 models (Bopp et al, 2013). Nevertheless, despite the claims of some authors e.g. Borges et al (2005) that the coastal oceans are playing a very large role, the role likely remains small relative to the gyres, primarily because the surface area of the coastal ocean remains very small relative to the size of the ocean gyres. Furthermore, the heterogeneous response of the coastal ocean means that observationally the response coastal ocean remains more poorly constrained than the ocean. Nevertheless, as more observations are collected, resolution improves in ESMs allowing a better the representation of the key processes in the coastal ocean well better it is likely these models will focus more on these regions, as this is where the impacts of climate change and variability are felt most acutely.

Borges, A. V., Bruno Delille, and M. Frankignoulle. "Budgeting sinks and sources of CO2 in the coastal ocean: Diversity of ecosystems counts." *Geophysical Research Letters* 32.14 (2005).

In the second paragraph of this section, the second sentence beginning with "Furthermore it appears that globally…." doesn't make sense to me, as I don't understand what it is that lies outside the range. Usually regional analyses reveal why global results occur, and it looks like the Southern Hemisphere is the main contributor to the ACCESS- ESM1 global seasonal flux anomaly? The text does go on to state this, so you may just consider removing or modifying that sentence. Later in that same paragraph, the text states that the Northern Hemisphere has fluxes larger than observed but within the range of CMIP5; however, to me it looks like there are several months when ACCESS- ESM1 is outside the range (i.e. Jan, Feb, May). Since the same disclaimer was put on this paragraph (which again I would suggest moving to the beginning of the paragraph if these figures are kept), I would consider a different format, such as a table or bar graph, or even keeping the seasonal figures but not showing the seasonal cycle in the observations if you don't believe them, and instead using a solid line.

We have rewritten this section to clarify our intended meaning, we have also used the regional plots to explain where and why the spatial biases exist. As above we have also chosen to keep the figures as these are both helpful in our explanation and consistent with previous CMIP5 assessment paper e.g. Anav et 2013. The text now states:

The anomaly of the seasonal cycle of the sea-air CO_2 fluxes was assessed against observations of W13 and CMIP5, shown in Fig 14 for the period 1986-2005. Here we see that ACCESS-ESM1 has larger global amplitude of sea-air CO2 fluxes than observed (W13) and simulated but close to the upper value of the range from CMIP5 models. We also see that globally the phase of sea-air CO2 fluxes is not well captured in ACCESS-ESM1, lying outside the range of the CMIP5 models. To better understand why there are differences between ACCESS-ESM1, CMIP5 and W13 we separate the response of sea-air CO2 into the same regions as for NPP, again following Anav et al (2013).

ACCESS-ESM1 appears to capture well the phase of sea-air CO2 fluxes in the subtropical gyres. In the northern subtropical gyre in particular, we see that the amplitude and phase of the seasonal cycle in ACCESS-ESM1 shows very good agreement with W13, in contrast with other ESMs (CMIP5). In the southern subtropical gyres, while the ACCESS-ESM1 appears to overestimate the amplitude relative to the observations, we see very good agreement with CMIP5 models. As anticipated the tropical ocean shows very little seasonality, nevertheless we do see good agreement with CMIP5 models. However, the comparison of ACCESS-ESM1 against observations (while shown) is not very meaningful as W13 is based on values of oceanic pCO2 from Takahashi et al, 2009 which does not include El Nino years.

The largest differences are seen in the representation of sea-air CO2 fluxes in the high latitude ocean. In the high latitude northern hemisphere, we see that the magnitude is larger than either CMIP5 or W13 and shows poor phasing. While the magnitude of the seasonal cycle in the Southern Ocean lies within the upper range of CMIP5 again poor phasing is seen. That the seasonal cycle is out of phase suggests that during the summer the solubility response likely dominates over the NPP response, leading to an out-gassing in the summer and uptake in the winter, as discussed in Lenton et al (2013). Consequently, we see that the poor global phasing in global sea-air CO2 fluxes is likely due to the solubility dominated response of the high latitudes during the summer.

<u>Anthropogenic Inventory</u>

The discussion is again quite subjective. Also, in Figure 15, there is a reference to Key et al. (2004) that is confusing and not in the text (I'm unsure if the figure appears in both papers?) Upon finishing the paper, I see the reference in the conclusions, but you might want to somehow clarify or include it in the text during this discussion.

Key et al (2004) refers to the data used in the paper to compare with ACCESS-ESM1 while the estimated anthropogenic CO_2 uptake is from Sabine et al (2004). We have ensured that the correct and appropriate references are used in the figure caption, section 5.2.3 and the conclusions.

Atmospheric CO2

I think it would be better to combine the first paragraph with the next sentence following it
(starting "Therefore, our…").

We will do this in a revised version.

Conclusions

I think the conclusions would be stronger if some quantitative assessments were provided. I also noticed the order of the conclusions was not the same as the order presented in the text, but neither was it combined in a more succinct fashion to group conclusions. Simply reordering the paragraphs would make the text more consistent.

We will reorder the paragraphs in a revised version and also include some quantitative assessments.

In the second- to- last paragraph, I was confused by the statement that "Seasonally the ACCESS- ESM1 appears biased toward the Southern Hemisphere." I'm not sure how to interpret this statement, as to me it's using both temporal and spatial references combined in an unclear fashion. This same type of statement occurs later in the sentence "Globally the annual mean is well captured but biased to low latitudes." I recommend reworking those sentences to clarify the intent of the text. Also, a quantification of how it is "well captured" would strengthen the statement.

We have now rephrased this paragraph to better reflect our intended meaning and removed the very qualitative comments. We have not added numbers to the conclusions but instead have added more quantitative analysis to the text that is referred to in the conclusion, the text now states:

Globally integrated sea-air CO2 fluxes are well captured and we reproduce very well the cumulative uptake estimate from the Global Carbon Project (LeQuere et al, 2014) and our anthropogenic uptake agrees very well with observed GLODAP value of Sabine et al (2004). The spatial distribution of sea-air CO2 fluxes is also well reproduced by CMIP5 models and observations. At the same time global ocean NPP also shows good agreement with observations and lies well within the range of CMIP5 models. However seasonal biases do exist in sea-air CO2 fluxes and NPP, potentially related to biases in mixed layer depth and surface temperature that are present in ACCESS-ESM1; and will need to be addressed in later versions of ACCESS-ESM1.

Technical Comments

Overall I have very few technical comments. My main technical comments are on comma usage. I personally find well- placed commas can aid in reading the text, but am not a comma expert, so feel free to take or leave the suggestions.

All technical comments will be considered and corrected in a revised version.

The carbon cycle in the Australian Community Climate and Earth System Simulator (ACCESS-ESM1). 2. Historical simulations

Tilo Ziehn[1], Andrew Lenton[2], Rachel M. Law[1], Richard J. Matear[2], and Matthew A. Chamberlain[2]

[1]CSIRO Oceans and Atmosphere, PMB 1, Aspendale, Victoria, Australia
[2]CSIRO Oceans and Atmosphere, Hobart, Tasmania, Australia

Correspondence to: Tilo Ziehn (tilo.ziehn@csiro.au)

Abstract. Over the last decade many climate models have evolved into earth system models (ESMs), which are able to simulate both physical and biogeochemical processes through the inclusion of additional components such as the carbon cycle. The Australian Community Climate and Earth System Simulator (ACCESS) has been recently extended to include land and ocean carbon cycle components in its ACCESS-ESM1 version. A detailed description of ACCESS-ESM1 components including results from pre-industrial simulations is provided in Part 1. Here, we focus on the evaluation of ACCESS-ESM1 over the historical period (1850-2005) in terms of its capability to reproduce climate and carbon related variables. Comparisons are performed with observations, if available, but also with other ESMs to highlight common weaknesses. We find that climate variables controlling the exchange of carbon are well reproduced~~, however~~ . However, the aerosol forcing in ACCESS-ESM1 is somewhat ~~over-sensitive to anthropogenic aerosols~~ larger than in other models, which leads to an overly strong cooling response in the land from about 1960 onwards. The land carbon cycle is evaluated for two scenarios: running with a prescribed leaf area index (LAI) and running with a prognostic LAI. We overestimate the seasonal ~~amplitude~~ mean (1.7 vs. 1.4) and peak amplitude (2.0 vs. 1.8) of the prognostic LAI at the global scale, which is common amongst CMIP5 ESMs. However, the prognostic LAI is our preferred choice, because it allows for the vegetation feedback through the coupling between LAI and the leaf carbon pool. ~~Globally~~ Our globally integrated land-atmosphere ~~and~~ flux over the historical period is 98 PgC for prescribed LAI and 137 PgC for prognostic LAI, which is in line with estimates of land-use emissions (ACCESS-ESM1 does not include land-use change). The integrated ocean-atmosphere ~~fluxes and flux patterns are well reproduced and show good agreement with most recent observations.~~ flux is 83 PgC, which is in agreement with a recent estimate of 82 PgC from the Global Carbon Project for the period 1959 to 2005. The seasonal cycle of simulated atmospheric CO_2 is close to the observed seasonal cycle (up to of 1 ppm difference for station at Mace Head and up to 2 ppm for station at Mauna Loa), but shows a larger amplitude (up to 6 ppm) in the high northern latitudes. Overall, ACCESS-ESM1 performs well over the historical period, making it a useful tool to explore the change in land and oceanic carbon uptake in the future.

1 Introduction

Climate models are continuously evolving to include more processes and interactions at higher resolutions and their number has increased rapidly in recent years. In addition, a number of institutes worldwide have been developing earth system models

(ESMs), which are able to simulate both physical and biogeochemical processes ~~,~~through the inclusion of the land and ocean carbon cycles.

The evaluation of ESMs in terms of their capability to reproduce climate and carbon related variables over the historical period (i.e. 1850 to 2005) is crucial prior to using such models for future predictions. Comparisons are usually performed with observation based products, if available, but also with other ESMs to identify common weaknesses.

The performance of 18 ESMs that participated in the Coupled Model Intercomparison Project phase 5 (CMIP5) (Taylor et al., 2012) has been evaluated in Anav et al. (2013) for the present day climate. They found that all models correctly reproduce the main climate variables controlling the spatial and temporal variability of the carbon cycle. However, large differences exist when reproducing specific fields. In terms of the land carbon cycle, an overestimation of photosynthesis and leaf area index (LAI) was found for most of the models. In contrast, for the ocean an underestimation of the ~~primary production~~ net primary production (NPP) was noted for a number of models. Anav et al. (2013) also found significant regional variations in model performance.

Eight of these CMIP5 ESMs were also evaluated in Shao et al. (2013), highlighting that temporal correlations between annual-mean carbon cycle and climate variables vary substantially among the 8 models. Large inter-model disagreements were found for ~~net primary production (NPP)~~ NPP and heterotrophic respiration (Rh). In agreement with Anav et al. (2013), Shao et al. (2013) also noted that the CMIP5 historical simulations tend to overestimate photosynthesis and LAI.

Todd-Brown et al. (2013) compared and evaluated 11 CMIP5 ESMs in terms of their variations in soil carbon. The correct representation of soil carbon in the model is important in order to accurately predict future climate-carbon feedbacks. Soil carbon simulations of the 11 models were compared against empirical data from the Harmonized World Soil Database (HWSD) and from the Northern Circumpolar Soil Carbon Database (NCSCD). A large spread across all models was found (nearly 6 fold) and the spatial distribution of soil carbon, especially in the northern latitudes was found to be poor in comparison to HWSD and NCSCD, which means that most ESMs were poorly representing grid-scale soil carbon.

Frölicher et al. (2015) showed that CMIP5 models appeared to capture the observed pattern of anthropogenic carbon storage in the ocean, particularly in the Southern Ocean. However, overall they underestimate the magnitude of the observed oceanic global anthropogenic carbon storage since the pre-industrial.

The representation of the global carbon cycle in ESMs continues to be challenging. For example, large uncertainties exist for the climate-carbon feedback, which can be mainly attributed to terrestrial carbon cycle components (Friedlingstein et al., 2006; Arora et al., 2013). Terrestrial ecosystem models show large variations when driven with future climate scenarios (Shao et al., 2013; Friend et al., 2014) due to differences in model formulation and uncertainties in process parameters (Knorr and Heimann, 2001; Booth et al., 2012).

The Australian Community Climate and Earth System Simulator (ACCESS) participated in CMIP5, but in a climate model only version. A selection of CMIP5 simulations have now been performed with the ESM version of ACCESS, ACCESS-ESM1 (Law et al., 2015). Here, we present the performance of the land and ocean carbon cycle components of ACCESS-ESM1 over the historical period (1850-2005). First, we briefly assess ACCESS-ESM1 simulation of climate variables that are relevant to the carbon cycle (Sect. 3). We then focus on the response of the carbon cycle to the historical forcing (Sect. 4) and comparison

of various present-day simulated carbon variables with observations (Sect. 5). Law et al. (2015) provides complementary analysis of the ACCESS-ESM1 pre-industrial simulation.

2 Model configuration, simulations and comparison data

Historical simulations (Sect. 2.2) are performed with two model configurations (Sect. 2.1) and the results compared with other CMIP5 ESMs (Sect. 2.3) and a number of observed data products (Sect. 2.4).

2.1 Model configuration

ACCESS-ESM1 is based on the ACCESS climate model (Bi et al., 2013), but with the addition of biogeochemical components for ocean and land as described in part 1 of this paper (Law et al., 2015). The climate model version underlying the ESM version is ACCESS1.4, a minor update of the ACCESS1.3 version submitted to CMIP5 (Bi et al., 2013; Dix et al., 2013). The relationship between the ACCESS1.3, ACCESS1.4 and ACCESS-ESM1 versions is illustrated in Law et al. (2015, Fig. 1). Law et al. (2015) also showed that the climate simulations of the three model versions are very similar.

For the ACCESS-ESM1 version, ocean carbon fluxes are simulated by the World Ocean Model of Biogeochemistry And Trophic dynamics (WOMBAT) (Oke et al., 2013) and land carbon fluxes are simulated by the Community Atmosphere Biosphere Land Exchange (CABLE) model (Kowalczyk et al., 2006; Wang et al., 2011), which optionally includes nutrient limitation (nitrogen and phosphorus) for the terrestrial biosphere through its biogeochemical module, denoted CASA-CNP (Wang et al., 2010). This capability is important because nitrogen, phosphorus and carbon biogeochemical cycles are strongly coupled, and it has been demonstrated that nutrient limitation has a large impact on the productivity of terrestrial ecosystems (Wang et al., 2010; Goll et al., 2012; Zhang et al., 2013). Consequently, global land carbon uptake can be altered significantly. Here we run CASA-CNP in 'CNP' mode with both nitrogen and phosphorus limitation active. This differentiates the ACCESS-ESM1 simulations presented here from other ESM simulations for CMIP5, few of which included nitrogen and none of which included phosphorus.

As in Law et al. (2015), two model configurations are used, differing in their treatment of leaf area index (LAI). LAI is an important variable in climate models for describing the biophysical and biogeochemical properties of the land cover and in CABLE it can either be prescribed or simulated. When prescribed, monthly values based on MODIS observations are read in through an external file (Law et al., 2015, Sec. 3.1.1). The dataset used here is limited by having no interannual or longer time-scale variability. Additionally the same LAI is assigned to all plant funtional types (PFTs) within a grid-cell even though CABLE simulates multiple PFTs per grid-cell. With prescribed LAI there is no coupling between the LAI and the leaf carbon pool which means that vegetation feedbacks cannot be included. These limitations are removed by making LAI a prognostic variable with the LAI dependent on the simulated size of the leaf carbon pool. However if the leaf carbon pool is not well simulated then this would lead to a poor LAI simulation with consequent impacts for the climate simulation.

2.2 Simulations

All experiments are set up as concentration driven simulations, which means that (historical) atmospheric CO_2 concentrations are prescribed as an input to ACCESS-ESM1 and changes in the land and ocean carbon pools do not feed back on to atmospheric CO_2 concentrations following CMIP5 protocols (Taylor et al., 2012).

As noted above we run ACCESS-ESM1 in two configurations, with prescribed LAI (PresLAI) and prognostic LAI (ProgLAI). For PresLAI, the carbon cycle has no impact on the simulated climate whereas for ProgLAI, there is a small impact on the climate through biogeophysical feedbacks related to surface albedo, evaporation and transpiration (Law et al., 2015, Sec. 4.1). The difference in LAI will also have an impact on the land carbon fluxes, whereas the impact on the ocean carbon cycle is negligible, and therefore our analysis of the ocean carbon fluxes focuses only on one scenario (i.e. PresLAI).

Both configurations of ACCESS-ESM1 were run for 1000 years under pre-industrial climate conditions (year 1850) (Law et al., 2015) with the historical simulations starting from year 800 of these control runs. As noted in Law et al. (2015) the net carbon fluxes for land and ocean did not equilibrate to zero. At the end of the control run (i.e. year 800 to 955), global NEE is 0.3 $\mathrm{PgC\,yr^{-1}}$ for PresLAI and 0.08 $\mathrm{PgC\,yr^{-1}}$ for ProgLAI. The net autgassing from the ocean is about 0.6 $\mathrm{PgC\,yr^{-1}}$ at the end of the control run. We take this drift into account when we calculate the net uptake of carbon for land and ocean.

The historical simulations use external forcing for 1850-2005 such as increasing greenhouse gases, aerosols, changes in solar radiation and volcanic eruptions as used in previous ACCESS versions (Dix et al., 2013). For example, the prescribed atmospheric CO_2 increases from 285 ppm in 1850 to 379 ppm in 2005.

Volcanic eruptions in ACCESS-ESM1 are prescribed based on monthly global mean stratospheric volcanic aerosol optical depth (Sato et al., 2002) which is then averaged over four equal-area latitude zones, similar to the way it is done in the Hadley Centre Global Environmental Model (HadGEM) (Stott et al., 2006; Jones et al., 2011). Globally significant volcanoes within the historical period are Krakatoa (1883), Santa Maria (1903), Agung (1963), El Chichón (1982) and Pinatubo (1991). Tropospheric aerosols are either calculated interactively (i.e. sea salt and mineral dust) or are based on emission datasets (i.e. sulphate and organic carbon) and increase rapidly from 1950 (Dix et al., 2013, Fig. 4).

The simulations do not include any land-use change; the distribution of PFTs used in the pre-industrial simulation is used throughout the historical period.

2.3 Comparison with CMIP5 models

ACCESS-ESM1 is compared against other ESMs that participated in CMIP5 and are available on the Earth System Grid. The models used in this paper are shown in Table 1 with the references provided in Lenton et al. (2015). As not all years were available for these simulations, we focused on the period 1870-2005 and used only the first ensemble member for each ESM. In assessing the response of the CMIP5 models, we calculated the median and the 10th and 90th percentiles following Lenton et al. (2015). This allows us to both assess how well ACCESS-ESM1 captures the median and whether it falls into the range of existing CMIP5 models.

2.4 Observations

We use the following observational data products to compare against ACCESS-ESM1 outputs. Climate variables are assessed, where this is helpful for interpreting the carbon simulation. For example, the land carbon balance is mainly controlled by surface temperature and precipitation (Piao et al., 2009), whereas the ocean carbon balance is mainly influenced by sea surface temperature (SST) and mixed layer depth (MLD) (Martinez et al., 2009).

Land surface temperature and precipitation: Climate Research Unit (CRU) 1901-2013 time-series (TS) data set at version 3.22 (Harris et al., 2014; Jones and Harris, 2014), statistically interpolated to 0.5° x 0.5° from monthly observations at meteorological stations across the world's land area (excluding Antarctica). A low resolution version at 5° for land surface temperature anomalies (CRUTEM4, (Jones et al., 2012)) is used for the period 1850-1900.

Sea surface temperatures (SST): the high-resolution (1° x 1°) Hadley SST1 (Rayner et al., 2003) in the period 1870-2006.

Climatological mixed layer depths: de Boyer Montégut et al. (2004) for the historical period, based on the density mixed layer criteria of a change density of 0.03 $\mathrm{kg\,m^{-3}}$ from the surface.

Ocean net primary productivity (NPP): from SeaWIFS calculated with the VPGM algorithm of Behrenfeld and Falkowski (1997).

Global ocean and land carbon flux: Global Carbon Project (GCP) estimates of annual global carbon budget components and their uncertainties using a combination of data, algorithms, statistics and model estimates (Le Quéré et al., 2015). The GCP residual land sink is estimated as the difference of emissions from fossil fuel and cement production, emissions from land use and land cover change (LULCC), atmospheric $\mathrm{CO_2}$ growth rate and the mean ocean $\mathrm{CO_2}$ sink. The 2014 global carbon budget (Le Quéré et al., 2015) provides annual values for the period 1959 to 2013.

Gross primary production (GPP): upscaled data from the Flux Network (FLUXNET) using eddy covariance flux data and various diagnostic models (Beer et al., 2010). Gridded data at the global scale is provided by Jung et al. (2011) using a machine learning technique called model tree ensemble (MTE) to scale up FLUXNET observations. Global flux fields are available at a 0.5° x 0.5° spatial resolution and a monthly temporal resolution from 1982 to 2008.

LAI: global LAI derived from the third generation (3g) Global Inventory Modeling and Mapping Studies (GIMMS) normalized difference vegetation index (NDVI)3g data set. Neural networks were trained first with best-quality and significantly post-processed Moderate Resolution Imaging Spectroradiometer (MODIS) LAI and Very High Resolution Radiometer (AVHRR) GIMMS NDVI3g data for the overlapping period (2000 to 2009) to derive the final data set at 1/12° resolution and a temporal resolution of 15 days for the period 1981 to 2011 (Zhu et al., 2013).

Soil organic carbon (SOC): the Harmonized World Soil Database (HWSD) (FAO, 2012) represents the most comprehensive and detailed globally consistent database of soil characteristics that is currently available for global analysis. We use an upscaled and regridded version of the HWSD with the area weighted SOC calculated from the soil organic carbon (%), bulk density and soil depth (Wieder et al., 2014).

Sea-air $\mathrm{CO_2}$ *fluxes:* seasonal climatology of Wanninkhof et al. (2013) based on the 1° x 1° global measurements of oceanic $\mathrm{pCO_2}$ of Takahashi et al. (2009).

Anthropogenic carbon uptake: column inventory estimated from Sabine et al. (2004) from GLobal Ocean Data Analysis Project (GLODAP) (Key et al., 2004).

Atmospheric CO_2 concentrations: mean atmospheric CO_2 seasonal cycles derived from NOAA/ESRL flask samples as processed in the GLOBALVIEW (GLOBALVIEW-CO_2, 2011) data product. These seasonal cycles are designed to be representative of background, clean-air at any given location. Here, we assess the seasonal cycle for 4 locations with an averaging period of about 20 years for Mace Head (53.33° N, 9.90° W), about 25 years for Alert (82.45° N, 62.52° W), about 35 years for South Pole (89.98° S, 24.80° W) and about 40 years for Mauna Loa (19.53° N, 155.58° W).

~~*Sea surface temperatures (SST):* the high-resolution (1x 1) Hadley SST1 (Rayner et al., 2003) in the period 1870-2006~~

2.5 Performance evaluation

For climate variables such as land surface temperature and precipitation we calculate the model variability index (MVI) (Gleckler et al., 2008; Scherrer, 2011). The models (mod) variability at every grid point i is compared against the observerd (obs) variability and then averaged over the globe in the following way:

$$\mathrm{MVI} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{s_i^{mod}}{s_i^{obs}} - \frac{s_i^{obs}}{s_i^{mod}}\right)^2, \quad (1)$$

where s is the standard deviation and n the number of grid cells. Perfect model - observations agreement would result in an MVI of zero. The definition of a limit to decide if a model performs well or poor is rather arbitrary. However, Scherrer (2011) and Anav et al. (2013) have used a threshold of MVI< 0.5.

~~*Climatological mixed layer depths:* de Boyer Montégut et al. (2004) for the historical period, based on the density mixed layer criteria of a change density of 0.03 from the surface~~For a number of carbon related variables we calculate the inter-annual variability (IAV), defined as the standard deviation of detrended annual mean values.

~~*Anthropogenic carbon uptake:* column inventory estimated from Sabine et al. (2004) from GLobal Ocean Data Analysis Project (GLODAP)(Key et al., 2004).~~

~~*Sea-air fluxes:* seasonal climatology of Wanninkhof et al. (2013) based on the 1x 1global measurements of oceanic of Takahashi et al. (2009).~~

~~*Ocean net primary productivity:* net primary productivity from SeaWIFS calculated with the VPGM algorithm of Behrenfeld and Falkow~~

3 ACCESS-ESM1 climatology

3.1 Land temperature and precipitation

Carbon fluxes across the historical period will be directly influenced by increasing atmospheric CO_2 and indirectly influenced by changes in the climate, driven by the increasing atmospheric CO_2 and modulated by other external ~~forcing~~forcings, such as anthropogenic and volcanic aerosols. In addition, each climate simulation generates its own internal variability, with major

modes of climate variability such as the El Niño Southern Oscillation (ENSO) known to generate large variability in carbon exchange between the atmosphere and both the ocean and land (Zeng et al., 2005).

The evolution of temperature and precipitation in ACCESS-ESM1 (Fig. 1) over land shows similar characteristics to ACCESS1.3 historical simulations (Dix et al., 2013; Lewis and Karoly, 2014) as well as those of ACCESS1.4 (P. Vohlarik, pers. comm.). Global land surface air temperature anomalies (relative to 1901-1930) are shown in Fig. 1. Both ACCESS-ESM1 simulation scenarios (PresLAI and ProgLAI) show similar temperature anomalies over most of the historical period, being close to the observed anomalies through most of the period ~~but somewhat lower than observations from~~ (decadal mean difference smaller than 0.2 K), apart from the 1940s where the PresLAI scenario shows a larger negative anomaly (decadal mean difference of about 0.37 K), which will be discussed later. From about 1965-2005 anomalies are by up to 0.4 K (decadal mean difference) lower than observations for both scenarios. This is attributed by Lewis and Karoly (2014) to a likely overly strong cooling response in ACCESS1.3 to anthropogenic aerosols, offsetting the warming due to greenhouse gas increases for which ACCESS1.3 responds similarly to a CMIP5 mean (Lewis and Karoly, 2014, Figs. 2a, 3a). Strong aerosol cooling is supported by Rotstayn et al. (2015) who found that ACCESS1.3 showed a large global mean aerosol effective radiative forcing (ERF) over the historical period of $-1.56\,\mathrm{W\,m^{-2}}$ which is much larger than the IPCC best estimate ($-0.9\,\mathrm{W\,m^{-2}}$) (Boucher et al., 2013) but still within the uncertainty range.

The interannual variability in temperature is well reproduced by both ACCESS-ESM1 scenarios, showing an MVI of 0.3 (PresLAI) and 0.4 (ProgLAI) for the period 1901-2005. According to Anav et al. (2013) only a few CMIP5 models show an MVI of lower than 0.5 (although their calculation is based on present day, i.e. 1986-2005).

Both ACCESS-ESM1 simulations exhibit cooling following major volcanic eruptions (marked in Fig. 1). At first sight, the ProgLAI run seems to be more sensitive to volcanic eruptions, showing a stronger cooling particularly for the two most recent major eruptions, El Chichón in 1982 and Mt. Pinatubo in 1991. However, this difference might be due to a different ENSO phase for the two runs at the time of the eruptions. Lewis and Karoly (2014) assessed the temperature impact of Agung, El Chichón and Pinatubo in three ACCESS1.3 simulations (e.g. their Fig. 7) and mean temperature anomalies from the two ACCESS-ESM1 simulations lie within or only slightly outside the ACCESS1.3 ensemble range. It is worth noting that Lewis and Karoly (2014) found that the simulated temperature anomalies from volcanoes tended to be larger in ACCESS than observed, and this was common across CMIP5 models.

Differences in the year to year temperature anomalies between the two ACCESS-ESM1 scenarios are likely due to internal climate variability. For example, between the years 1940 and 1950, the PresLAI run shows a large negative temperature anomaly and the ProgLAI run shows a positive anomaly. The negative anomaly for the PresLAI is probably related to a strong La Niña event (Nino3 index of -1.2) around the year 1945 (Fig. 1c), whereas in the ProgLAI case we see a small El Niño event (Nino3 index of 0.6) around the same time.

The temperature anomalies hide an absolute temperature difference between the two ACCESS-ESM1 simulations; the ProgLAI scenario produces a slightly warmer climate (0.56 K difference in mean land surface air temperature averaged over 1850-2005) than the PresLAI run. This is consistent with the difference in surface air temperature found for the pre-industrial simulations (Law et al., 2015, Sec. 4.1). As noted in Law et al. (2015) the warmer climate can be explained by the difference in

LAI, which is generally higher in the prognostic case. This leads to a lower albedo, especially for evergreen needleleaf forests during the winter months in the northern hemisphere, and consequently to an increase in absorbed radiation. The difference in LAI for both scenarios is explored in more detail in section 5.1.2. Compared to the observations the ACCESS-ESM1 runs show a cooler land surface air temperature by about 0.5 K for the ProgLAI scenario and 1.1 K for the PresLAI scenario averaged over 1901-2005.

Precipitation anomalies over the land are presented in Fig. 1b. Larger differences in the anomalies for the two ACCESS-ESM1 simulations can be observed around the years 1870 to 1880, where the PresLAI scenario shows a positive anomaly and the ProgLAI scenario shows a mainly negative anomaly. The difference over the remaining time period for the two runs is generally small. ACCESS-ESM1 simulations compare well with observed rainfall anomalies until about ~~1950.~~ 1960 (decadal mean difference smaller than 8 $\mathrm{mm\,yr^{-1}}$), with the exeption of the period 1911-1920 for PresLAI (decadal mean difference of about 12 $\mathrm{mm\,yr^{-1}}$) and the period 1951-1960 for ProgLAI (decadal mean difference of about 17 $\mathrm{mm\,yr^{-1}}$). After that, observed anomalies are mostly higher than the simulation results (decadal mean difference of up to 41 $\mathrm{mm\,yr^{-1}}$), a feature also seen in the ACCESS1.3 historical ensemble (Lewis and Karoly, 2014, Fig. 6a). The comparison of absolute rainfall for the two ACCESS-ESM1 scenarios suggests a dryer climate (approx. 20 $\mathrm{mm\,yr^{-1}}$) for the ProgLAI run.

For precipitation we calculate an MVI of 1.7 (PresLAI) and 1.8 (ProgLAI) for the period 1901-2005, which suggests that the IAV is not well represented in ACCESS-ESM1. However, according to Anav et al. (2013) none of the CMIP5 models had an MVI close to the threshold of 0.5. Also note that for the calculation of the MVI for precipitation we had to exclude 60 land points (mainly coastal points) due to inconsistancies in the regridding.

A reduction in precipitation can be observed following the eruption of major volcanoes for both ACCESS-ESM1 scenarios, apart from the 1903 Santa Maria eruption and the 1982 El Chichón eruption, where the PresLAI scenario does not show a strong anomaly and the ProgLAI anomaly is likely too late to be due to the volcano. As for temperature, the precipitation anomalies lie within or close to the ACCESS1.3 ensemble of anomalies presented by Lewis and Karoly (2014, Fig. 9).

3.2 Sea surface temperature and mixed layer depth

To assist in the assessment of responses of the ocean ~~net primary productivity (NPP)~~ NPP and sea-air CO_2 fluxes, the responses of SST and mixed layer depth are first assessed.

The ocean response from ACCESS-ESM1 is compared with the time series of HadiSST v1 (Rayner et al., 2003) ~~.~~ in Figure 2. Here we see~~that the simulated SST does not increase as much as observations over the historical period due~~ , that there is a warm bias in the early part of the historical period. This warm bias in ACCESS-ESM1 is the same as reported by Bi et al. (2013) over the period 1870-1899 in ACCESS 1.3 (0.26 K). ~~Followed the observed acceleration in warming in the~~ In the period ~~1910-1970,~~ ~~we see~~ 1870-1970 we see that the warming of the oceans appears to be less climate sensitive than the observations. However, by the end of the historical simulation (1970-2005) we notice that ACCESS-ESM1 captures well the observed response of HadiSST in the later period~~(1970-2005)~~.

However~~despite little~~ , despite little global bias in the latter period we see that the ACCESS-ESM1 SST response, consistent with ACCESS 1.3 (Bi et al., 2013), produces ~~very heterogeneous~~ strong spatial differences from observations. Fig. 3 shows

that clear spatially coherent differences between ACCESS-ESM1 has a strong warm summer bias in SST at the higher latitudes which is weaker but present in the winter. This summer bias is largest in the Southern Ocean where large values are present (and observations (1986-2005). Some of these regions show a strong summer warming bias (>3 K). Away from the higher latitudes, there does not appear to be strong seasonal biases, with the exception of the subtropical North Atlanticwhich has a coherent biastowards cooler temperatures. K) in areas such as the high latitude Southern and Pacific Ocean, while in other regions such as the subtropical Atlantic, a strong cooling bias is present during the same season. This is in contrast to other regions, such as the high latitude North Atlantic, that has a strong year round warming bias. These biases are broadly consistent with known errors associated with the UK Met Office Unified Model (Williams et al., 2015), which is employed as the atmospheric model in ACCESS-ESM1. Our SST response is also broadly consistent with other ESMs such as HadGEM2 (Martin et al., 2011) that also use the UK Met Office Unified Model.

The magnitude of the interannual variability of simulated SST is a of similar magnitude as the observations. In response to large aerosol injections associated with volcanic eruptions, overlain on Fig. 2, we see that the ocean does capture a net cooling, as expected (e.g. Stenchikov et al., 2009) and consistent with observations. Interestingly, the magnitude of the cooling is sometimes less than observed in HadiSST v1 despite the stronger than observed aerosol response in ACCESS-ESM1.

Ocean mixed layer depths are compared with the observations from de Boyer Montégut et al. (2004)based on following de Boyer Montégut et al. (2004), based on more than 880000 depth profiles from research ships and ARGO profiles, and based on a 0.03 $\mathrm{kg\,m^{-3}}$ density change from the surface. Significant advances in autonomous measurement platforms have allowed the mixed layer to be increasingly well constrained in all seasons across the global ocean.

Overall we see in the mid and lower latitudes that the mixed layer depth is deeper than observed in all seasons (Figure 4shows that ACCESS-ESM1 appears to slightly overestimate the depth of the). However the very large values likely represent the differences in the positions of fronts between the relatively coarse resolution model relative to the observations rather than very large differences (Lenton et al., 2013). In the higher latitudes winter mixed layers . Winter mixed layers close to or deeper than observed are well captured by ACCESS-ESM1 (Figure 4). This is encouraging given that many ocean models tend to underestimate winter mixed layer depths (Sallée et al., 2013; Downes et al., 2015). At the same time Simulating winter mixed layers correctly is critical for setting interior ocean properties supplying nutrients to the upper ocean to fuel the biologically active growing season (Rodgers et al., 2014). However in contrast to the winter, ACCESS-ESM1 appears to systematically underestimate mixed layer depths in the high latitude Southern Ocean in summer. This Southern Oceanunderestimate ocean in summer, 60% (or 30-40 m) in the Southern Ocean, Pacific and Atlantic Oceans. In the Southern Ocean, in particular, the underestimation of summer mixed layer depths is consistent with Sallée et al. (2013) Sallée et al. (2013) and Huang et al. (2014) who showed that most CMIP5 models tend to underestimate summer mixed layer depths. That this summer bias is also seen in the ocean only version of ACCESS Downes2015suggests that this bias may be related to the formulation of mixed layer depth in the ocean model, rather than due solely to the summer warming bias. Little bias is seen in summer mixed layers in the higher latitude Northern HemisphereHuang et al. (2014) attributed this to a lack of vertical mixing in CMIP5 rather than sea surface forcing related to individual models, this is consistent with Downes et al. (2015), who showed that these biases are also present in the ocean only simulations of ACCESS-ESM1.

4 ACCESS-ESM1 carbon cycle response to historical forcing

The increase in atmospheric CO_2 over the historical period is expected to have a direct impact on both land and ocean carbon fluxes. Additionally there may be indirect impacts from the change in climate caused by the increasing atmospheric CO_2. These impacts are explored firstly for land carbon and then for ocean carbon.

4.1 Land carbon response

The direct impact of increasing atmospheric CO_2 is seen clearly in the simulated global land gross primary production (GPP) (Fig. 5a), with increasing GPP for both simulations. The ProgLAI case gives the larger increase, with fluxes for the final 10 years of the simulation being 19% larger than for the first 10 years, compared to an increase of 11% in the PresLAI case. This is due to increasing LAI in the ProgLAI simulation (Fig. 5b) compared to the prescribed LAI which is annually repeating with no increase. Thus the PresLAI case captures only the direct CO_2 fertilisation effect of more efficient photosynthesis per leaf area while the ProgLAI case also allows the growing leaf biomass to increase the global total assimilation. The inter-annual variability (IAV) in GPP over the whole historical period for the ProgLAI run is $2.6\,\mathrm{PgC\,yr^{-1}}$, considerably larger than in the PresLAI case ($1.7\,\mathrm{PgC\,yr^{-1}}$), but within the range of other CMIP5 models. We also notice a large decadal variability of global GPP for the ProgLAI case, which is much weaker in the PresLAI case (1.9 vs. 1.3 $\mathrm{PgC\,yr^{-1}}$). Natural variability of the climate is the main driver for the IAV in GPP for the PresLAI case. The larger variability in the ProgLAI case is due to the stronger response to volcanic cooling and climate, causing an increase in LAI and a positive feedback through increased GPP. In the PresLAI case, without the LAI feedback, the impact of volcanic cooling is sometimes largely offset by natural climate variability, for example in the Pinatubo (1991) case.

The difference between the two simulations is less obvious for the net ecosystem exchange (Fig. 5c). NEE is a relatively small flux that represents the difference between respiration (heterotrophic and autotrophic) and GPP. In the current set up of ACCESS-ESM1 we do not include disturbances such as fire and LULCC, which means that in this case NEE also represents the net flux of carbon from the land to the atmosphere. Both simulations generally produce small land sinks over most of the historical period, with some tendency to an increasing sink from the 1920s, followed by a possible reduction in the sink from the mid 1990s to 2005. The IAV is relatively large and similar for both scenarios (1.4 vs. 1.3 $\mathrm{PgC\,yr^{-1}}$) and likely caused by variations in GPP (Piao et al., 2009; Jung et al., 2011) that are moderated by respiration, especially in the ProgLAI case. Law et al. (2015, Table 2) found similar IAV in the preindustrial simulation with larger GPP IAV in the ProgLAI case offset by positively correlated leaf respiration IAV. Decadal variability for the ProgLAI run is larger than for the PresLAI run (0.7 vs. 0.3 $\mathrm{PgC\,yr^{-1}}$).

Larger decadal variability in the ProgLAI run can be explained by the stronger response to volcanic eruptions. In principle, aerosols scatter incoming solar radiation and therefore have a mainly cooling effect. Hence, an increase in aerosol emissions leads to a decrease in global temperature which in turn increases GPP in the tropics and reduces plant respiration globally in both cases (PresLAI and ProgLAI) and therefore increases NEE. However, whereas in the PresLAI case the LAI is kept at a

constant level, in the ProgLAI case the LAI is allowed to increase with the leaf carbon pools (Fig. 5b). This leads to a further increase in GPP at the same time (Fig. 5a) which further increases NEE in the ProgLAI case.

Due to the fact that during the control run ~~NEE~~ our net carbon flux did not equilibrate to zero (Law et al., 2015, Sec. 4.2.2), we calculate the carbon uptake for both scenarios by subtracting the mean ~~NEE~~ net flux over the corresponding part of the control run. ~~In this way we~~ We estimate a total uptake of carbon to the land (using the net ecosystem production (NEP), with NEP $= -1 \times$ NEE) over the historical period of ~~1289~~8 PgC for the PresLAI scenario and ~~154~~137 PgC for the ProgLAI scenario. The increase in biomass over the historical period is 70 PgC for PresLAI and 87 PgC for ProgLAI, (see also Table 2). This is similar to results from CMIP5 models that also do not consider LULCC. For, example the Beijing Climate Center Climate System Model (BCC-CSM1.1) estimates an increase in biomass of about 83 PgC over the historical period and the Institute of Numerical Mathematics Coupled Model (INM-CM4.0) reports an increase of about 70 PgC (Jones et al., 2013). The increase in combined soil and litter carbon over the historical period is smaller in ACCESS-ESM1 (28 PgC for PresLAI and 49 PgC for ProgLAI) than in the two CMIP5 models without LULCC (64 PgC for both, BCC-CSM1.1 and INM-CM4.0).

We can compare the total carbon uptake (here cumulative NEP) from ACCESS-ESM1 with other models and estimates in two ways:

1. Comparison against land-use emission estimates:

   The observation based cumulative historical land carbon uptake is estimated to be -11 ± 47 PgC ~~Arora et al. (2011)~~(Arora et al., 2011 which suggests an almost neutral behaviour of the land over that period. Since we do not include disturbances in our model, we do not expect our simulations to match those results. However, we can compare our calculated cumulative uptake against estimates of land-use emissions to see if they are in a similar range. For example, Houghton (2010) reports land-use emissions of 108–188 PgC for 1850-2000, comparable to the ACCESS-ESM1 cumulative uptakes.

2. Comparison against CMIP5 estimates of cumulative NEP:

   Simulation results from CMIP5 ESMs that include LULCC provide a large range for the total carbon uptake. Shao et al. (2013, Table 4), for example, reports the separate contributions of ~~net ecosystem production (NEP)~~ NEP and disturbance to cumulative land carbon uptake for eight CMIP5 models. While NEP ranges from 24-1730 (median 387) PgC and disturbance ranges from 3-1729 PgC, the range for land uptake is smaller with two outlying models (-120 and 211 PgC) and the remainder ranging from -59 to 18 PgC. ~~Jones et al. (2013) reports a similar range in land carbon storage across 13 CMIP5 models that include land-use change (-124 to 50 for 1850-2005). They attribute the wide range to uncertainty in the strength of fertilization effects and differences in the way land use change is implemented.~~ The estimates of cumulative NEP from ACCESS-ESM1 are at the low end of the CMIP5 range reported in Shao et al. (2013), possibly due to the inclusion of nitrogen (N) and phosphorus (P) limitation; Zhang et al. (2013) found a reduction of 1850-2005 NEP from 210 PgC for a carbon-only simulation to 85 PgC with N and P limitation when using CABLE in a low resolution earth system model.

4.2 Ocean response to historical forcing

Figure 6 shows that, consistent with other CMIP5 models, there is no statistically significant trend of ocean NPP globally over the historical period. The global mean NPP from ACCESS-ESM1 of ~~46~~51 $\mathrm{PgC\,yr^{-1}}$ is close to that calculated from the SeaWIFS data of ~~52~~50 $\mathrm{PgC\,yr^{-1}}$ for 1998-2005. Furthermore it is also in agreement with estimates, based on observations, of global NPP of between 45-50 $\mathrm{PgC\,yr^{-1}}$ (Behrenfeld and Falkowski, 1997). The ACCESS-ESM1 NPP is larger than the median CMIP5 model value of 37 PgC, however NPP in CMIP5 models is associated with a very large range (Anav et al., 2013).

The evolution of sea-air CO_2 fluxes in the period 1850-2005 is shown in Fig. 7. Overlain on this plot is the timing of the major volcanic eruptions, the estimated sea-air CO_2 flux from the Global Carbon Project (GCP) (Le Quéré et al., 2015) and results from the CMIP5 model archive. We also take into account the drift over the corresponding part of the control run. Here we see very good agreement with the CMIP5 models in the period 1870-1960, with the ACCESS-ESM1 sitting close to the median of the CMIP5 models, and well within the range of the CMIP5 models. After 1960, ACCESS-ESM1 shows greater uptake than the median of CMIP5 models, and appears to more closely follow the observed value from the GCP, lying at the 10th percentile of the CMIP5 range. For 1960-2005, ACCESS-ESM1 gives a mean sea-air CO_2 flux of $1.8 \pm 0.1\,\mathrm{PgC\,yr^{-1}}$ in good agreement with the estimated GCP value of $1.9 \pm 0.3\,\mathrm{PgC\,yr^{-1}}$, and larger than the estimate from CMIP5 models of $1.56 \pm 0.1\,\mathrm{PgC\,yr^{-1}}$. For 1986-2005, the sea-air CO_2 is $2.2 \pm 0.1\,\mathrm{PgC\,yr^{-1}}$ from ACCESS-ESM1, the same as from the GCP ($2.2 \pm 0.2\,\mathrm{PgC\,yr^{-1}}$), and larger than the median CMIP5 model value of $1.8 \pm 0.1\,\mathrm{PgC\,yr^{-1}}$. ~~This result highlights~~ The cumulative uptake of carbon by air-sea CO_2 fluxes in the period 1959-2005 from ACCESS-ESM1 is 83 PgC which is good agreement with the GCP value of 82 PgC (Le Quéré et al., 2015) over the same period. These results highlight that ACCESS-ESM1 show good skill at capturing the globally integrated ocean carbon uptake at the global scale.

5 Evaluation of the present day carbon cycle

The last 20 years of the historical simulation (1986-2005) is used to evaluate the simulated carbon cycle against observation based products. Analysis considers the land, ocean and atmosphere in turn.

5.1 Land carbon

5.1.1 GPP

Both ACCESS-ESM1 runs (PresLAI and ProgLAI) provide a mean GPP of about 130 $\mathrm{PgC\,yr^{-1}}$ for 1986-2005. The observation based estimate of Jung et al. (2011) suggests a GPP of about 119 $\mathrm{PgC\,yr^{-1}}$ for the same period. Other studies also suggest a global GPP within the same range: Beer et al. (2010) reports an estimate also based on FLUXNET data of $123 \pm 8\,\mathrm{PgC\,yr^{-1}}$ for the period 1998-2005; Ziehn et al. (2011) used plant traits to constrain parameters of the Farquhar photosynthesis model and estimated the global GPP for the same period to be 121 $\mathrm{PgC\,yr^{-1}}$ (95% confidence interval from 110 to 130 $\mathrm{PgC\,yr^{-1}}$) and the IPCC in its AR4 report states a global value of 120 PgC for 1995 (Denman et al., 2007).

~~The ACCESS-ESM1 simulation results of global GPP agree well with observation based estimates and other studies, although they are somewhat higher.~~ If compared with other CMIP5 earth system models which were divided into two groups by Anav et al. (2013), ACCESS-ESM1 lies in the middle of the lower group with the range 106 to 140 $\mathrm{PgC\,yr^{-1}}$. It was also noted by Anav et al. (2013), that the group of CMIP5 models with a GPP above 150 PgC did not include nitrogen limitation and might therefore overestimate GPP. ACCESS-ESM1 contains both nitrogen and phosphorus limitation, which ~~ensures~~ may provide a more realistic simulation of carbon uptake by the terrestrial biosphere.

A number of studies that base their estimates on observations suggest that a global GPP of about 120 $\mathrm{PgC\,yr^{-1}}$ may be somewhat too low. For example, Welp et al. (2011) provides a best guess of 150-175 $\mathrm{PgC\,yr^{-1}}$ and (Koffi et al., 2012) an estimate of $146 \pm 19\,\mathrm{PgC\,yr^{-1}}$. However, the estimate by Jung et al. (2011) is based on the largest set of observations and also provides a spatial distribution of GPP. In the following, we therefore use this product for the validation of the ACCESS-ESM1 land carbon component.

The mean annual cycle of GPP as simulated by the ACCESS-ESM1 is shown in Fig. 8 for both scenarios as Anav et al. (2013, Fig. 8). Observation based estimates by Jung et al. (2011) are also shown for comparison. At the global scale both ACCESS-ESM1 runs show a similar behaviour and they both overestimate GPP by about 2 $\mathrm{PgC\,month^{-1}}$ (peak amplitude) if compared with the observations as discussed earlier. However, when we split GPP into its contributions from three latitudinal regions we notice larger differences between the two ACCESS-ESM1 simulations. The ProgLAI simulation shows a much more productive northern region (by about 2 $\mathrm{PgC\,month^{-1}}$) and a lower GPP in the tropics ~~which compensate~~ (by about 0.2 $\mathrm{PgC\,month^{-1}}$), which compensated for at the global scale. Overall, both ACCESS-ESM1 simulations show good agreement with the observations in terms of the amplitude, with only a small bias of up to 2.2 $\mathrm{PgC\,month^{-1}}$ for the globe and the northern hemisphere. In contrast, a large number of CMIP5 models produce a strong positive bias during June-August on a global scale and for the northern hemisphere (Anav et al., 2013). Agreement with observations in terms of the phase is generally good, accept for the Tropics, where ACCESS-ESM1 fails to accurately reproduce the phase. However, as noted by Anav et al. (2013) this is common amongst CMIP5 models.

The spatial distribution of GPP is presented in Fig. 9 along with its IAV for the last 20 years of the historical period. Generally there is good agreement in the spatial pattern of GPP between ACCESS-ESM1 with prescribed LAI and the observation based estimate (95 % of all land points have errors smaller than 0.5 $\mathrm{kgC\,m^{-2}\,yr^{-1}}$). However, there are some small differences mainly in tropical regions (i.e. central Africa). The ACCESS-ESM1 ProgLAI run shows a larger GPP in the NH, mostly in the boreal regions, and a lower GPP for large parts of South-America (86 % of all land points have errors smaller than 0.5 $\mathrm{kgC\,m^{-2}\,yr^{-1}}$). Comparing the IAV of GPP for the two ACCESS-ESM1 runs reveals large differences. Whereas the PresLAI run shows little variability for most areas, the ProgLAI run shows large hotspots in South-America and Southeast Australia of up to 0.5 $\mathrm{kgC\,m^{-2}\,yr^{-1}}$ which are caused by the LAI feedback as discussed previously. The observation based estimate of GPP shows large areas of variability over the continents, but the distribution and magnitude are quite different to the ACCESS-ESM1 runs. However, as pointed out in Anav et al. (2013) one of the limitations of the GPP observational product is the magnitude of the IAV.

5.1.2 LAI

Global LAI estimates are mainly derived from satellite observations and various products are available. The prescribed LAI in ACCESS-ESM1 is based on MODIS observations (Yang et al., 2006) with no IAV. If compared with the observation based estimates of Zhu et al. (2013), which uses a combination of MODIS and AVHRR data, over the last 20 years of the historical period (mean of 1.4), we notice that our current prescribed LAI is somewhat smaller (mean of 1.3), but agrees well in terms of its seasonal cycle (Fig. 10). There is a number of reasons why remote sensing LAI products differ from each other, i.e. because different sensors and algorithms are used (Los et al., 2000).

The prognostic LAI which is calculated by CASA-CNP is significantly higher at the global scale (mean: 1.7) and also shows a different seasonality with its peak in August, whereas the observations suggest the peak is in July (Fig. 10). In CABLE the phenology phase is currently prescribed and the leaf onset might be defined as too late for deciduous vegetation which leads to a shift in the LAI peak by about one month.

The global seasonal cycle of LAI is mainly influenced by the northern extra-tropics and we notice that leaf coverage throughout the year and especially in autumn and winter is too high in the ProgLAI case. We clearly overestimate the mean LAI (observations suggest a mean of 1.3) and underestimate the seasonal variability. On a PFT level the main contributor to this is evergreen needle leaf forest which produces a large value (mean 3.8) over the whole year with only a very small seasonal cycle. In the tropics we underestimate LAI by a significant amount (mean of 1.5 in comparison to 2.3 as suggested by observations). This is mainly due to C4 grass showing an LAI which is about a factor of 5 smaller than the observations. Law et al. (2015) attributes the low simulated LAI of C4 grass to a large sensitivity to rainfall and the inability of CABLE to grow back C4 grass after a die back.

The overestimation of the LAI for evergreen needle leaf forest and the underestimation for C4 grass have a direct impact on GPP, which is also too large for evergreen needle leaf and too low for C4 grass. In CABLE, the calculation of GPP is related to APAR (absorbed photosynthetic active radiation) which is the product of FPAR (fraction of photosynthetically active radiation) and PAR (photosynthetically active radiation) with FPAR calculated from the LAI.

At the global scale, most CMIP5 earth system models also tend to overestimate LAI (Anav et al., 2013, Fig. 11), ranging from 1.5 in December-January to almost 3.5 in June-August. Anav et al. (2013) reports that only 2 models captured the main feature of the global LAI pattern, whereas the remaining 16 models overestimate the global LAI with some even exceeding a mean of 2.4. At the regional scale the ACCESS-ESM1 prognostic LAI is within the CMIP5 range for both hemispheres, but below the CMIP5 range for the Tropics.

5.1.3 NEE

We compare our NEE results against estimates of the residual land sink from the global carbon project (GCP) (Le Quéré et al., 2015) for 1959–2005 (Fig. 5c). The mean residual land sink and interannual variability for this period is estimated to be about $1.9\pm1.0\,\mathrm{PgC\,y^{-1}}$ compared to $1.4\pm1.3\,\mathrm{PgC\,y^{-1}}$ for PresLAI and $1.8\pm1.6\,\mathrm{PgC\,y^{-1}}$ for ProgLAI. In all cases the IAV is large relative to the mean uptake, but more so in the ACCESS-ESM1 simulations. The large IAV makes it difficult to be definitive

about land uptake trends over this period, though there is some suggestion of slightly increasing uptake in the GCP budget estimates but slightly decreasing uptake in the ACCESS-ESM1 simulations. This might be better assessed using an ensemble of simulations and extending the analysis closer to 2015 through use of the RCP scenario simulations. Simulations without anthropogenic aerosols would also be useful to determine whether the relatively strong cooling due to tropospheric aerosols in ACCESS-ESM1 is impacting the decadal evolution of land carbon uptake.

5.1.4 CNP pool sizes

The amount of carbon, nitrogen and phosphorus stored in the biomass and soil of terrestrial ecosystems as simulated by ACCESS-ESM1 is compared against other estimates from the literature. Here, we refer to the terrestrial biomass as the sum of living above ground (leaf and wood) and below ground (roots) material. All mean pool sizes and spatial distributions derived from ACCESS-ESM1 are calculated over the last 20 years of the historical period (1986-2005).

Carbon pool sizes simulated with ACCESS-ESM1 are in general smaller for the PresLAI scenario as shown in Table 2. The total carbon in the terrestrial biomass amounts to 670 PgC (PresLAI) and 807 PgC (ProgLAI). The IPCC (Prentice et al., 2001) reports two different estimates of 466 PgC and 654 PgC for the global plant carbon stock, depending on the data being used. This would imply that our plant carbon pools are somewhat to large, especially for the ProgLAI scenario. However, ~~other~~ we have to take into account account that we do not consider LULCC, which might be the reason why we overestimate the size of our carbon pools. Other studies such as Houghton et al. (2009) suggest a range of 800-1300 PgC for the global terrestrial biomass. The large range is a result of inconsistent definitions of forest, uncertain estimates of forest area, paucity of ground measurements and the lack of reliable mechanisms for upscaling ground measurements to larger areas (Houghton et al., 2009).

A large number of observational based estimates for global soil organic carbon (SOC) exists with most studies reporting a global estimate of about 1500 PgC (Scharlemann et al., 2014). SOC pools simulated by ACCESS-ESM1 are somewhat smaller with 1050 PgC for the PresLAI scenario and about 1200 PgC for the ProgLAI scenario. However, these numbers agree well with the best estimate of 1260 PgC derived from the HWSD (FAO, 2012) and considering the large range of 510 - 3040 PgC of global SOC simulated by CMIP5 models (Todd-Brown et al., 2013) this is an encouraging ~~results~~result.

The Harmonized World Soil Database (HWSD) also provides a spatial distribution of the SOC density which is shown in Fig.11 along with the results from ACCESS-ESM1. In general there is good agreement between the two ACCESS-ESM1 scenarios, showing a similar pattern, but with a slightly larger density in the NH boreal region for the ProgLAI run. The agreement between the HWSD and ACCESS-ESM1 is also generally good. However, the HWSD suggest localized hot spots of high SOC density in North America and Siberia which are not covered by ACCESS-ESM1. We also underestimate SOC in the tropics especially in the maritime continent region. On the other hand, both ACCESS-ESM1 scenarios suggest a high SOC density in the north Asian region, which is not apparent in the HWSD.

In addition to other environmental constraints such as water, light and temperature, carbon storage by terrestrial ecosystems may also be limited by nutrients, predominantly nitrogen and phosphorus (Wang and Houlton, 2009; Wang et al., 2010; Zhang et al., 2013). However, few estimates are available of total nitrogen and phosphorus pool sizes and their global spatial distribution is even more uncertain.

Simulated nitrogen pool sizes are shown in Table 2, and there is only a small difference between the two ACCESS-ESM1 scenarios. Our estimate for the nitrogen in the terrestrial biomass is about 6.5 PgN. Estimates based on field data reconstructions range from about 3.5 PgN (Schlesinger, 1997) to 10 PgN (Davidson, 1994) which places the ACCESS-ESM1 results right in the middle of that range. Soil organic nitrogen pools are simulated to be about 85 PgN for both ACCESS-ESM1 scenarios which is slightly low if compared with estimates based on field data (95 PgC (Post et al., 1985) to 140 PgC (Batjes, 1996)).

The terrestrial phosphorus cycle at present day is even less constrained than the nitrogen cycle and modelling and empirical estimates vary greatly. ACCESS-ESM1 results suggest a total of 0.35 PgP in the terrestrial biosphere which is ~~slightly~~ lower than the estimated range of 0.5 - 1 PgP by Smil (2000). Organic soil phosphorus pool sizes differ to some extent between the two ACCESS-ESM1 scenarios. The PresLAI model run simulates a pool size of about 10 PgP and the ProgLAI model run gives a pool size of about 12 PgP (see Table 2). Other estimates range from about 5 PgP to about 200 PgP with the upper end being assessed as unrealistic (Smil, 2000).

5.2 Ocean carbon

5.2.1 Net primary production

To assess the ~~NPP response in different latitude bands, the~~ seasonal anomaly of ocean NPP, calculated as the anomaly of vertically integrated primary productivity through the water column, the global ocean is broken down into 5 regions, following ~~Anav et al. (2013)~~(Anav et al., 2013). Figure 6 shows the NPP seasonal anomaly from ACCESS-ESM1, CMIP5 models and SeaWIFS ~~. Seasonally, at~~ over the (SeaWIFS) observational period 1998-2005. At the global ocean scale, seasonally we see that the magnitude of NPP from ACCESS-ESM1 is less than the amplitude of CMIP5 and SeaWIFS, with poor phasing. This likely reflects the biases in ACCESS-ESM1 toward lower latitudes, reflecting excess nutrient supply~~and utilization~~ , and utilization, to the upper oligotrophic ocean (Law et al., 2015) ~~.~~

associated with deeper than observed mixed layers. In the northern and southern subtropical gyres ACCESS-ESM1 (18 N-49 N and 19 S-44 S respectively) appears to overestimate the amplitude of the observed seasonal cycle when compared with SeaWIFS. ~~In the northern subtropical gyre ACCESS-ESM1 appears to lag by up to 3 months compared with SeaWIFS and CMIP5, which show good agreement. However in the southern subtropical gyre, there is poor agreement in the phasing between ESM (ACCESS-ESM1~~ Again this overestimate of NPP is associated with deeper than observed mixed layers which increase nutrient supply to the oligotrophic upper ocean. The phase of the NPP in these regions, where agreement between observations and CMIP5 ~~) and SeaWIFS, which are two monthsahead (CMIP5) and two months behind (ACCESS-ESM1) respectively.~~ is very good, is delayed by about three months. This delay may also be explained by a combination of higher (than observed) concentrations of nutrients and slower than expected biological productions associated with cool biases, particularly in the Atlantic Ocean allowing the bloom to occur later.

In the high latitude northern hemisphere~~the amplitude of ACCESS-ESM1, consistent with other ESMs (CMIP5), underestimates the amplitude~~ , the magnitude of the seasonal cycle ~~. Again the phase~~ of NPP is not well captured in ACCESS-ESM1. While CMIP5 appears also to underestimate the magnitude of the seasonal cycle~~is delayed in ACCESS-ESM1 by up to two~~

~~months relative to SeaWIFS, but is quite different from other ESMs (CMIP5) which proceed SeaWIFS by up to 2months. In~~ , ACCESS-ESM1 is lower again. In contrast, in the Southern Ocean the ~~high latitude Southern Ocean , the magnitude~~ amplitude of the seasonal cycle ~~is reproduced well~~ of NPP in ACCESS-ESM1 ~~, as opposed to CMIP5 ESMs. However the phase of ACCESS-ESM1 appears to be delayed by several months. Finally in the~~ shows good agreement with observations. However in the high latitude oceans the phase of NPP is delayed by about 2 months. This delay may be attributed to the too shallow mixed layers that exist in these regions, which means that it is only when mixed layers start to deepen that biological productivity can start to occur. As a result the remaining growing season is shorter (than observed) leading to a reduced total productivity. This may in part explain why the total NPP northern hemisphere is much less than observed.

Interestingly, in the tropical ocean we see very good agreement in the amplitude of the seasonal cycle with CMIP5 and SeaWIFS. We note however, that comparing the phase of the seasonal cycle from ESMs (ACCESS-ESM1 and CMIP5) with SeaWIFS is not very meaningful in this region, as they all simulate their own ENSO cycle with their own timing. Therefore, any comparison over a 20 year period between models has the potential to be biased by the number of El Niño or La Niña events.

5.2.2 Sea-air CO_2 fluxes

Figure 13 shows that, in the period 1986-2005, ACCESS-ESM1 is in good agreement with the spatial pattern and the magnitude of sea-air CO_2 fluxes of Wanninkhof et al. (2013), hereafter referred to as W13. In the Southern Ocean (44 S-90 S), which is an important net sink of carbon, ACCESS-ESM1 (-0.77 $\mathrm{PgC\,yr^{-1}}$) captures a larger annual mean uptake than the sea-air CO_2 flux of W13 who only estimated an uptake of -0.18 $\mathrm{PgC\,yr^{-1}}$. In the Southern subtropical gyres (44 S-18 S) ACCESS-ESM1 (-0.39 $\mathrm{PgC\,yr^{-1}}$) captures, but overestimates, the observed sea-air flux of W13 (-0.23 $\mathrm{PgC\,yr^{-1}}$). In contrast in the Northern Hemisphere ACCESS-ESM1 underestimates the uptake at -0.36 $\mathrm{PgC\,yr^{-1}}$ and -0.19 $\mathrm{PgC\,yr^{-1}}$ in the subtropical, and (sub) polar regions respectively, while W13 estimated the uptake at -0.69 $\mathrm{PgC\,yr^{-1}}$ and -0.54 $\mathrm{PgC\,yr^{-1}}$ over the same regions. The uptake in the tropical ocean is well captured, showing very good agreement between ACESS-ESM1 and W13 who estimate an uptake of -0.56 $\mathrm{PgC\,yr^{-1}}$ and -0.57 $\mathrm{PgC\,yr^{-1}}$. Spatially the interannual variability in sea-air CO_2 flux is presented in a companion paper (Law et al., 2015).

The anomaly of the seasonal cycle of the sea-air CO_2 fluxes ~~were~~ was assessed against observations of W13 and CMIP5, shown in Fig.14 ~~.~~ for the period 1986-2005. Here, we see that ACCESS-ESM1 ~~shows that globally the amplitude of the~~ has a larger global amplitude of sea-air CO_2 fluxes ~~appears larger~~ than observed (W13) and simulated, but close to the upper value of the range from CMIP5 models. ~~Furthermore it appears~~ We also see that globally the phase of sea-air CO_2 fluxes ~~are~~ is not well captured in ACCESS-ESM1, ~~explaining why it lies~~ lying outside the range of the CMIP5 models. ~~Many of these global differences are due to representation of the Southern Ocean in~~ To better understand why there are differences between ACCESS-ESM1, ~~which does not capture well the amplitude nor phase of the seasonal cycle. This inability to reproduce the observed response may well reflect the strong summer bias in warming, and the subsequent limited NPP in this region. This suggests that during the summer the solubility response likely dominates over the NPP response, leading to an out-gassing in the summer and uptake in the winter, as discussed in Lenton et al. (2013). In the high latitude Northern Hemisphere, the~~

~~seasonal cycle~~ CMIP5 and W13 we separate the response of sea-air CO_2 ~~fluxes in ACCESS-ESM1 appears also to be larger than W13, but within the range of CMIP5 models~~into the same regions as for NPP, again following Anav et al. (2013).

ACCESS-ESM1 appears to capture well the phase of sea-air CO_2 fluxes in the subtropical gyres. In the northern subtropical gyre in particular, we see that the amplitude and phase of the seasonal cycle in ACCESS-ESM1 shows very good agreement with W13, in ~~constrast~~ contrast with other ESMs (CMIP5). In the southern subtropical gyres, while the ACCESS-ESM1 appears to overestimate the amplitude relative to the observations, we see very good agreement with CMIP5 models. As anticipated the tropical ocean shows very little seasonality, nevertheless we do see good agreement with CMIP5 models. However, the comparison of ACCESS-ESM1 against observations (while shown) is not very meaningful as W13 is based on values of oceanic pCO_2 from Takahashi et al. (2009), which does not include El Niño years.

The largest differences are seen in the representation of sea-air CO_2 fluxes in the high latitude ocean. In the high latitude northern hemisphere, we see that the magnitude is larger than either CMIP5 or W13 and shows poor phasing. While the magnitude of the seasonal cycle in the Southern Ocean lies within the upper range of CMIP5 again poor phasing is seen. That the seasonal cycle is out of phase suggests that during the summer the solubility response likely dominates over the NPP response, leading to an out-gassing in the summer and uptake in the winter, as discussed in Lenton et al. (2013). Consequently, we see that the poor global phasing in global sea-air CO_2 fluxes is likely due to the solubility dominated response of the high latitudes during the summer.

5.2.3 Anthropogenic inventory

The global inventory of anthropogenic carbon from ACCESS-ESM1 is compared with the uptake from GLODAP (Sabine et al., 2004) for the year 1994 in Fig. 15. Here we see that the spatial pattern of the column inventory of anthropogenic carbon is very well reproduced, with the large storage occurring in the North Atlantic and large uptake in the Southern Ocean. The inventory for the period 1850–1994 in ACCESS-ESM1 is ~~130~~132 PgC, which is close to the estimated value from GLODAP of 118 ± 19 PgC (Sabine et al., 2004) over the same domain. This suggests that despite a somewhat limited representation of the seasonal cycle of sea-air CO_2 fluxes in key regions of anthropogenic uptake such as the Southern Ocean, that ACCESS-ESM1 is doing a very good job, spatially and temporally, of capturing and storing anthropogenic carbon. If the entire domain (including the Arctic Ocean) the is integrated the anthropogenic uptake is 143 PgC over the same period.

5.3 Atmospheric CO_2

The land and ocean carbon fluxes have been put into two atmospheric tracers as described in Law et al. (2015, Sec. 2.4). These tracers have no impact on the model simulation but allow the atmospheric CO_2 distribution to be assessed. A reasonable simulation of known features of atmospheric CO_2 can increase our confidence in the simulated carbon fluxes. For example the seasonal cycle of atmospheric CO_2 is strongly driven by the seasonality in land carbon fluxes.

Therefore, our simulated seasonality can be realistically compared to present day atmospheric CO_2 observations.

The seasonal cycle of atmospheric CO_2 is shown for four locations at different latitudes (Fig. 16, note the different vertical scale in the upper and lower panels). Seasonal cycles from the PresLAI and ProgLAI cases are calculated as the mean over the

last 20 years of the historical period (1986-2005) with the annual mean removed from each year. The seasonality is plotted for the contribution from the land carbon fluxes only and for both the land and ocean carbon fluxes combined. The model output was taken from the nearest grid point to each location with the exception of Mace Head, where the model was sampled further west to better approximate the observations which are selected for clean-air (ocean) conditions.

As observed, the amplitude of the seasonal cycle decreases from north to south. At Alert (82° N, Fig.16(a)) both model simulations overestimate the seasonal amplitude by up to 6 ppm with the growing season starting earlier than currently observed. The ocean carbon fluxes contribute little to seasonality at this latitude. At Mace Head (53° N, Fig.16(b)) the simulated seasonal cycle is comparable to that observed with only a small difference in the seasonal amplitude (smaller than 2 ppm), while at Mauna Loa (20° N,Fig.16(c)) the ProgLAI case better represents the observed seasonality than the PresLAI case.

Seasonal cycles in the southern hemisphere (e.g. South Pole) are more challenging to simulate correctly as they are made up of roughly equal contributions from local land fluxes, northern hemisphere land fluxes and ocean fluxes. Figure16(d) shows for the PresLAI case that the simulated seasonality from the land carbon fluxes is shifted in phase when the ocean carbon contribution is included but the phase shift is away from the observed seasonality. This phase shift is not apparent for the case with ProgLAI.

6 Conclusions

The evaluation of ACCESS-ESM1 over the historical period is an essential step before using the model to predict future uptake of carbon by land and oceans. Here, we performed two different scenarios for the evaluation of the land carbon cycle: running ACCESS-ESM1 with a prescribed LAI and a prognostic LAI. Running with a prognostic LAI is our preferred choice, since this includes the vegetation feedback through the coupling between LAI and the leaf carbon pool. However, results have shown that we overestimate the amplitude of the prognostic LAI annual cycle in the northern and southern hemisphere ~~,~~and underestimate it in the tropics. In future versions we need to improve the performance of the prognostic LAI, particularly for evergreen needle leaf and C4 grass.

ACCESS-ESM1 shows a strong cooling response to anthropogenic aerosols, which is offsetting the warming due to increases in greenhouse gases. The aerosol radiative forcing over the historical period is much stronger than the IPCC best estimate, but still within the uncertainty range. The impact of the cooling due to anthropogenic aerosols in ACCESS-ESM1 needs to be quantified in future work.

The land carbon uptake over the historical period is ~~by about 20 %~~ about 40 % larger for the run with prognostic LAI in comparison to the run with prescribed LAI. This is mainly due to the stronger response to volcanic eruptions which increases GPP in the tropics and reduces plant respiration globally~~and ~~, therefore increases NEE.

Globally integrated sea-air CO_2 fluxes are well captured and we reproduce very well the cumulative uptake estimate from the Global Carbon Project (Le Quéré et al., 2015) and our anthropogenic uptake agrees very well with observed GLODAP value of Sabine et al. (2004). The spatial distribution of sea-air CO_2 fluxes is also well reproduced by CMIP5 models and observations. At the same time global ocean NPP also shows good agreement with observations and lies well within the range

of CMIP5 models. However, seasonal biases do exist in sea-air CO_2 fluxes and NPP, potentially related to biases in mixed layer depth and surface temperature that are present in ACCESS-ESM1; and will need to be addressed in later versions of ACCESS-ESM1.

Simulated carbon pool sizes are generally within the range of estimates provided in the literature. Simulated soil organic carbon has been compared against the Harmonized World Soil Database, finding very good agreement in the spatial distribution and the total size. Nitrogen and phosphorus limitation were active in our simulations and pool sizes seem reasonable if compared with other estimates. However, nitrogen and phosphorus cycles are poorly constrained and only a few global estimates exist with large uncertainties.

ACCESS-ESM1 has the capability of putting land and ocean carbon fluxes into tracers, which provides a way of assessing simulated atmospheric CO_2 concentrations. The simulated seasonal cycle is close to the observed, but we overestimate the amplitude in the high northern latitude by up to 6 ppm and we also notice small phase shifts.

~~Globally integrated sea-air fluxes are well reproduced, capturing well the most recent observations of the Global Carbon Project (Le Quéré et al., 2015) and anthropogenic uptake of Sabine et al. (2004); Key et al. (2004). The spatial distribution of sea-air fluxes is also well reproduced. Seasonally the ACCESS-ESM1 appears biased toward the Southern Hemisphere with too much uptake and too little in the North, while the tropics are well captured. These differences appear to be strongly related to the dynamical response of the model. Nevertheless in most regions the results of ACCESS-ESM1 lie within the range of published CMIP5 ESMs. Globally the annual mean ocean NPP is well captured, but is somewhat biased to the low latitudes, when compared with observations, reflecting excess nutrient delivery to the lower latitude ocean in ACCESS-ESM1.~~

Overall, land and ocean carbon modules provide realistic simulations of land and ocean carbon exchange, suggesting that ACCESS-ESM1 is a valuable tool to explore the change in land and oceanic uptake in the future.

Code availability

Code availability varies for different components of ACCESS-ESM1. The UM is licensed by the UK Met Office and is not freely available. CABLE2 is available from https://trac.nci.org.au/svn/cable/. See https://trac.nci.org.au/trac/cable/wiki/CableRegistration for information on registering to use the CABLE repository. MOM4p1 and CICE are freely available under applicable registration or copyright conditions. For MOM4p1 see http://data1.gfdl.noaa.gov/~arl/pubrel/r/mom4p1/src/mom4p1/doc/mom4_manual.html. For CICE see http://oceans11.lanl.gov/trac/CICE. For access to the MOM4p1 code with WOMBAT as used for ACCESS-ESM1, please contact Hailin Yan (Hailin.Yan@csiro.au). The OASIS3-MCT 2.0 coupler code is available from http://oasis.enes.org.

Acknowledgements. This research is supported by the Australian Government Department of the Environment, the Bureau of Meteorology and CSIRO through the Australian Climate Change Science Programme. The research was undertaken on the NCI National Facility in Canberra, Australia, which is supported by the Australian Commonwealth Government. The authors wish to acknowledge use of the Ferret program for some of the analysis and graphics in this paper. Ferret is a product of NOAA's Pacific Marine Environmental Laboratory.

(Information is available at http://ferret.pmel.noaa.gov/Ferret/). We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modelling groups (listed in Table 1 of this paper) for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals.

References

Anav, A., Friedlingstein, P., Kidston, M., Bopp, L., Ciais, P., Cox, P., Jones, C., Jung, M., Myneni, R., and Zhu, Z.: Evaluating the Land and Ocean Components of the Global Carbon Cycle in the CMIP5 Earth System Models, J. Climate, 26, 6801–6843, 2013.

Arora, V. K., Scinocca, J. F., Boer, G. J., Christian, J. R., Denman, K. L., Flato, G. M., Kharin, V. V., Lee, W. G., and Merryfield, W. J.: Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases, Geophysical Research Letters, 38, n/a–n/a, doi:10.1029/2010GL046270, http://dx.doi.org/10.1029/2010GL046270, l05805, 2011.

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. Climate, 26, 5289–5314, 2013.

Batjes, N.: Total carbon and nitrogen in the soils of the world, European Journal of Soil Science, 47, 151–163, doi:10.1111/j.1365-2389.1996.tb01386.x, http://dx.doi.org/10.1111/j.1365-2389.1996.tb01386.x, 1996.

Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., Rödenbeck, C., Arain, M. A., Baldocchi, D., Bonan, G. B., Bondeau, A., Cescatti, A., Lasslop, G., Lindroth, A., Lomas, M., Luyssaert, S., Margolis, H., Oleson, K. W., Roupsard, O., Veenendaal, E., Viovy, N., Williams, C., Woodward, F. I., and Papale, D.: Terrestrial Gross Carbon Dioxide Uptake: Global Distribution and Covariation with Climate, Science, 329, 834–838, doi:10.1126/science.1184984, http://www.sciencemag.org/content/329/5993/834.abstract, 2010.

Behrenfeld, M. J. and Falkowski, P. G.: Photosynthetic rates derived from satellite-based chlorophyll concentration, Limnology and Oceanography, 42, 1–20, doi:10.4319/lo.1997.42.1.0001, http://dx.doi.org/10.4319/lo.1997.42.1.0001, 1997.

Bi, D., Dix, M., Marsland, S. J., O'Farrell, S., Rashid, H. A., Uotila, P., Hirst, A. C., Kowalczyk, E., Golebiewski, M., Sullivan, A., Yan, H., Hannah, N., Franklin, C., Sun, Z., Vohralik, P., Watterson, I., Zhou, X., Fiedler, R., Collier, M., Ma, Y., Noonan, J., Stevens, L., Uhe, P., Zhu, H., Griffies, S. M., Hill, R., Harris, C., and Puri, K.: The ACCESS coupled model: description, control climate and evaluation, Aus. Meteor. Oceanogr. J., 63, 41–64, 2013.

Booth, B. B. B., Jones, C. D., Collins, M., Totterdell, I. J., Cox, P. M., Sitch, S., Huntingford, C., Betts, R. A., Harris, G. R., and Lloyd, J.: High sensitivity of future global warming to land carbon cycle processes, Environmental Research Letters, 7, 024 002, http://stacks.iop.org/1748-9326/7/i=2/a=024002, 2012.

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., Kerminen, V.-M., Kondo, Y., Liao, H., Lohmann, U., Rasch, P., Satheesh, S. K., Sherwood, S., Stevens, B., and Zhang, X.: Clouds and Aerosols, book section 7, p. 571–658, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, doi:10.1017/CBO9781107415324.016, www.climatechange2013.org, 2013.

Davidson, E.: Climate Change and Soil Microbial Processes: Secondary Effects are Hypothesised from Better Known Interacting Primary Effects, in: Soil Responses to Climate Change, edited by Rounsevell, M. and Loveland, P., vol. 23 of *NATO ASI Series*, pp. 155–168, Springer Berlin Heidelberg, doi:10.1007/978-3-642-79218-2_10, http://dx.doi.org/10.1007/978-3-642-79218-2_10, 1994.

de Boyer Montégut, C., Madec, G., Fischer, A. S., Lazar, A., and Iudicone, D.: Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology, Journal of Geophysical Research: Oceans, 109, n/a–n/a, doi:10.1029/2004JC002378, http://dx.doi.org/10.1029/2004JC002378, c12003, 2004.

Denman, K. L., Brasseur, G., Chidthaisong, A., Ciais, P., Cox, P. M., Dickinson, R. E., Hauglustaine, D., Heinze, C., Holland, E., Jacob, D., Lohmann, U., Ramachandran, S., Dias, Wofsy, S. C., and Zhang, X.: Couplings Between Changes in the Climate System and Biogeochemistry, in: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report

of the Intergovernmental Panel on Climate Change, edited by Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M., and Miller, H. L., Cambridge University Press, Cambridge, UK and New York, NY, 2007.

Dix, M., Vohralik, P., Bi, D., Rashid, H., Marsland, S., O'Farrell, S., Uotila, P., Hirst, T., Kowalczyk, E., Sullivan, A., Yan, H., Franklin, C., Sun, Z., Watterson, I., Collier, M., Noonan, J., Rotstayn, L., Stevens, L., Uhe, P., and Puri, K.: The ACCESS coupled model: documentation of core CMIP5 simulations and initial results, Aus. Meteor. Oceanogr. J., 63, 83–99, 2013.

Downes, S. M., Farneti, R., Uotila, P., Griffies, S. M., Marsland, S. J., Bailey, D., Behrens, E., Bentsen, M., Bi, D., Biastoch, A., Böning, C., Bozec, A., Canuto, V. M., Chassignet, E., Danabasoglu, G., Danilov, S., Diansky, N., Drange, H., Fogli, P. G., Gusev, A., Howard, A., Ilicak, M., Jung, T., Kelley, M., Large, W. G., Leboissetier, A., Long, M., Lu, J., Masina, S., Mishra, A., Navarra, A., Nurser, A. G., Patara, L., Samuels, B. L., Sidorenko, D., Spence, P., Tsujino, H., Wang, Q., and Yeager, S. G.: An assessment of Southern Ocean water masses and sea ice during 1988–2007 in a suite of interannual CORE-II simulations, Ocean Modelling, 94, 67 – 94, doi:http://dx.doi.org/10.1016/j.ocemod.2015.07.022, http://www.sciencedirect.com/science/article/pii/S1463500315001420, 2015.

FAO: Harmonized World Soil Database (version 1.2), fAO/IIASA/ISRIC/ISSCAS/JRC, 2012.

Friedlingstein, P., Cox, P., Betts, R., Bopp, L., von Bloch, 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. Climate, 19, 3337–3353, doi:10.1175/JCLI3800.1, 2006.

Friend, A. D., Lucht, W., Rademacher, T. T., Keribin, R., Betts, R., Cadule, P., Ciais, P., Clark, D. B., Dankers, R., Falloon, P. D., Ito, A., Kahana, R., Kleidon, A., Lomas, M. R., Nishina, K., Ostberg, S., Pavlick, R., Peylin, P., Schaphoff, S., Vuichard, N., Warszawski, L., Wiltshire, A., and Woodward, F. I.: Carbon residence time dominates uncertainty in terrestrial vegetation responses to future climate and atmospheric CO2, Proceedings of the National Academy of Sciences, 111, 3280–3285, doi:10.1073/pnas.1222477110, http://www.pnas.org/content/111/9/3280.abstract, 2014.

Frölicher, T. L., Sarmiento, J. L., Paynter, D. J., Dunne, J. P., Krasting, J. P., and Winton, M.: Dominance of the Southern Ocean in anthropogenic carbon and heat uptake in CMIP5 models, Journal of Climate, 28, 862–886, 2015.

Gleckler, P. J., Taylor, K. E., and Doutriaux, C.: Performance metrics for climate models, Journal of Geophysical Research: Atmospheres, 113, n/a–n/a, doi:10.1029/2007JD008972, http://dx.doi.org/10.1029/2007JD008972, d06104, 2008.

GLOBALVIEW-CO_2: Cooperative Atmospheric Data Integration Project - Carbon Dioxide, www.esrl.noaa.gov/gmd/ccgg/globalview/co2/co2_version.html, nOAA ESRL, Boulder, Colorado, 2011.

Goll, D. S., Brovkin, V., Parida, B. R., Reick, C. H., Kattge, J., Reich, P. B., van Bodegom, P. M., and Niinemets, U.: Nutrient limitation reduces land carbon uptake in simulations with a model of combined carbon, nitrogen and phosphorus cycling, Biogeosciences, 9, 3547–3569, doi:10.5194/bg-9-3547-2012, http://www.biogeosciences.net/9/3547/2012/, 2012.

Harris, I., Jones, P., Osborn, T., and Lister, D.: Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset, International Journal of Climatology, 34, 623–642, doi:10.1002/joc.3711, http://dx.doi.org/10.1002/joc.3711, 2014.

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

Houghton, R. A., Hall, F., and Goetz, S. J.: Importance of biomass in the global carbon cycle, Journal of Geophysical Research: Biogeosciences, 114, n/a–n/a, doi:10.1029/2009JG000935, http://dx.doi.org/10.1029/2009JG000935, g00E03, 2009.

Huang, C. J., Qiao, F., and Dai, D.: Evaluating CMIP5 simulations of mixed layer depth during summer, Journal of Geophysical Research: Oceans, 119, 2568–2582, doi:10.1002/2013JC009535, http://dx.doi.org/10.1002/2013JC009535, 2014.

Jones, C., Robertson, E., Arora, V., Friedlingstein, P., Shevliakova, E., Bopp, L., Brovkin, V., Hajima, T., Kato, E., Kawamiya, M., Liddicoat, S., Lindsay, K., Reick, C. H., Roelandt, C., Segschneider, J., and Tjiputra, J.: Twenty-First-Century Compatible CO2 Emissions and Airborne Fraction Simulated by CMIP5 Earth System Models under Four Representative Concentration Pathways, J. Climate, 26, 4398–4413, doi:10.1175/jcli-d-12-00554.1, http://dx.doi.org/10.1175/jcli-d-12-00554.1, 2013.

Jones, C. D., Hughes, J. K., Bellouin, N., Hardiman, S. C., Jones, G. S., Knight, J., Liddicoat, S., O'Connor, F. M., Andres, R. J., Bell, C., Boo, K.-O., Bozzo, A., Butchart, N., Cadule, P., Corbin, K. D., Doutriaux-Boucher, M., Friedlingstein, P., Gornall, J., Gray, L., Halloran, P. R., Hurtt, G., Ingram, W. J., Lamarque, J.-F., Law, R. M., Meinshausen, M., Osprey, S., Palin, E. J., Parsons Chini, L., Raddatz, T., Sanderson, M. G., Sellar, A. A., Schurer, A., Valdes, P., Wood, N., Woodward, S., Yoshioka, M., and Zerroukat, M.: The HadGEM2-ES implementation of CMIP5 centennial simulations, Geoscientific Model Development, 4, 543–570, doi:10.5194/gmd-4-543-2011, http://www.geosci-model-dev.net/4/543/2011/, 2011.

Jones, P. and Harris, I.: CRU TS3.22: Climatic Research Unit (CRU) Time-Series (TS) Version 3.22 of High Resolution Gridded Data of Month-by-month Variation in Climate (Jan. 1901- Dec. 2013), doi:doi:10.5285/18BE23F8-D252-482D-8AF9-5D6A2D40990C, http://dx.doi.org/10.5285/18BE23F8-D252-482D-8AF9-5D6A2D40990C, 2014.

Jones, P. D., Lister, D. H., Osborn, T. J., Harpham, C., Salmon, M., and Morice, C. P.: Hemispheric and large-scale land-surface air temperature variations: An extensive revision and an update to 2010, Journal of Geophysical Research: Atmospheres, 117, n/a–n/a, doi:10.1029/2011JD017139, http://dx.doi.org/10.1029/2011JD017139, d05127, 2012.

Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., Arneth, A., Bernhofer, C., Bonal, D., Chen, J., Gianelle, D., Gobron, N., Kiely, G., Kutsch, W., Lasslop, G., Law, B. E., Lindroth, A., Merbold, L., Montagnani, L., Moors, E. J., Papale, D., Sottocornola, M., Vaccari, F., and Williams, C.: Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations, Journal of Geophysical Research: Biogeosciences, 116, n/a–n/a, doi:10.1029/2010JG001566, http://dx.doi.org/10.1029/2010JG001566, g00J07, 2011.

Key, R. M., Kozyr, A., Sabine, C. L., Lee, K., Wanninkhof, R., Bullister, J. L., Feely, R. A., Millero, F. J., Mordy, C., and Peng, T.-H.: A global ocean carbon climatology: Results from Global Data Analysis Project (GLODAP), Global Biogeochemical Cycles, 18, n/a–n/a, doi:10.1029/2004GB002247, http://dx.doi.org/10.1029/2004GB002247, gB4031, 2004.

Knorr, W. and Heimann, M.: Uncertainties in global terrestrial biosphere modeling: 1. A comprehensive sensitivity analysis with a new photosynthesis and energy balance scheme, Global Biogeochemical Cycles, 15, 207–225, doi:10.1029/1998GB001059, http://dx.doi.org/10.1029/1998GB001059, 2001.

Koffi, E. N., Rayner, P. J., Scholze, M., and Beer, C.: Atmospheric constraints on gross primary productivity and net ecosystem productivity: Results from a carbon-cycle data assimilation system, Global Biogeochemical Cycles, 26, n/a–n/a, doi:10.1029/2010GB003900, http://dx.doi.org/10.1029/2010GB003900, gB1024, 2012.

Kowalczyk, E. A., Wang, Y. P., Law, R. M., Davies, H. L., McGregor, J. L., and Abramowitz, G.: The CSIRO Atmosphere Biosphere Land Exchange (CABLE) model for use in climate models and as an offline model, CSIRO Marine and Atmospheric Research technical paper 13, 2006.

Law, R. M., Ziehn, T., Matear, R. J., Lenton, A., Chamberlain, M. A., Stevens, L., Wang, Y.-P., Bi, D., and Yan, H.: The carbon cycle in the Australian Climate and Earth System Simulator (ACCESS-ESM1). 1. Model description and pre-industrial simulation, Geoscientific Model Development, xx, xx–xx, 2015.

Le Quéré, C., Moriarty, R., Andrew, R. M., Peters, G. P., Ciais, P., Friedlingstein, P., Jones, S. D., Sitch, S., Tans, P., Arneth, A., Boden, T. A., Bopp, L., Bozec, Y., Canadell, J. G., Chini, L. P., Chevallier, F., Cosca, C. E., Harris, I., Hoppema, M., Houghton, R. A., House, J. I., Jain,

A. K., Johannessen, T., Kato, E., Keeling, R. F., Kitidis, V., Klein Goldewijk, K., Koven, C., Landa, C. S., Landschützer, P., Lenton, A., Lima, I. D., Marland, G., Mathis, J. T., Metzl, N., Nojiri, Y., Olsen, A., Ono, T., Peng, S., Peters, W., Pfeil, B., Poulter, B., Raupach, M. R., Regnier, P., Rödenbeck, C., Saito, S., Salisbury, J. E., Schuster, U., Schwinger, J., Séférian, R., Segschneider, J., Steinhoff, T., Stocker, B. D., Sutton, A. J., Takahashi, T., Tilbrook, B., van der Werf, G. R., Viovy, N., Wang, Y.-P., Wanninkhof, R., Wiltshire, A., and Zeng, N.: Global carbon budget 2014, Earth System Science Data, 7, 47–85, doi:10.5194/essd-7-47-2015, http://www.earth-syst-sci-data.net/7/47/2015/, 2015.

Lenton, A., Tilbrook, B., Law, R. M., Bakker, D., Doney, S. C., Gruber, N., Ishii, M., Hoppema, M., Lovenduski, N. S., Matear, R. J., McNeil, B. I., Metzl, N., Mikaloff Fletcher, S. E., Monteiro, P. M. S., Rödenbeck, C., Sweeney, C., and Takahashi, T.: Sea-air CO_2 fluxes in the Southern Ocean for the period 1990-2009, Biogeosciences, 10, 4037–4054, doi:10.5194/bg-10-4037-2013, http://www.biogeosciences.net/10/4037/2013/, 2013.

Lenton, A., McInnes, K. L., and O'Grady, J. G.: Marine projections of warming and ocean acidification in the Australasian region, Australian Meteorological and Oceanographic Journal, in press, 2015.

Lewis, S. and Karoly, D.: Assessment of forced responses of the Australian Community Climate and Earth System Simulator (ACCESS) 1.3 in CMIP5 historical detection and attribution experiments, Australian Meteorological and Oceanographic Journal, 64, 87–101, 2014.

Los, S. O., Collatz, G. J., Sellers, P. J., Malmstrsm, C. M., Pollack, N. H., DeFries, R. S., Bounoua, L., Parris, M. T., Tucker, C. J., and Dazlich, D. A.: A global 9-year biophysical land-surface data set from NOAA AVHRR data, J. Hydrometeor., 1, 183–199, 2000.

Martin, T. H. D. T. G. M., Bellouin, N., Collins, W. J., Culverwell, I. D., Halloran, P. R., Hardiman, S. C., Hinton, T. J., Jones, C. D., McDonald, R. E., McLaren, A. J., O'Connor, F. M., Roberts, M. J., Rodriguez, J. M., Woodward, S., Best, M. J., Brooks, M. E., Brown, A. R., Butchart, N., Dearden, C., Derbyshire, S. H., Dharssi, I., Doutriaux-Boucher, M., Edwards, J. M., Falloon, P. D., Gedney, N., Gray, L. J., Hewitt, H. T., Hobson, M., Huddleston, M. R., Hughes, J., Ineson, S., Ingram, W. J., James, P. M., Johns, T. C., Johnson, C. E., Jones, A., Jones, C. P., Joshi, M. M., Keen, A. B., Liddicoat, S., Lock, A. P., Maidens, A. V., Manners, J. C., Milton, S. F., Rae, J. G. L., Ridley, J. K., Sellar, A., Senior, C. A., Totterdell, I. J., Verhoef, A., Vidale, P. L., and Wiltshire, A.: The HadGEM2 family of Met Office Unified Model climate configurations, Geoscientific Model Development, 4, 723–757, doi:10.5194/gmd-4-723-2011, http://www.geosci-model-dev.net/4/723/2011/, 2011.

Martinez, E., Antoine, D., D'Ortenzio, F., and Gentili, B.: Climate-Driven Basin-Scale Decadal Oscillations of Oceanic Phytoplankton, Science, 326, 1253–1256, doi:10.1126/science.1177012, http://www.sciencemag.org/content/326/5957/1253.abstract, 2009.

Oke, P. R., Griffin, D. A., Schiller, A., Matear, R. J., Fiedler, R., Mansbridge, J., Lenton, A., Cahill, M., Chamberlain, M. A., and Ridgeway, K.: Evaluation of a near-global eddy-resolving ocean model, Geosci. Model Dev., 6, 591–615, doi:10.5194/gmd-6-591-2013, 2013.

Piao, S., Ciais, P., Friedlingstein, P., de Noblet-Ducoudré, N., Cadule, P., Viovy, N., and Wang, T.: Spatiotemporal patterns of terrestrial carbon cycle during the 20th century, Global Biogeochemical Cycles, 23, n/a–n/a, doi:10.1029/2008GB003339, http://dx.doi.org/10.1029/2008GB003339, gB4026, 2009.

Post, W., Pastor, J., Zinke, P. J., and Stangenberger, A. G.: Global patterns of soil nitrogen storage, Nature, 317, 613–616, 1985.

Prentice, I. C., Farquhar, G. D., Fasham, M. J. R., Goulden, M. L., Heimann, M., Jaramillo, V. J., Kheshgi, H. S., Quéré, C. L., Scholes, R. J., and Wallace, D.: The carbon cycle and atmospheric carbon dioxide, in: Climate Change 2001: The Scientific Basis. Contri- bution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change, edited by Houghton, J. T., Ding, Y., Griggs, D. J., Noguer, M., van der Linden, P. J., Dai, X., Maskell, K., and Johnson, C., Cambridge University Press, Cambridge, UK and New York, NY, 2001.

Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., Kent, E. C., and Kaplan, A.: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, Journal of Geophysical Research: Atmospheres, 108, n/a–n/a, doi:10.1029/2002JD002670, http://dx.doi.org/10.1029/2002JD002670, 4407, 2003.

Rodgers, K. B., Aumont, O., Mikaloff Fletcher, S. E., Plancherel, Y., Bopp, L., de Boyer Montégut, C., Iudicone, D., Keeling, R. F., Madec, G., and Wanninkhof, R.: Strong sensitivity of Southern Ocean carbon uptake and nutrient cycling to wind stirring, Biogeosciences, 11, 4077–4098, doi:10.5194/bg-11-4077-2014, http://www.biogeosciences.net/11/4077/2014/, 2014.

Rotstayn, L. D., Collier, M. A., and Luo, J.-J.: Effects of declining aerosols on projections of zonally averaged tropical precipitation, Environmental Research Letters, 10, 044 018, http://stacks.iop.org/1748-9326/10/i=4/a=044018, 2015.

Sabine, C. L., Feely, R. A., Gruber, N., Key, R. M., Lee, K., Bullister, J. L., Wanninkhof, R., Wong, C. S., Wallace, D. W. R., Tilbrook, B., Millero, F. J., Peng, T.-H., Kozyr, A., Ono, T., and Rios, A. F.: The Oceanic Sink for Anthropogenic CO2, Science, 305, 367–371, doi:10.1126/science.1097403, http://www.sciencemag.org/content/305/5682/367.abstract, 2004.

Sallée, J.-B., Shuckburgh, E., Bruneau, N., Meijers, A. J. S., Bracegirdle, T. J., and Wang, Z.: Assessment of Southern Ocean mixed-layer depths in CMIP5 models: Historical bias and forcing response, Journal of Geophysical Research: Oceans, 118, 1845–1862, doi:10.1002/jgrc.20157, http://dx.doi.org/10.1002/jgrc.20157, 2013.

Sato, M., Hansen, J., Lacis, A., and Thomason, L.: Stratospheric Aerosol Optical Thickness, http://data.giss.nasa.gov/modelforce/strataer/, 2002.

Scharlemann, J. P., Tanner, E. V., Hiederer, R., and Kapos, V.: Global soil carbon: understanding and managing the largest terrestrial carbon pool, Carbon Management, 5, 81–91, doi:10.4155/cmt.13.77, http://dx.doi.org/10.4155/cmt.13.77, 2014.

Scherrer, S. C.: Present-day interannual variability of surface climate in CMIP3 models and its relation to future warming, International Journal of Climatology, 31, 1518–1529, doi:10.1002/joc.2170, http://dx.doi.org/10.1002/joc.2170, 2011.

Schlesinger, W.: Biogeochemistry: An Analysis of Global Change, Academic Press, 1997.

Shao, P., Zeng, X., Sakaguchi, K., Monson, R. K., and Zeng, X.: Terrestrial Carbon Cycle: Climate Relations in Eight CMIP5 Earth System Models, J. Climate, 26, 8744–8764, 2013.

Smil, V.: PHOSPHORUS IN THE ENVIRONMENT: Natural Flows and Human Interferences, Annual Review of Energy and the Environment, 25, 53–88, doi:10.1146/annurev.energy.25.1.53, http://dx.doi.org/10.1146/annurev.energy.25.1.53, 2000.

Stenchikov, G., Delworth, T. L., Ramaswamy, V., Stouffer, R. J., Wittenberg, A., and Zeng, F.: Volcanic signals in oceans, Journal of Geophysical Research: Atmospheres, 114, n/a–n/a, doi:10.1029/2008JD011673, http://dx.doi.org/10.1029/2008JD011673, d16104, 2009.

Stott, P. A., Jones, G. S., Lowe, J. A., Thorne, P., Durman, C., Johns, T. C., , and Thelen, J.-C.: Transient climate simulations with the hadgem1 climate model: causes of past warming and future climate change., J. Climate, 19, 2763–2782, http://dx.doi.org/10.1175/JCLI3731.1, 2006.

Takahashi, T., Sutherland, S. C., Wanninkhof, R., Sweeney, C., Feely, R. A., Chipman, D. W., Hales, B., Friederich, G., Chavez, F., Sabine, C., Watson, A., Bakker, D. C., Schuster, U., Metzl, N., Yoshikawa-Inoue, H., Ishii, M., Midorikawa, T., Nojiri, Y., Körtzinger, A., Steinhoff, T., Hoppema, M., Olafsson, J., Arnarson, T. S., Tilbrook, B., Johannessen, T., Olsen, A., Bellerby, R., Wong, C., Delille, B., Bates, N., and de Baar, H. J.: Climatological mean and decadal change in surface ocean pCO2, and net sea–air {CO2} flux over the global oceans, Deep Sea Research Part II: Topical Studies in Oceanography, 56, 554 – 577, doi:http://dx.doi.org/10.1016/j.dsr2.2008.12.009, http://www.sciencedirect.com/science/article/pii/S0967064508004311, surface Ocean {CO2} Variability and Vulnerabilities, 2009.

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

Todd-Brown, K. E. O., Randerson, J. T., Post, W. M., Hoffman, F. M., Tarnocai, C., Schuur, E. A. G., and Allison, S. D.: Causes of variation in soil carbon simulations from CMIP5 Earth system models and comparison with observations, Biogeosciences, 10, 1717–1736, doi:10.5194/bg-10-1717-2013, http://www.biogeosciences.net/10/1717/2013/, 2013.

Wang, Y.-P. and Houlton, B. Z.: Nitrogen constraints on terrestrial carbon uptake: Implications for the global carbon-climate feedback, Geophysical Research Letters, 36, n/a–n/a, doi:10.1029/2009GL041009, http://dx.doi.org/10.1029/2009GL041009, l24403, 2009.

Wang, Y. P., Law, R. M., and Pak, B.: A global model of carbon, nitrogen and phosphorus cycles for the terrestrial biosphere, Biogeosciences, 7, 2261–2282, doi:10.5194/bg-7-2261-2010, 2010.

Wang, Y. P., Kowalczyk, E., Leuning, R., Abramowitz, G., Raupach, M. R., Pak, B., van Gorsel, E., and Luhar, A.: Diagnosing errors in a land surface model (CABLE) in the time and frequency domains, J. Geophys. Res., 116, G01 034, doi:10.1029/2010JG001385, 2011.

Wanninkhof, R., Park, G. H., Takahashi, T., Sweeney, C., Feely, R., Nojiri, Y., Gruber, N., Doney, S. C., McKinley, G. A., Lenton, A., Le Quéré, C., Heinze, C., Schwinger, J., Graven, H., and Khatiwala, S.: Global ocean carbon uptake: magnitude, variability and trends, Biogeosciences, 10, 1983–2000, doi:10.5194/bg-10-1983-2013, http://www.biogeosciences.net/10/1983/2013/, 2013.

Welp, L., Keeling, R., Meijer, H., Bollenbacher, A., Piper, S., Yoshimura, K., Francey, R., Allison, C., and Wahlen, M.: Interannual variability in the oxygen isotopes of atmospheric CO2 driven by El Nino, Nature, 477, 579–582, doi:10.1038/nature10421, 2011.

Wieder, W. R., Boehnert, J., and Bonan, G. B.: Evaluating soil biogeochemistry parameterizations in Earth system models with observations, Global Biogeochemical Cycles, 28, 211–222, doi:10.1002/2013GB004665, http://dx.doi.org/10.1002/2013GB004665, 2014.

Williams, K. D., Harris, C. M., Bodas-Salcedo, A., Camp, J., Comer, R. E., Copsey, D., Fereday, D., Graham, T., Hill, R., Hinton, T., Hyder, P., Ineson, S., Masato, G., Milton, S. F., Roberts, M. J., Rowell, D. P., Sanchez, C., Shelly, A., Sinha, B., Walters, D. N., West, A., Woollings, T., and Xavier, P. K.: The Met Office Global Coupled model 2.0 (GC2) configuration, Geoscientific Model Development, 8, 1509–1524, doi:10.5194/gmd-8-1509-2015, http://www.geosci-model-dev.net/8/1509/2015/, 2015.

Yang, W., Tan, B., Huang, D., Rautiainen, M., Shabanov, N., Wang, Y., Privette, J., Huemmrich, K., Fensholt, R., Sandholt, I., Weiss, M., Ahl, D., Gower, S., Nemani, R., Knyazikhin, Y., and Myneni, R.: MODIS leaf area index products: from validation to algorithm improvement, Geoscience and Remote Sensing, IEEE Transactions on, 44, 1885–1898, doi:10.1109/TGRS.2006.871215, 2006.

Zeng, N., Mariotti, A., and Wetzel, P.: Terrestrial mechanisms of interannual CO2 variability, Global Biogeochemical Cycles, 19, n/a–n/a, doi:10.1029/2004GB002273, http://dx.doi.org/10.1029/2004GB002273, gB1016, 2005.

Zhang, Q., Pitman, A. J., Wang, Y. P., Dai, Y. J., and Lawrence, P. J.: The impact of nitrogen and phosphorous limitation on the estimated terrestrial carbon balance and warming of land use change over the last 156 yr, Earth Syst. Dynam., 4, 333–345, doi:10.5194/esd-4-333-2013, 2013.

Zhu, Z., Bi, J., Pan, Y., Ganguly, S., Anav, A., Xu, L., Samanta, A., Piao, S., Nemani, R. R., and Myneni, R. B.: Global Data Sets of Vegetation Leaf Area Index (LAI)3g and Fraction of Photosynthetically Active Radiation (FPAR)3g Derived from Global Inventory Modeling and Mapping Studies (GIMMS) Normalized Difference Vegetation Index (NDVI3g) for the Period 1981 to 2011, Remote Sensing, 5, 927, doi:10.3390/rs5020927, http://www.mdpi.com/2072-4292/5/2/927, 2013.

Ziehn, T., Kattge, J., Knorr, W., and Scholze, M.: Improving the predictability of global CO2 assimilation rates under climate change, Geophysical Research Letters, 38, n/a–n/a, doi:10.1029/2011GL047182, http://dx.doi.org/10.1029/2011GL047182, l10404, 2011.

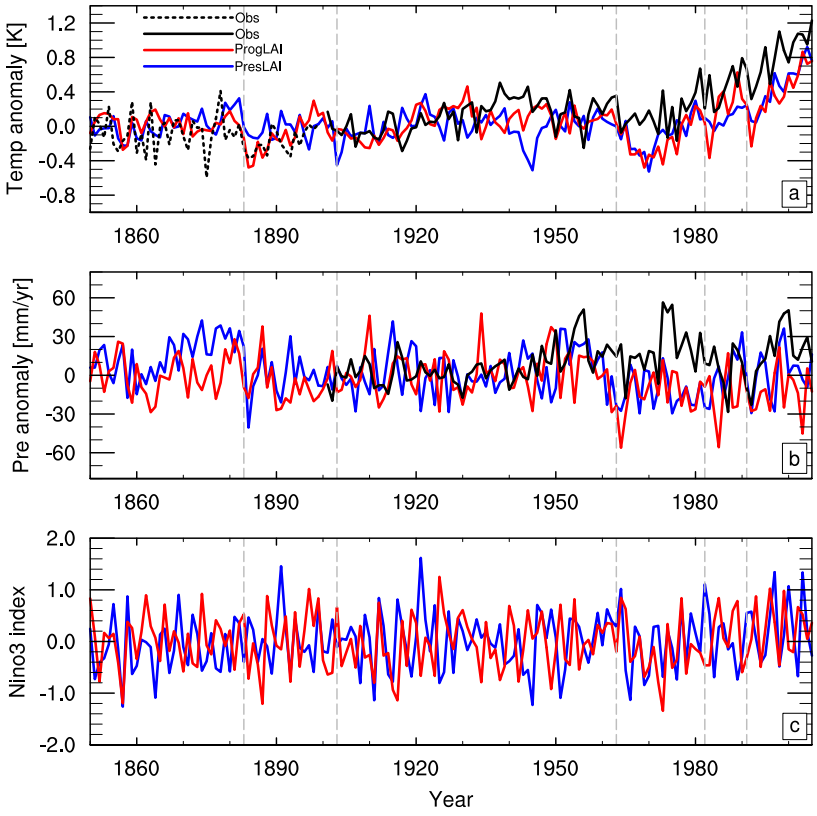


Figure 1. Anomalies (reference period: 1901-1930) for **(a)** globally averaged surface air temperature and **(b)** globally averaged precipitation for land points only for ACCESS-ESM1 (PresLAI, blue; ProgLAI, red) and observed CRU (black, dashed before 1901). Major volcanic eruptions are marked with dashed lines: Krakatoa (1983), Santa Maria (1903), Mt. Agung (1963), El Chichón (1982) and Mt. Pinatubo (1991).

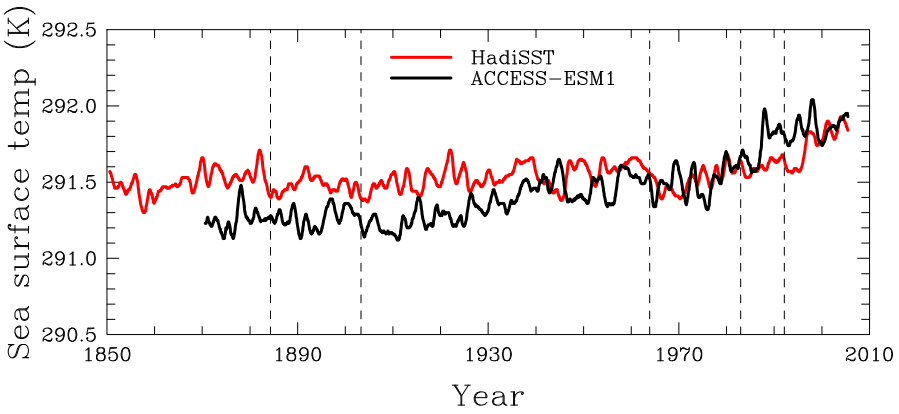


Figure 2. Globally averaged sea surface temperature (K) between 1850- 2005, red is ACCESS-ESM1 and black is HadiSST (Rayner et al., 2003). Major volcanic eruptions are marked with dashed lines: Krakatoa (1983), Santa Maria (1903), Mt. Agung (1963), El Chichón (1982) and Mt. Pinatubo (1991).

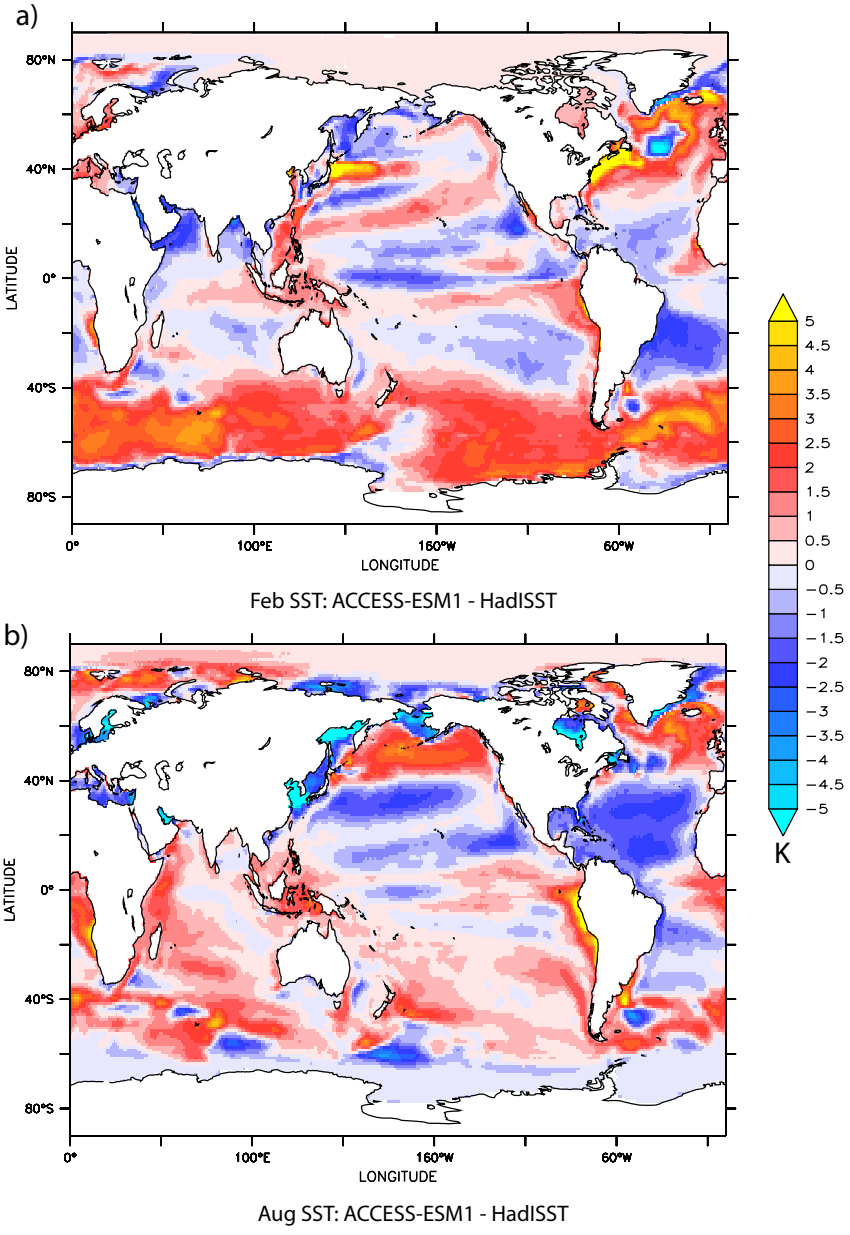


Figure 3. Differences in sea surface temperature (K) between ACCESS-ESM1 and HadiSST for **(a)** February and **(b)** August.

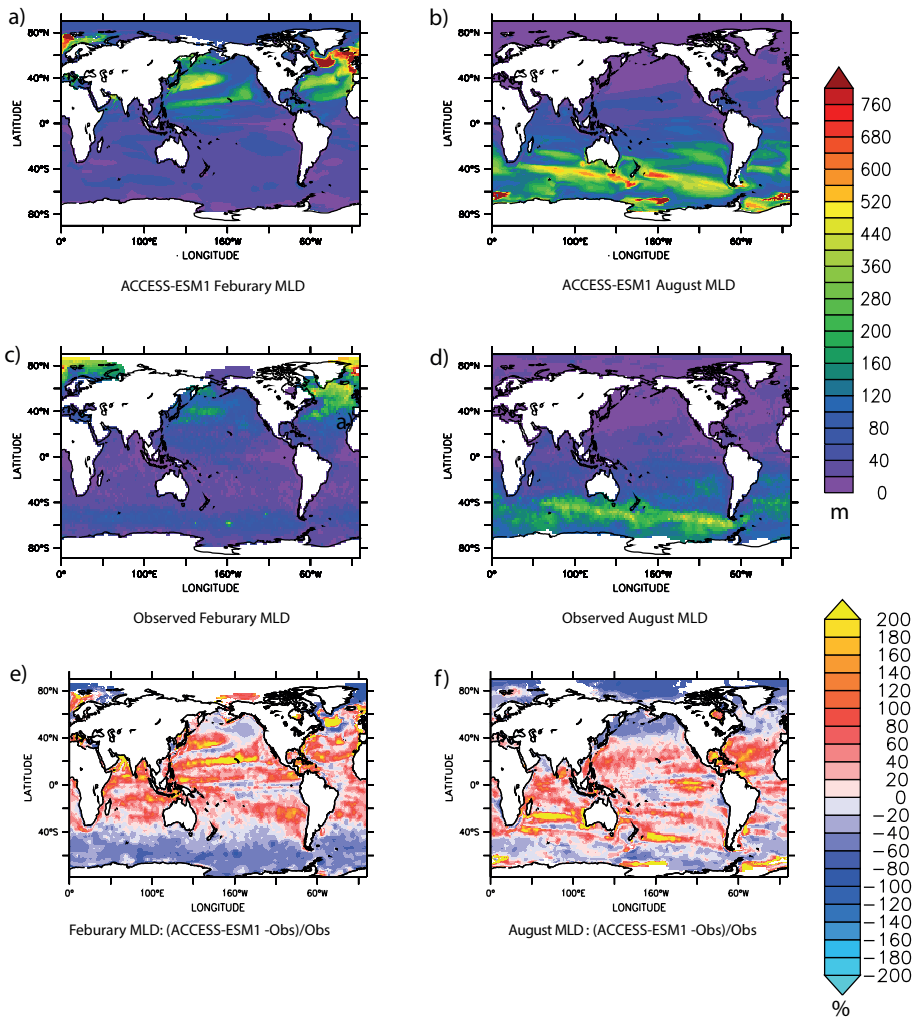


Figure 4. Differences in mixed layer depth between ACCESS-ESM1 and observations ~~from~~ de Boyer Montégut et al. (2004) for **(a,c)** February and for **(b,d)** August. Panels **(e,f)** show the percentage difference between de Boyer Montégut et al. (2004) and ACCESS-ESM1 calculated as ((OBS–ACCESS-ESM1)/OBS)*100. The mixed layer is calculated based on a 0.03 ~~kg/m^3~~ $\mathrm{kg\,m^{-3}}$ density change from the surface ocean. ~~Differences are shown for **(a,c)** February and for **(b,d)** August.~~

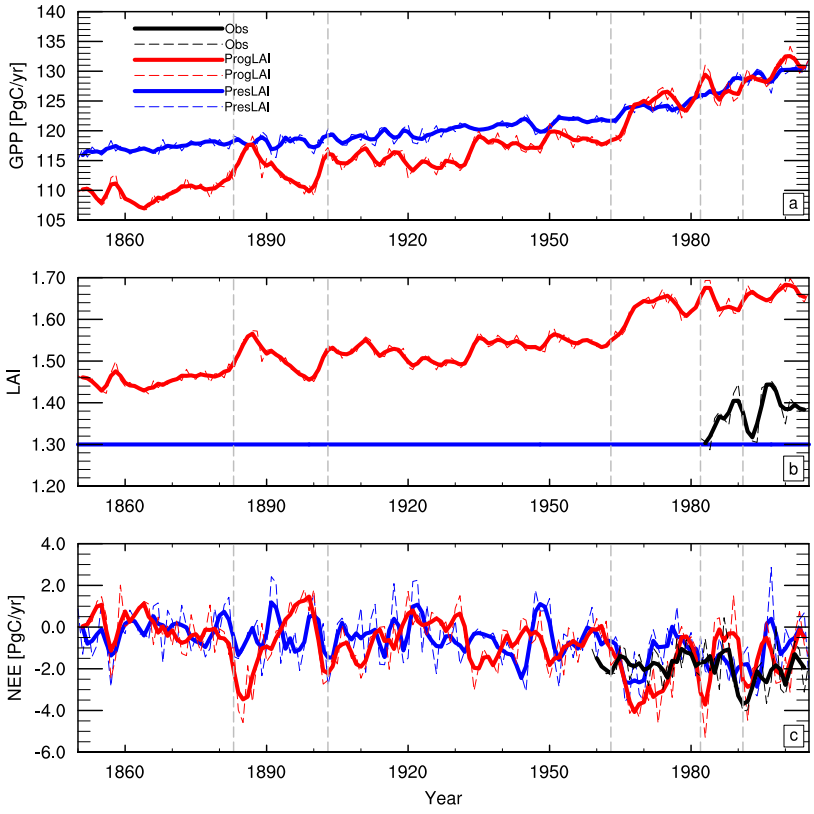


Figure 5. Temporal evolution of **(a)** GPP ($\mathrm{PgC\,yr^{-1}}$), **(b)** LAI and **(c)** NEE ($\mathrm{PgC\,yr^{-1}}$). GCP estimates for NEE are shown for comparison in black for the years 1959-2005. ACCESS-ESM1 results are shown for PresLAI (blue line) and ProgLAI (red line) with annual values marked in thin dashed lines and a 5 yr running mean in heavy solid lines. Major volcanic eruptions are marked with dashed lines: Krakatoa (1983), Santa Maria (1903), Mt. Agung (1963), El Chichón (1982) and Mt. Pinatubo (1991).

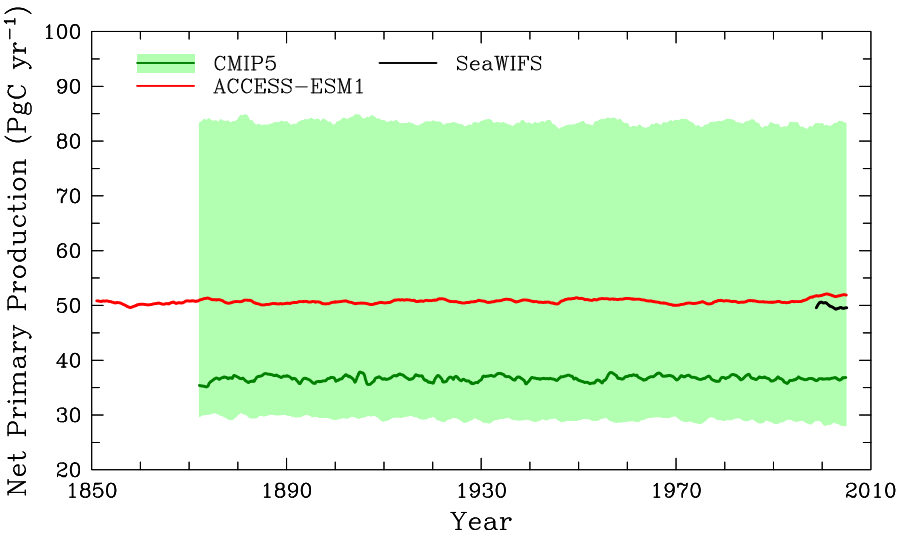


Figure 6. Comparison of Integrated Net Primary Production ($\mathrm{PgC\,yr^{-1}}$) in the period 1850-2005 between CMIP5 and ACCESS-ESM1. The solid red line represents the integrated carbon uptake in $\mathrm{PgC\,yr^{-1}}$ from ACCESS-ESM1, while the green line represents the median of the CMIP5, model with the range overlain (as shaded area) as the 10th and 90th percentiles. Overlain on this plot are the observed values from SeaWIFS over the period 1998-2005 in black.

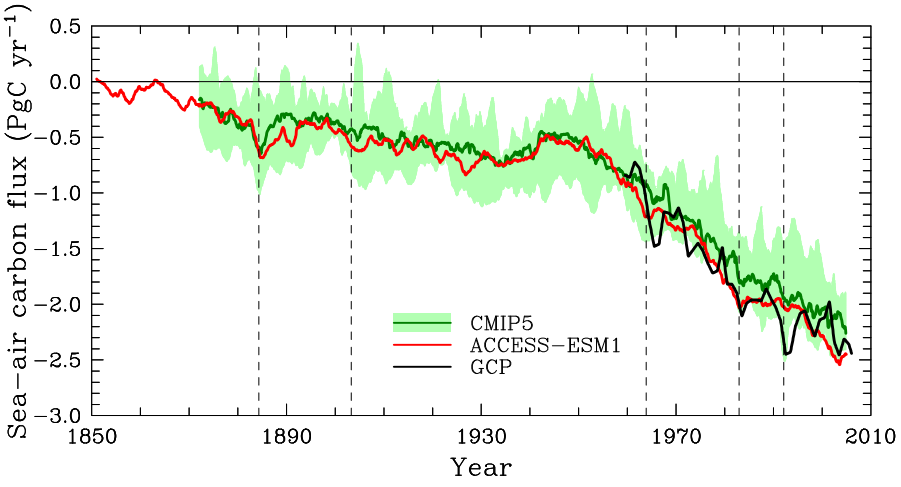


Figure 7. Comparison of sea-air CO_2 fluxes ($\mathrm{PgC\,yr^{-1}}$) in the period 1850-2005 carbon uptake from ACCESS-ESM1. The solid green line represents the median of the CMIP5, while the shaded are represents the 10th and 90th percentiles of the CMIP5 model. Overlain on this is the estimated sea-air fluxes from the Global Carbon Project (Le Quéré et al., 2015) in black; and the timing of major volcano eruptions over the historical period.

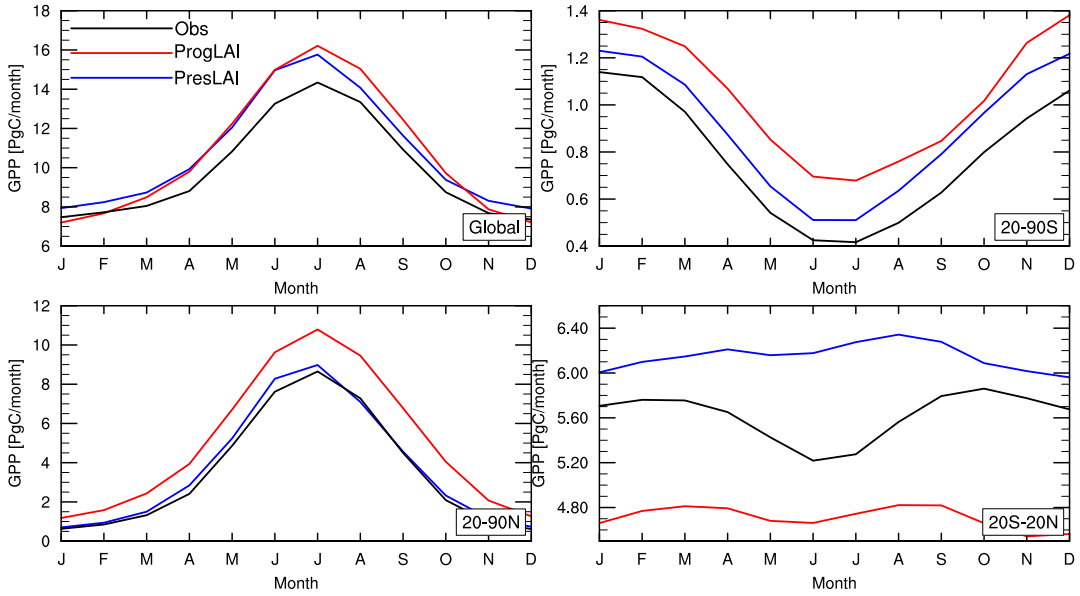


Figure 8. Mean annual cycle of GPP ($\mathrm{PgC\,month^{-1}}$) for the period 1986-2005. ACCESS-ESM1 results are shown in blue (PresLAI) and red (ProgLAI). Observation based estimates are shown in black.

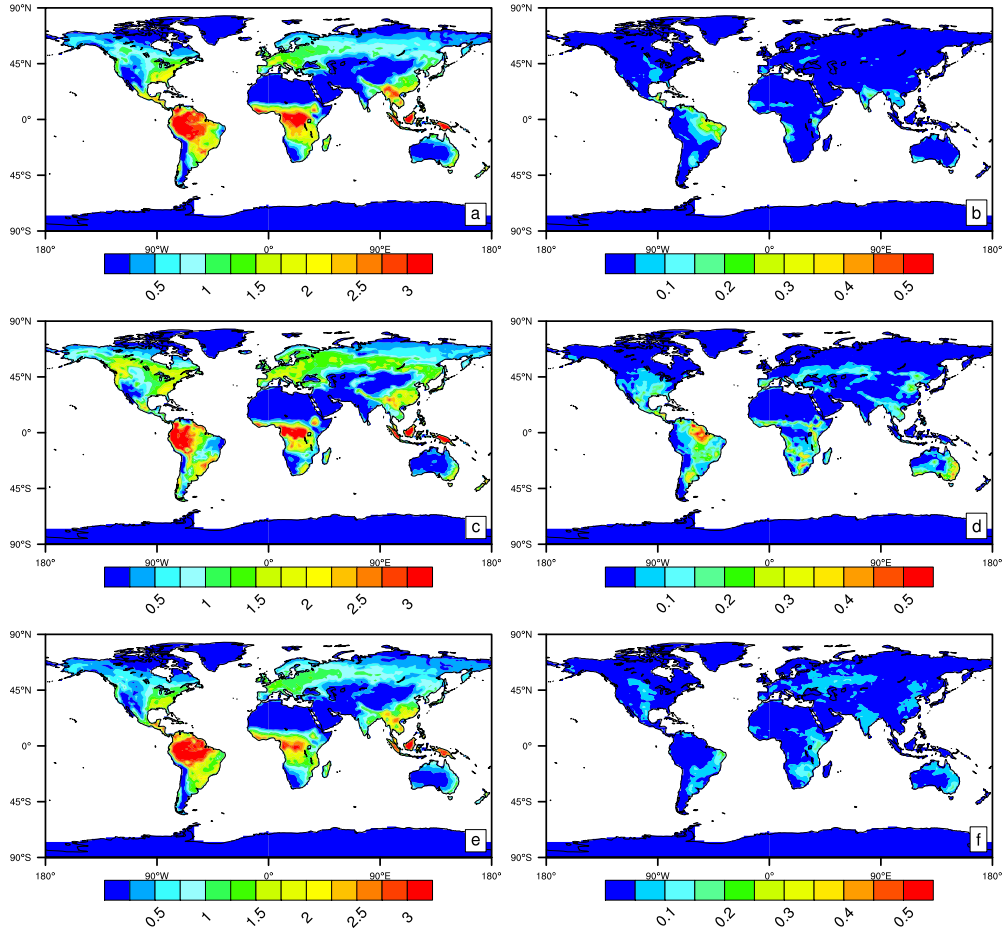


Figure 9. Spatial distribution of **(a,c,e)** GPP and **(b,d,f)** GPP IAV ($\mathrm{kgC\,m^{-2}\,yr^{-1}}$) for **(a,b)** PresLAI, **(c,d)** ProgLAI and **(e,f)** observation based estimates.

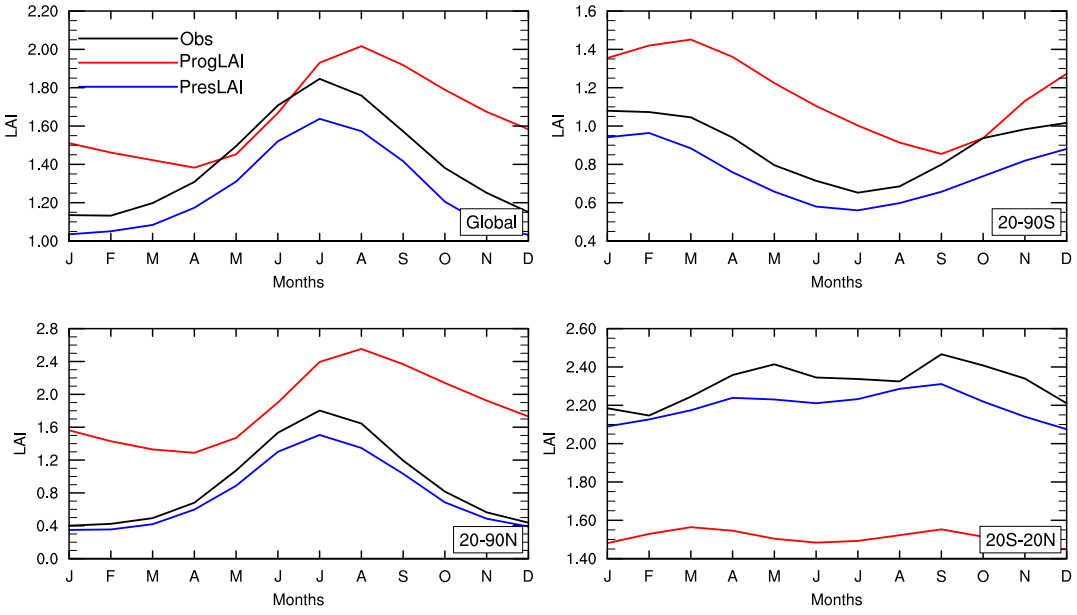


Figure 10. Mean annual cycle of LAI for the period 1986-2005. ACCESS-ESM1 results are shown in blue (scenario with prescribed LAI) and red (scenario with prognostic LAI). Observation based estimates are shown in black.

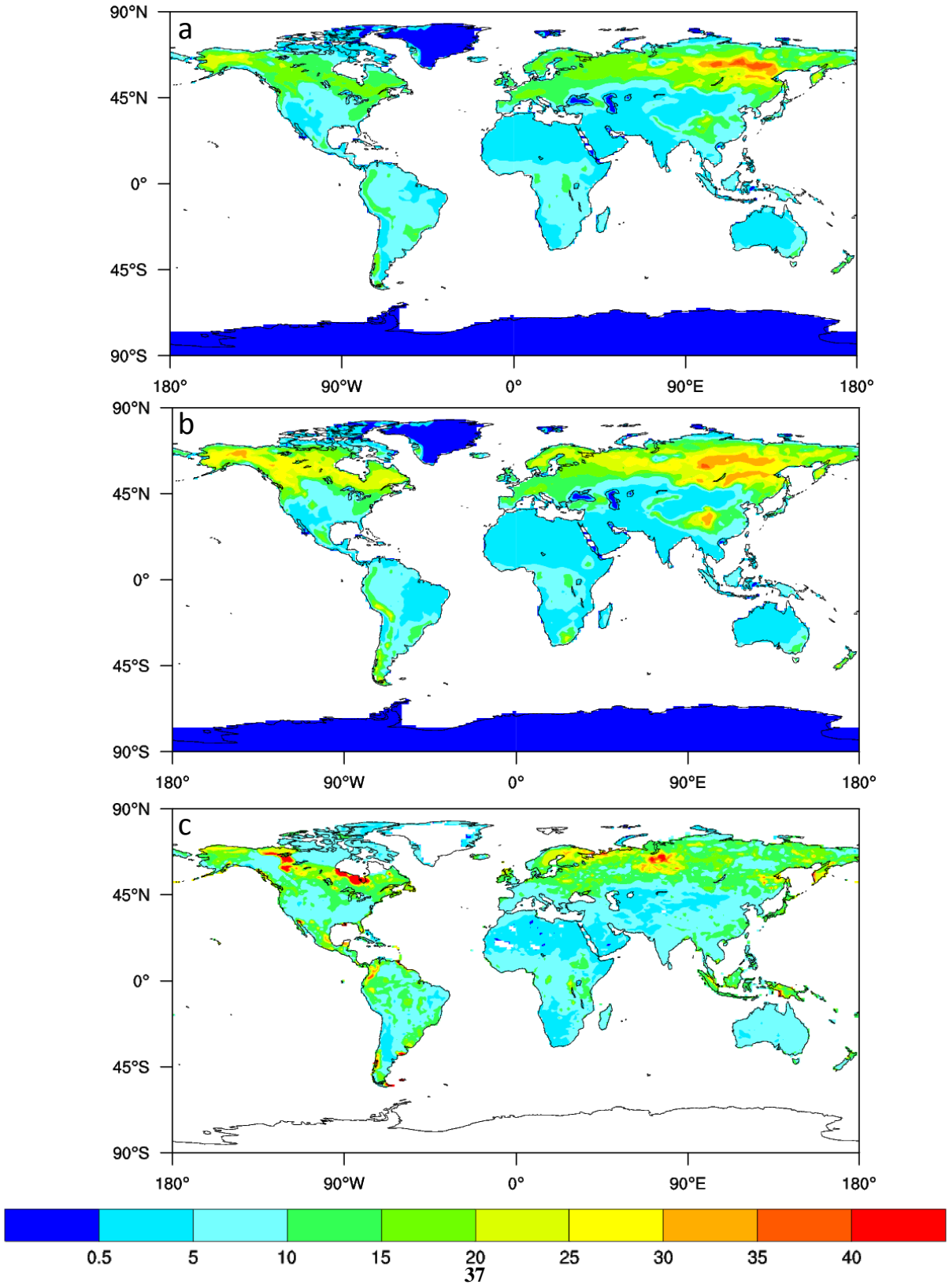
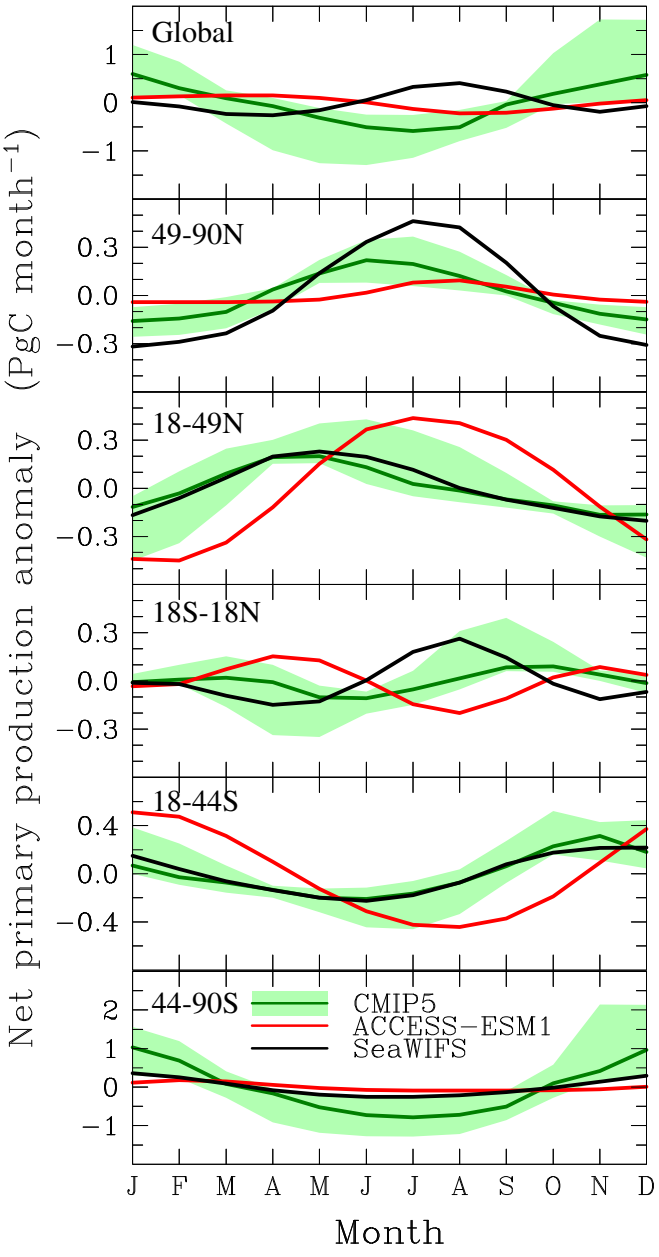


Figure 11. Spatial distribution of organic soil carbon ($\mathrm{kgC\,m^{-2}}$) (a) using prescribed LAI, (b) using prognostic LAI and (c) observation based estimated from HWSD.



~~The~~ seasonal cycle of ~~integrated net primary production~~ NPP anomalies (PgC month^{-1}) from ACCESS-ESM1 in red and SeaWIFS (Behrenfeld and Falkowski, 1997) in black calculated over the period 1998-2005. Overlain on this plot is the CMIP5 the median (solid green line) and the range 10th and 90th percentiles (shaded).

~~The~~ seasonal cycle of ~~integrated net primary production~~ NPP anomalies (PgC month^{-1}) from ACCESS-ESM1 in red and SeaWIFS (Behrenfeld and Falkowski, 1997) in black calculated over the period 1998-2005. Overlain on this plot is the CMIP5 the median (solid green line) and the range 10th and 90th percentiles (shaded).

Figure 12. The ~~integrated sea-air fluxes over the period 1986-2005 from **(a)** ACCESS-ESM1 and **(b)** Wanninkhof et al. (2013)~~ ~~The~~ seasonal cycle of ~~integrated net primary production~~ NPP anomalies (PgC month^{-1}) from ACCESS-ESM1 in red and SeaWIFS (Behrenfeld and Falkowski, 1997) in black calculated over the period 1998-2005. Overlain on this plot is the CMIP5 the median (solid green line) and the range 10th and 90th percentiles (shaded).

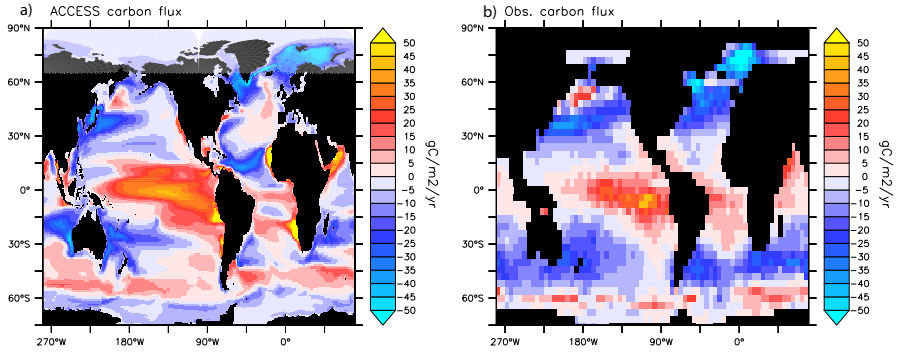


Figure 13. The integrated sea-air CO_2 fluxes over the period ~~1998-2005~~1986-2005 from **(a)** ACCESS-ESM1 and **(b)** Wanninkhof et al. (2013).

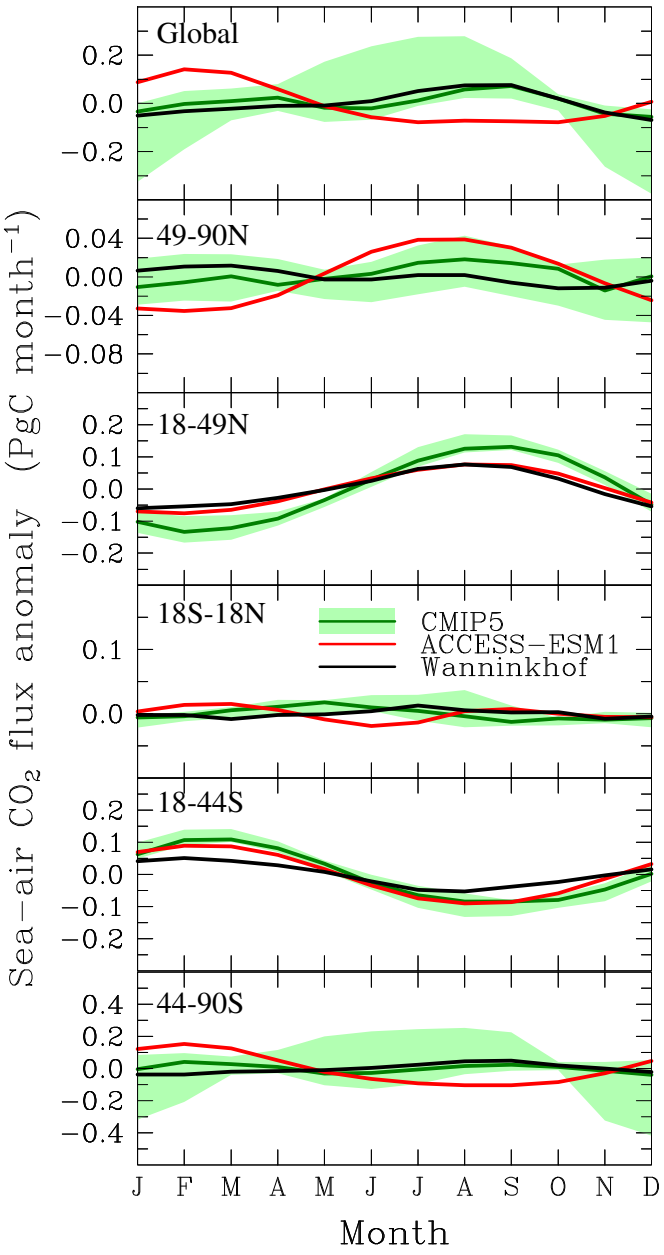


Figure 14. The seasonal cycle (1986-2005) of sea-air CO_2 flux anomalies ($\mathrm{PgC\,month^{-1}}$) from ACCESS-ESM1 (red line) and observations ((Wanninkhof et al., 2013); black line). Overlain is the CMIP5 median (solid green line) and the range as the 10th and 90th percentiles (shaded).

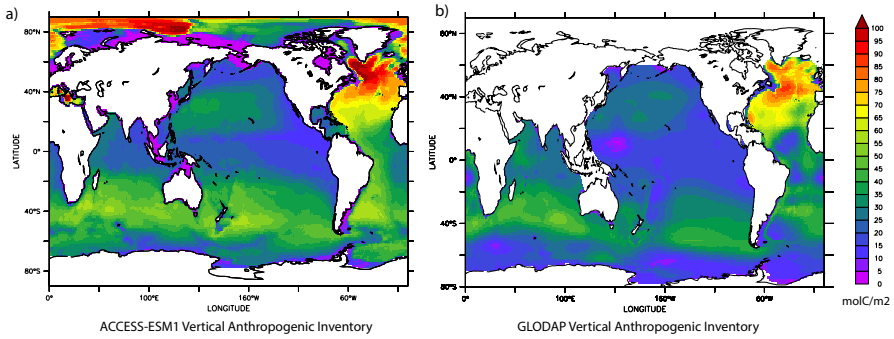


Figure 15. Column inventory of Anthropogenic Carbon in the ocean (~~molC/m2~~$\mathrm{molC\,m^{-2}}$) from **(a)** ACCESS-ESM1 and from **(b)** GLODAP (~~Sabine et al. (2004); Key et al. (2004)~~ Key et al. (2004) for 1994.

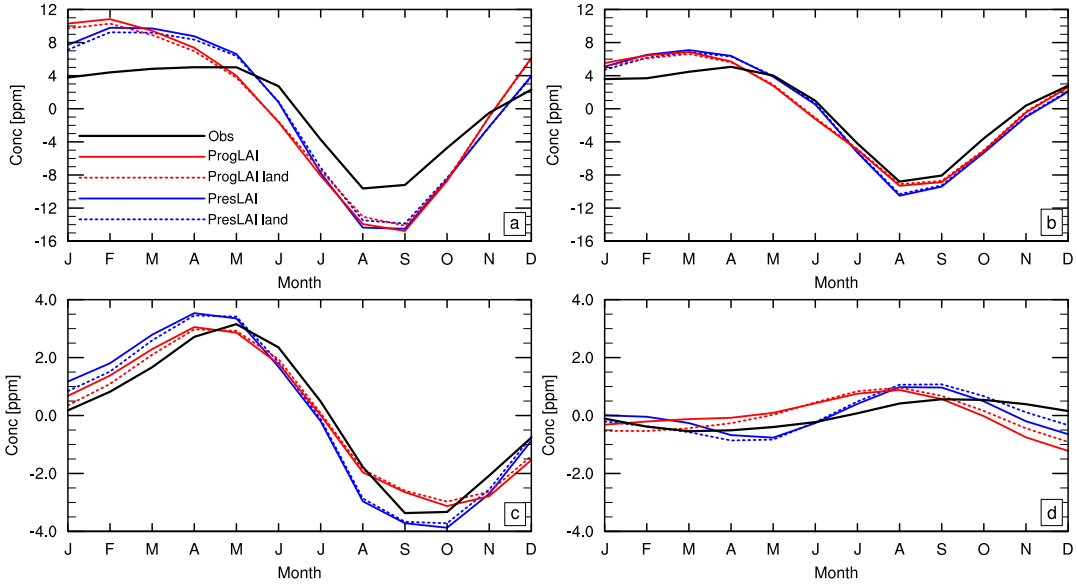


Figure 16. Mean seasonal cycle of atmospheric CO_2 for the period 1986-2005 from land carbon fluxes (dashed lines) and both land and ocean carbon fluxes (solid line). The prescribed LAI case is shown in blue, the prognostic LAI case in red and observations based on flask data from GLOBALVIEW in black for **(a)** Alert (82.45° N, 62.52° W), **(b)** Mace Head (53.33° N, 9.90° W), **(c)** Mauna Loa (19.53° N, 155.58° W) and **(d)** South Pole (89.98° S, 24.80° W).

Table 1. The CMIP5 models used to assess the ocean response of ACCESS-ESM1 over the historical period in the study. Reference for all models are provided in Lenton et al. (2015).

Model Name	Institute ID	Modelling Group
CanESM2	CCCMA	Canadian Centre for Climate Modelling and Analysis
HadGEM-ES	MOHC	Met Office Hadley Centre (additional HadGEM2-ES
	(additional realizations by INPE)	realizations contributed by Instituto Nacional de Pesquisas Espaciais)
GFDL-ESM2M	NOAA GFDL	NOAA Geophysical Fluid Dynamics Laboratory
ISPL-CM5A-LR	IPSL	Institut Pierre-Simon Laplace
IPSL-CM5A-MR	IPSL	Institut Pierre-Simon Laplace
MPI-ESM-MR	MPI-M	Max-Planck-Institut für Meteorologie
		(Max Planck Institute for Meteorology)

Table 2. Mean carbon (C), Nitrogen (N) and phosphorus (P) pools sizes in Pg for ~~the period~~ pre-industrial (780-799) and present day (1986-2005). Historical changes (1850-2005) for C are also shown. Biomass comprises leaf, wood and root pool.

	Pre-industrial						Present day						Historical change	
	PresLAI			ProgLAI			PresLAI			ProgLAI			PresLAI	ProgL
Pool	C	N	P	C	N	P	C	N	P	C	N	P	ΔC	ΔC
Biomass	611	5.7	0.31	731	6.15	0.33	670	6.2	0.34	807	~~6.84~~ 6.8	0.37	69.5	87.2
Litter	117	0.85	0.04	149	1.02	0.05	126	0.9	0.05	163	1.1	0.06	7.6	12.3
SOC	1034	82	9.6	1187	86.1	11.9	1050	83.4	~~10.14~~ 10.1	1217	~~88.53~~ 88.5	~~12.59~~ 12.6	20.5	37
$\sum$	1762	88.6	10.0	2067	93.3	12.3	1846	90.5	10.5	2187	96.4	13.0	97.6	136.5