Author Response to Review

Response 1

Response 1 Tracked Changes

Response 2

Response 2 Tracked Changes

Response to Reviewer #1

Thank you for your careful review of our manuscript, your suggestions have helped us improve our manuscript. We have shortened our manuscript by removing all of the text you identified in your detailed comments and we have removed Appendix A. Also, the detailed summary in section 4.1 has been removed and replaced with a much more succinct version that better links with the cited literature. Finally, we have added sentences to section 4.2 contextualizing our work with that of the climate-carbon feedbacks literature. Our detailed responses can be found below.

The nomencalture of anthropogenic carbon is handled quite confusingly, with Cant meaning very different things in different places in the manuscript without proper distinction. In some places, Cant(z,t) is used to describe the time-and depth-dependent anthropogenic carbon in the 1-d model, in other places, either the depth or time dependency is left away (e.g. line 142). In equation 7, then again Cant means the surface ocean concentration only. I suppose that is also, what is meant on the lefthand side of equation 8, while on the right hand side the depth- and time-dependent field is meant. In that form, the equation cannot hold in the interior of the ocean. I suggest to clarify what is meant in each instance e.g. by adding additional superscripts, e.g. CML ant for the mixed-layer Cant.

We understand that our current naming convention is potentially very confusing. We have followed your suggestions and updated the symbols to the following:

Name	**Symbol**	**Units**
Time, depth, and space dependent anthropogenic carbon concentration	$C_{ant}(x, y, z, t)$	mmol m^{-3}
Mixed layer C_{ant} concentration	$C_{ant}^{ML}(t)$	mmol m^{-3}
Atmospheric anthropogenic carbon inventory	$C_{ant}^{ATM}(t)$	Pg C

I think readers are able to understand the linear equation 4 and its consequence equation 5 without the lengthy explanation in lines 114 to 124.

We agree and have updated this section

A similar statement holds for the explanation of the impulse response function 7 from lines 179 to 195; this is a fairly standard mathematical technique and does not need to be explained in so much detail.

We have removed this lengthy explanation, which we agree is too verbose.

An important point in the methods chapter is line 220 to 222, where it is stated that the relation between CantML(t) andpCO2ocn contains the effects of changing buffer factor and of changing temperature. Here would be a good place to discuss how the decomposition of effects made here is related to the more traditional feedback analysis that results on the two feedback factors β and γ. I must admit that I wondered why it is necessary to have fitted relation between the two (Appendix B), rather than using standard carbonate chemistry and calculating one from the other, assuming constant alkalinity

We agree that this needs to be better linked to the introduction. Quantifying carbon climate feedbacks requires a different set of simulations, but their effects are included in our analysis. We have added a discussion of this to discussion section 4.2:

"While uncertainty in the mean state of ocean circulation is most important over the next 60 years, as warming increases, the magnitude of climate-carbon feedbacks related to ocean circulation increase. For simulations made with the one-dimensional model and the CESM, we simulate the effects of real-world climate-carbon feedbacks. The strength of ocean climate-carbon feedbacks (o) in CESM is weaker (-2.4 Pg C K^{-1}) than the CMIP5 multi-model mean (-7.8 Pg C K^{-1}) (Arora et al., 2013; Friedlingstein et al. 2015). Compared to Cchem, the decline in ocean carbon uptake due to climate-carbon feedbacks in high emission scenarios is an order of magnitude smaller. The one-dimensional model only simulates the reduced CO2 solubility feedback, but is close to the CESM response to warming (Figure 2d), thus indicating solubility effects dominate climate carbon feedbacks prior to 2080. The remainder of the climate-carbon feedback is related to changes in ocean circulation. In simulations out the year 2300 with CESM (Randerson et al. 2015), or simulations with models featuring a rapidly declining AMOC (Sarmiento et al., 1996), AMOC collapse plays a large role in reducing carbon uptake. The small effect of changing ocean circulation in our simulation is likely because AMOC has yet to collapse by 2080 (Randerson et al. 2015). While assuming that climate-carbon feedbacks related to ocean circulation are small prior to 2080 is consistent with the behavior of the CESM (Randerson et al. 2015), this may not be hold true for the Earth System itself. The uncertainties associated with the timing and magnitude of climate-carbon feedbacks can be avoided by mitigating climate change (Randerson et al. 2015)."

We have added the following justification for the fitted solution to Appendix B:

"There are two main reasons for using a fitted solution. First is that if every variable other than temperature and carbon are held constant, using the full carbonate system equations provides no additional accuracy, thus the additional model complexity is unjustified. Second, we want to be consistent with other models: the concentration scenarios used in CMIP5 (RCP4.5, RCP8.5) are generated using the same model and thus the same representation of ocean chemistry."

Line 314: "Assuming ocean circulation remains constant" To what extent is that assumption justified?

We justify this assumption in an updated discussion section 4.2. This updated text is in the response to your comment that starts with "An important point in the methods chapter is line 220 to 222..."

I do not completely understand, how the fields shown in Figure 4 have been calculated. Is this the zonal average of the 3-d model output minus the expected profile from the 1-dimensional model? How does advection enter this picture?

We have updated our description of Equation 6, which was understandably hard to interpret. Your interpretation is close, except that the calculation of the expected Cant is derived from the CESM zonal mean Cant in 1990. Three dimensional advection is thus accounted for.

Section 4.1 contains very little discussion and is in large parts a repetition of results. Then, in the end (lines 468-471), some other results on future Cant uptake are cited, but without giving any connection to what this study has shown. Is there a relation between these results, or is it just inserted here without too much meaning?

Section 4.1 has been revised for clarity and brevity. We have removed the discussion of the land sink.

"We find . . .": Is this really a finding, or not rather an assumption that went into the methodology?

We agree that the current wording is imprecise. Theory predicts that the vertical Cant profile would behave as a function of pCO2atm under exponential forcing but we show that this is true in a complex model. We have updated this line to be more precise:

"We find that an exponentially increasing pCO2atm allows for the vertical Cant profile to behave as a function of pCO2atm, as predicted by the historical scaling."

Appendix A again explains the impulse response function, even repeats equation 7, and gives details about using it that can be found in many textbooks. I suggest to remove it (including Figures A1 and A2), or shorten it to not more than one paragraph.

We agree and have done as you suggested, shortening it to a paragraph.

Appendix B: The coefficients shown in equation B3 to B7 all have different units, but these are not shown.

Units are not provided in the original reference for these empirically fit coefficients, and thus we do not consider it necessary to add here.

Lines 297 and 298: should one of the mentioned scenarios be RCP8.5 rather than RCP4.5?

We removed one of the sentences, it was repetitive.

References: Many references give two web addresses for the cited papers, one being the http-form of the doi (I would rather have the doi without the https://doi.org in front), and the journal address. Are really both necessary?

Thank you for noting this. We have fixed the problem.

Response to Reviewer #1

Thank you for your careful review of our manuscript, your suggestions have helped us improve our manuscript. We have shortened our manuscript by removing all of the text you identified in your detailed comments and we have removed Appendix A. Also, the detailed summary in section 4.1 has been removed and replaced with a much more succinct version that better links with the cited literature. Finally, we have added sentences to section 4.2 contextualizing our work with that of the climate-carbon feedbacks literature. Our detailed responses can be found below.

The nomencalture of anthropogenic carbon is handled quite confusingly, with Cant meaning very different things in different places in the manuscript without proper distinction. In some places, Cant(z,t) is used to describe the time-and depth-dependent anthropogenic carbon in the 1-d model, in other places, either the depth or time dependency is left away (e.g. line 142). In equation 7, then again Cant means the surface ocean concentration only. I suppose that is also, what is meant on the lefthand side of equation 8, while on the right hand side the depth- and time-dependent field is meant. In that form, the equation cannot hold in the interior of the ocean. I suggest to clarify what is meant in each instance e.g. by adding additional superscripts, e.g. CML ant for the mixed-layer Cant.

We understand that our current naming convention is potentially very confusing. We have followed your suggestions and updated the symbols to the following:

Name	**Symbol**	**Units**
Time, depth, and space dependent anthropogenic carbon concentration	$C_{ant}(x, y, z, t)$	mmol m^{-3}
Mixed layer C_{ant} concentration	$C_{ant}^{ML}(t)$	mmol m^{-3}
Atmospheric anthropogenic carbon inventory	$C_{ant}^{ATM}(t)$	Pg C

I think readers are able to understand the linear equation 4 and its consequence equation 5 without the lengthy explanation in lines 114 to 124.

We agree and have updated this section

A similar statement holds for the explanation of the impulse response function 7 from lines 179 to 195; this is a fairly standard mathematical technique and does not need to be explained in so much detail.

We have removed this lengthy explanation, which we agree is too verbose.

An important point in the methods chapter is line 220 to 222, where it is stated that the relation between CantML(t) andpCO2ocn contains the effects of changing buffer factor and of changing temperature. Here would be a good place to discuss how the decomposition of effects made here is related to the more traditional feedback analysis that results on the two feedback factors β and γ. I must admit that I wondered why it is necessary to have fitted relation between the two (Appendix B), rather than using standard carbonate chemistry and calculating one from the other, assuming constant alkalinity

We agree that this needs to be better linked to the introduction. Quantifying carbon climate feedbacks requires a different set of simulations, but their effects are included in our analysis. We have added a discussion of this to discussion section 4.2:

"While uncertainty in the mean state of ocean circulation is most important over the next 60 years, as warming increases, the magnitude of climate-carbon feedbacks related to ocean circulation increase. For simulations made with the one-dimensional model and the CESM, we simulate the effects of real-world climate-carbon feedbacks. The strength of ocean climate-carbon feedbacks (o) in CESM is weaker (-2.4 Pg C K^{-1}) than the CMIP5 multi-model mean (-7.8 Pg C K^{-1}) (Arora et al., 2013; Friedlingstein et al. 2015). Compared to Cchem, the decline in ocean carbon uptake due to climate-carbon feedbacks in high emission scenarios is an order of magnitude smaller. The one-dimensional model only simulates the reduced CO2 solubility feedback, but is close to the CESM response to warming (Figure 2d), thus indicating solubility effects dominate climate carbon feedbacks prior to 2080. The remainder of the climate-carbon feedback is related to changes in ocean circulation. In simulations out the year 2300 with CESM (Randerson et al. 2015), or simulations with models featuring a rapidly declining AMOC (Sarmiento et al., 1996), AMOC collapse plays a large role in reducing carbon uptake. The small effect of changing ocean circulation in our simulation is likely because AMOC has yet to collapse by 2080 (Randerson et al. 2015). While assuming that climate-carbon feedbacks related to ocean circulation are small prior to 2080 is consistent with the behavior of the CESM (Randerson et al. 2015), this may not be hold true for the Earth System itself. The uncertainties associated with the timing and magnitude of climate-carbon feedbacks can be avoided by mitigating climate change (Randerson et al. 2015)."

We have added the following justification for the fitted solution to Appendix B:

"There are two main reasons for using a fitted solution. First is that if every variable other than temperature and carbon are held constant, using the full carbonate system equations provides no additional accuracy, thus the additional model complexity is unjustified. Second, we want to be consistent with other models: the concentration scenarios used in CMIP5 (RCP4.5, RCP8.5) are generated using the same model and thus the same representation of ocean chemistry."

Line 314: "Assuming ocean circulation remains constant" To what extent is that assumption justified?

We justify this assumption in an updated discussion section 4.2. This updated text is in the response to your comment that starts with "An important point in the methods chapter is line 220 to 222..."

I do not completely understand, how the fields shown in Figure 4 have been calculated. Is this the zonal average of the 3-d model output minus the expected profile from the 1-dimensional model? How does advection enter this picture?

We have updated our description of Equation 6, which was understandably hard to interpret. Your interpretation is close, except that the calculation of the expected Cant is derived from the CESM zonal mean Cant in 1990. Three dimensional advection is thus accounted for.

Section 4.1 contains very little discussion and is in large parts a repetition of results. Then, in the end (lines 468-471), some other results on future Cant uptake are cited, but without giving any connection to what this study has shown. Is there a relation between these results, or is it just inserted here without too much meaning?

Section 4.1 has been revised for clarity and brevity. ~~The cited works provide the additional context of what is going on with~~ We have removed the discussion of the land sink~~during rapid mitigation, and what happens to the ocean sink beyond 2080.~~ .

"We find . . .": Is this really a finding, or not rather an assumption that went into the methodology?

We agree that the current wording is imprecise. Theory predicts that the vertical Cant profile would behave as a function of pCO2atm under exponential forcing but we show that this is true in a complex model. We have updated this line to be more precise:

"We find that an exponentially increasing pCO2atm allows for the vertical Cant profile to behave as a function of pCO2atm, as predicted by the historical scaling."

Appendix A again explains the impulse response function, even repeats equation 7, and gives details about using it that can be found in many textbooks. I suggest to remove it (including Figures A1 and A2), or shorten it to not more

than one paragraph.

We agree and have done as you suggested, shortening it to a paragraph.

Appendix B: The coefficients shown in equation B3 to B7 all have different units, but these are not shown.

~~Thank you for noticing this omission, units have been added~~Units are not provided in the original reference for these empirically fit coefficients, and thus we do not consider it necessary to add here.

Lines 297 and 298: should one of the mentioned scenarios be RCP8.5 rather than RCP4.5?

We removed one of the sentences, it was repetitive.

References: Many references give two web addresses for the cited papers, one being the http-form of the doi (I would rather have the doi without the https://doi.org in front), and the journal address. Are really both necessary?

~~We have the same preference but we are using the Biogeoscience's LaTeX template which automatically formats the references~~Thank you for noting this. We have fixed the problem.

Response to Reviewer #2

Thank you for your detailed review of our manuscript. Following your suggestions, we have carefully edited our manuscript for both clarity and conciseness. Variable names for the various forms of C_{ant} have been changed for clarity, and the introduction and methods have been rewritten. Another large change is the removal of the detailed description of the IRF model in the methods and Appendix. Please consider our detailed responses below:

1) Structure and conciseness of the manuscript (mainly introduction and methods):

***The concept of an exponential growth of emissions leading to a constant sink rate (under assumptions) is quite central in this work, and it needs to be introduced and put into context. Currently this concept is first mentioned in passing in line 44. It needs to be introduced to the reader before the sink rate is mentioned.**

Thank you, we introduce this at line 39-40 as the sink rates are defined:

“This result is as expected because the theoretical prediction of constant sink efficiency is only valid if CO2 emissions are strictly exponential."

***line 50-53: " Nearly every nation..." I don’t see that this sentence adds anything new here, consider deleting.**

Agreed, we have removed this line.

***The authors claim that "In the RCP8.5 scenario (Meinshausen et al., 2011), pCO2_atm increases exponentially..." (line 71). To me it is unclear to which degree the RCP8.5 emissions or concentraions can be approximated by an exponential, but this is not very relevant to this study either (since the baseline is an idealized exponential growth). RCP8.5 is the outcome of an advanced modelling exercise, so the emissions are not strictly exponential.**

We have decided to keep this in the manuscript. This helps the reader understand that the behavior of the sinks under RCP8.5 is similar to idealized scenarios with exponential emissions and exponential pCO_2^{atm} increase. Under the nearly exponential pCO_2^{atm} of RCP8.5, declines in k_m are dominated by changing ocean circulation and changing buffer capacity. In the other scenarios, changes to the atmospheric growth rate play a role in the decline of k_m.

***Throughout the manuscript, the authors mention exponential historical emissions. It should be made clear that this is an idealization (e.g. by saying "roughly exponential" or similar). This is already the case in some places but missing in others.**

We agree that this makes things unclear and we will update the manuscript accordingly.

***lines 54-67: Again, all this could be much more concise. The feedback studies mentioned are considering a single (exponential) concentration pathway, so they cannot quantify an uptake efficiency for different emission pathways. (And yes, in addition to this, they cannot quantify the contribution of a changing buffer factor, either).**

We understand that to someone familiar with climate-carbon feedbacks and the ocean carbon cycle, these lines may be unnecessary, however will keep these lines in the manuscript given the broader readership of Biogeosciences.

***line 74-78: Consider moving this up to lines 40-50 where k_s in general is introduced.**

We agree and have moved it up to where k_S is introduced.

***lines 102-106: This text is not necessary, we do not need a summary of subsections at the beginning of a section. Please consider removing. The same applies for lines 271-277.**

After further review, we agree and have removed the summaries.

***line 108: What is the first sentence of 2.1 supposed to tell us? Just start with "The efficiency metric (eta) used here is defined as k_m ..."**

We agree, and we now begin that line as you suggested

***The text explaining equation 4 (lines 112-118) can be shortened substantially: "The historical scaling for ocean C_ant uptake (F_ant) is defined as: (equation 4). The overset "*" indicates that a variable has been extrapolated using the historical scaling. Here, we diagnose F_ant(1990) from the CESM large ensemble simulations. For example...".**

With another look we agree, and have shortened it as suggested.

***Delete unnecessary words, e.g. "mathematically".**

After careful review we have removed many unnecessary words, including "mathematically".

***Lines 129-130: "While k_M remains constant,...". This does not reflect the logic of this manuscript. The authors use the exponential scaling to define a baseline against which simulated quantities are compared. The actual k_M is not and does not need to be constant for this exercise.**

As you have noticed this line is misleading and is clearly out of place. We have updated the manuscript to the following:

"We apply the historical scaling to C_{ant} concentration…"

***Line 144: "The CESM provides a realistic simulation of the response of the ocean carbon cycle to climate change." What do the authors mean by "realistic"? I would suggest to delete this sentence.**

We have removed the sentence following your suggestion.

***Section 2.3: Impulse response functions are a well established tool in climate modelling. It is useful to give a short explanation for those readers that are not familiar and highlight those aspects that relevant for this study, but otherwise the authors should refer the reader to the literature and shorten section 2.3 substantially. Likewise the Appendix A is not necessary. The most important assumption of IRFs is constant circulation. The most important aspect of the Joos-IRF (hidden in the Appendix A) is the fact that, contrary to atmospheric IRFs, the mixed layer IRF take ocean carbon chemistry (including changes in SST) into account.**

We have greatly shortened this section and removed the appendix as you suggested. The response to SST is included in the main text. The temperature variable in the pCO2 equation in the appendix is initial global mean SST. We have updated the text so that there is a separate variable for this temperature (T_{pi})

***Throughout section 2, it is unclear what the symbol C_ant(t) is supposed to denote. In equation 7 it is the anthropogenic carbon content of the mixed layer, but otherwise C_ant (often) seems to denote the total ocean anthropogenic carbon. The authors also frequently use the expression "C_ant air-sea flux", which I would suggest to replace by "air-sea flux of anthropogenic carbon" (and use F_ant as a short form of this if necessary).**

We understand that the current usage of variables is very confusing, and appreciate your suggestions. We have updated the ambiguous C_{ant} variables to the following:

Name	**Symbol**	**Units**
Time, depth, and space dependent anthropogenic carbon concentration	$C_{ant}(x, y, z, t)$	mmol m^{-3}
Mixed layer C_{ant} concentration	$C_{ant}^{ML}(t)$	mmol m^{-3}
Atmospheric anthropogenic carbon inventory	$C_{ant}^{ATM}(t)$	Pg C

***Section 2.3: The remaining description of the 1d-model is verbose and confusing. Apparently, the authors "extend" the mixed layer IRF downward by "plugging" a diffusion equation under the mixed layer IRF? Yes, the downward flux can be determined "by residual", but then how are the profiles of C_ant(z) calculated?**

These lines should have been removed before submission given that we do not show the downward flux in this version of the manuscript or the $C_{ant}(z)$ from the 1D model. $C_{ant}(z)$ is calculated from the CESM. The 1D diffusion representation of ocean physics for conceptual purposes.

***Section 2.5: Equation 12 can be derived by assuming F_ant = F_ant(pCO2atm(t),pCO2ocn(t)). Then, later, it is additionally assumed pCO2ocn=pCO2ocn(C_ant,T). This should be made clearer. (Again C_ant here means surface C_ant).**

We would like to make this more clear, but from this comment it is not clear what you would like clarified. Hopefully changing C_{ant} to $C_{ant}^{ML}(t)$ makes things more clear.

***Section 2.5: The equation del F_ant/del pCO2atm = del F_ant/del pCO2ocn, is this based on Equation 10? Then a minus sigh is missing. Also, the dependency of the transfer velocity on temperature is neglected in this step. How does it follow from Equation 12 that "The pCO2ocn closely follows pCO2atm, and the sign of their growth rates is the same"?**

Thank you for catching the missing minus sign. In our simplified model the transfer velocity is independent of temperature. Also, the statement "The pCO2ocn closely follows..." is not derived from the equation 12. It's a useful heuristic for understanding the equation that is derived from the behavior of the model.

***Section 3.1: If a paragraph begins with "In the RCP4.5 scenario, changes to the spatial pattern lie somewhere between RCP8.5 and the 1.5C scenario" both scenarios should have been discussed already. This is not the case here**

Thank you for catching this mistake, we no longer reference RCP8.5 before discussing it.

2) In my opinion, the term "historical scaling" used by the authors is misleading or at best confusing. The sink rate has not been constant over a substantial part of the historical period in observations (Raupach et al. 2014, cited) as well as in the model experiments used in this study (Fig S1). If the authors wish to replace the previously used term "transient steady state", why not just saying what it is, e.g. "exponential scaling" (or something similar that does not refer to the historical period)?

Thank you for your feedback, we see how this is confusing in the current version of the manuscript so we have updated the introduction. Exponential CO2 emissions growth is a necessary, but not sufficient condition to ensure that $F_{ant} = \overset{*}{F}_{ant}$. The historical scaling of F_{ant} holds if the following conditions are met: the impacts of climate change are small, ocean chemistry is relatively unchanged, and emissions continue at a exponential rate. Over the historical period, changes in these conditions are small, therefore $F_{ant} \approx \overset{*}{F}_{ant}$. In Figure 3a from Devries (2014), it is evident that observational estimates of the increase in F_{ant}, is nearly proportional to the increase in pCO_2^{atm} , therefore we can assert that variability in k_M doesn't make the long term change in F_{ant} inconsistent with the historical scaling. We refer to the scaling as the historical scaling because the necessary conditions are only satisfied over the historical period, in the RCPs these conditions are not all satisfied.

4) Choice of time periods:

The time period 1920-2006 is not the "historical period". From a CMIP5 perspective this would be 1850-2006. Why do the authors choose 1920 as a starting year? Likewise, why do the authors not use the last 20 years of the scenarios, which would be most interesting period in the mitigation scenarios?

The choice of time period was set by the length of NCAR's simulations. The historical period of the CESM ensembles begins in 1920, which differs from the CMIP5 protocol (1850 starting year). The CESM ensemble for RCP4.5 ends in 2080, while the RCP8.5 and 1.5C scenarios end in 2100. From 2080-2100 in the 1.5C scenario, the air-sea flux remains close to 0, thus 2080 is a natural cut off.

5) 1d-model evaluation:

To me it is not given that the 1d diffusion model has skill in reproducing the CESM global mean C_ant profiles for all scenarios, but this is the basis of the analysis of the "gradient effect".

We will update our manuscript to emphasize that our diagnosis of the gradient effect is not based on the 1D diffusion model. Changes to downward anthropogenic carbon transport are either due to changes in the circulation or the gradient of C_{ant}. In experiments with ocean GCMs, changes to downward anthropogenic carbon transport due to changing ocean circulation are small (Winton et al. 2013, Bronselaer and Zana 2020). Thus, regardless of whether the circulation is parameterized as diffusive as in the 1D model, or a mix of diffusive and advective processes as in the CESM, the change in vertical transport is largely due to changes in the vertical gradient of C_{ant}. In conclusion, we can diagnose the gradient effect directly from the CESM simulations.

The space saved by shortening Section 2 could be invested in presenting a brief evaluation of the (full) 1d-model compared to CESM. How well does the fitted 1d-model reproduce F_ant in the 3 different scenarios?

The 1D-model's representation of F_{ant} is shown in Figure 2d. F_{ant} as simulated by the 1D model is almost identical to the CESM simulations of F_{ant}, as a result of the tuning process.

More important, how well are vertical profiles of C_ant simulated? To me it is not given that the 1d diffusion model has skill in reproducing the CESM global mean C_ant profiles for all scenarios, but this is the basis of the analysis of the "gradient effect".

In order to clarify any potential confusion, we have updated the manuscript to make it more clear that the profiles of $C_{ant}(z)$ shown are only from the CESM. In fact, with the pulse response form of the model, we do not simulate the 1D model's vertical profile.

6) Section 3.4:

This section is not easy to follow. What is the main point here? I guess it is the fact that the ocean uptake in the strong mitigation scenario after 2040 is maintained by the ocean through continuous downward mixing (otherwise the surface ocean would start outgassing because pCO2atm declines already). Could the authors please add some easy to understand explanations here?

Your interpretation is correct. This section quantifies how the atmospheric growth rate supports the air-sea flux in the RCP4.5 and RCP8.5 scenarios, and acts to decrease the air-sea flux in the 1.5C scenario.

Further, related to my point 5) above, how realistic is this process simulated by a 1d-model?

This is the most simplistic way to view the ocean anthropogenic carbon air-sea uptake, but has been shown by many authors to be very useful in the study of the carbon cycle (Joos et al., 2013; Raupach et al., 2014; cited).

In reality we would have upwelling of waters that have been last in contact with the atmosphere in preindustrial times, that can potentially sustain ocean uptake even under declining CO2, but this is not the case in the 1d-diffusive model. Here the processes must be different. Could the authors please comment on this?

Advection and diffusion both act to mix anthropogenic carbon downwards in the ocean. Although the upward and downward advective fluxes are not necessarily colocated (e.g. Southern Ocean upwelling, subtropical downwelling), we are only interested in the integrated effect of advection on the global air-sea flux. Thus, the vertical mixing of anthropogenic carbon can be conceptualized as a 1D diffusion process.

Secondly, the HILDA model, which the mixed layer response function is derived from, includes a representation of advection and diffusion. As the impacts of climate change on ocean circulation increase, the advective and diffusive processes respond differently; however, over the next 100 years changes in uptake related to these transport processes are small (Winton et al., 2013, Bronselaer and Zana 2020).

References

Bronselaer, B., Zanna, L. Heat and carbon coupling reveals ocean warming due to circulation changes. Nature 584, 227–233 (2020). doi:10.1038/s41586-020-2573-5

DeVries, T. The oceanic anthropogenic CO2 sink: Storage, air-sea fluxes, and transports over the industrial era. Global Biogeochem. Cycles 28, 631– 647 (2014), doi:10.1002/2013GB004739.

Joos, F., Roth, R., Fuglestvedt, J. S., Peters, et al. Carbon dioxide and climate impulse response functions for the computation of greenhouse gas metrics: a multi-model analysis. Atmos. Chem. Phys. 13, 2793–2825 (2013), doi:10.5194/acp-13-2793-2013

Winton, M., S. M. Griffies, B. L. Samuels, J. L. Sarmiento, and T. L. Frölicher Connecting Changing Ocean Circulation with Changing Climate. J. Climate, 26, 2268–2278 (2020), doi:10.1175/JCLI-D-12-00296.1

Response to Reviewer #2

Thank you for your detailed review of our manuscript. Following your suggestions, we have carefully edited our manuscript for both clarity and conciseness. Variable names for the various forms of C_{ant} have been changed for clarity, and the introduction and methods have been rewritten. Another large change is the removal of the detailed description of the IRF model in the methods and Appendix. Please consider our detailed responses below:

1) Structure and conciseness of the manuscript (mainly introduction and methods):

***The concept of an exponential growth of emissions leading to a constant sink rate (under assumptions) is quite central in this work, and it needs to be introduced and put into context. Currently this concept is first mentioned in passing in line 44. It needs to be introduced to the reader before the sink rate is mentioned.**

~~Thank you for this suggestion, we will introduce this earlier in the manuscript~~Thank you, we introduce this at line 39-40 as the sink rates are defined:

"~~Exponential growth of~~ This result is as expected because the theoretical prediction of constant sink efficiency is only valid if CO2 ~~emissions leads to a declining sink rate as a result of climate change and reduced chemical capacity of the ocean. Slower than exponential CO2~~ emissions ~~results in atmospheric growth rate also driving a decline in sink rate."~~ are strictly exponential."

***line 50-53: " Nearly every nation..." I don't see that this sentence adds anything new here, consider deleting.**

Agreed, we have removed this line.

***The authors claim that "In the RCP8.5 scenario (Meinshausen et al., 2011), pCO2_atm increases exponentially..." (line 71). To me it is unclear to which degree the RCP8.5 emissions or concentraions can be approximated by an exponential, but this is not very relevant to this study either (since the baseline is an idealized exponential growth). RCP8.5 is the outcome of an advanced modelling exercise, so the emissions are not strictly exponential.**

We have decided to keep this in the manuscript. This helps the reader understand that the behavior of the sinks under RCP8.5 is similar to idealized scenarios with exponential emissions and exponential pCO_2^{atm} increase. Under the nearly exponential pCO_2^{atm} of RCP8.5, declines in k_m are dominated by changing ocean circulation and changing buffer capacity. In the other scenarios, changes to the atmospheric growth rate play a role in the decline of k_m.

***Throughout the manuscript, the authors mention exponential historical emissions. It should be made clear that this is an idealization (e.g. by saying "roughly exponential" or similar). This is already the case in some places but missing in others.**

We agree that this makes things unclear and we will update the manuscript accordingly.

***lines 54-67: Again, all this could be much more concise. The feedback studies mentioned are considering a single (exponential) concentration pathway, so they cannot quantify an uptake efficiency for different emission pathways. (And yes, in addition to this, they cannot quantify the contribution of a changing buffer factor, either).**

We understand that to someone familiar with climate-carbon feedbacks and the ocean carbon cycle, these lines may be unnecessary, however will keep these lines in the manuscript given the broader readership of Biogeosciences.

***line 74-78: Consider moving this up to lines 40-50 where k_s in general is introduced.**

We agree and have moved it up to where k_S is introduced.

***lines 102-106: This text is not necessary, we do not need a summary of subsections at the beginning of a section. Please consider removing. The same applies for lines 271-277.**

After further review, we agree and have removed the summaries.

***line 108: What is the first sentence of 2.1 supposed to tell us? Just start with "The efficiency metric (eta) used here is defined as k_m ..."**

We agree, and we now begin that line as you suggested

***The text explaining equation 4 (lines 112-118) can be shortened substantially: "The historical scaling for ocean C_ant uptake (F_ant) is defined as: (equation 4). The overset "*" indicates that a variable has been extrapolated using the historical scaling. Here, we diagnose F_ant(1990) from the CESM large ensemble simulations. For example...".**

With another look we agree, and have shortened it as suggested.

***Delete unnecessary words, e.g. "mathematically".**

After careful review we have removed many unnecessary words, including "mathematically".

***Lines 129-130: "While k_M remains constant,...". This does not reflect the logic of this manuscript. The authors use the exponential scaling to define a baseline against which simulated quantities are compared. The actual k_M is not and does not need to be constant for this exercise.**

As you have noticed this line is misleading and is clearly out of place. We have updated the manuscript to the following:

"We apply the historical scaling to C_{ant} concentration…"

***Line 144: "The CESM provides a realistic simulation of the response of the ocean carbon cycle to climate change." What do the authors mean by "realistic"? I would suggest to delete this sentence.**

We have removed the sentence following your suggestion.

***Section 2.3: Impulse response functions are a well established tool in climate modelling. It is useful to give a short explanation for those readers that are not familiar and highlight those aspects that relevant for this study, but otherwise the authors should refer the reader to the literature and shorten section 2.3 substantially. Likewise the Appendix A is not necessary. The most important assumption of IRFs is constant circulation. The most important aspect of the Joos-IRF (hidden in the Appendix A) is the fact that, contrary to atmospheric IRFs, the mixed layer IRF take ocean carbon chemistry (including changes in SST) into account.**

We have greatly shortened this section and removed the appendix as you suggested. The response to SST is included in the main text. The temperature variable in the pCO2 equation in the appendix is initial global mean SST. We have updated the text so that there is a separate variable for this temperature (T_{pi})

***Throughout section 2, it is unclear what the symbol C_ant(t) is supposed to denote. In equation 7 it is the anthropogenic carbon content of the mixed layer, but otherwise C_ant (often) seems to denote the total ocean anthropogenic carbon. The authors also frequently use the expression "C_ant air-sea flux", which I would suggest to replace by "air-**

sea flux of anthropogenic carbon" (and use F_ant as a short form of this if necessary).

We understand that the current usage of variables is very confusing, and appreciate your suggestions. We have updated the ambiguous C_{ant} variables to the following:

Name	**Symbol**	**Units**
Time, depth, and space dependent anthropogenic carbon concentration	$C_{ant}(x,y,z,t)$	mmol m^{-3}
Mixed layer C_{ant} concentration	$C_{ant}^{ML}(t)$	mmol m^{-3}
Atmospheric anthropogenic carbon inventory	$C_{ant}^{ATM}(t)$	Pg C

***Section 2.3: The remaining description of the 1d-model is verbose and confusing. Apparently, the authors "extend" the mixed layer IRF downward by "plugging" a diffusion equation under the mixed layer IRF? Yes, the downward flux can be determined "by residual", but then how are the profiles of C_ant(z) calculated?**

These lines should have been removed before submission given that we do not show the downward flux in this version of the manuscript or the $C_{ant}(z)$ from the 1D model. $C_{ant}(z)$ is calculated from the CESM. The 1D diffusion representation of ocean physics for conceptual purposes.

***Section 2.5: Equation 12 can be derived by assuming F_ant = F_ant(pCO2atm(t),pCO2ocn(t)). Then, later, it is additionally assumed pCO2ocn=pCO2ocn(C_ant,T). This should be made clearer. (Again C_ant here means surface C_ant).**

We would like to make this more clear, but from this comment it is not clear what you would like clarified. Hopefully changing C_{ant} to $C_{ant}^{ML}(t)$ makes things more clear.

***Section 2.5: The equation del F_ant/del pCO2atm = del F_ant/del pCO2ocn, is this based on Equation 10? Then a minus sigh is missing. Also, the dependency of the transfer velocity on temperature is neglected in this step. How does it follow from Equation 12 that "The pCO2ocn closely follows pCO2atm, and the sign of their growth rates is the same"?**

Thank you for catching the missing minus sign. In our simplified model the transfer velocity is independent of temperature. Also, the statement "The pCO2ocn closely follows..." is not derived from the equation 12. It's a useful heuristic for understanding the equation that is derived from the behavior of the model.

***Section 3.1: If a paragraph begins with "In the RCP4.5 scenario, changes to the spatial pattern lie somewhere between RCP8.5 and the 1.5C scenario" both scenarios should have been discussed already. This is not the case here**

Thank you for catching this mistake, we no longer reference RCP8.5 before discussing it.

2) In my opinion, the term "historical scaling" used by the authors is misleading or at best confusing. The sink rate has not been constant over a substantial part of the historical period in observations (Raupach et al. 2014, cited) as well as in the model experiments used in this study (Fig S1). If the authors wish to replace the previously used term "transient steady state", why not just saying what it is, e.g. "exponential scaling" (or something similar that does not refer to the historical period)?

Thank you for your feedback, we see how this is confusing in the current version of the manuscript so we have updated the introduction. Exponential CO2 emissions growth is a necessary, but not sufficient condition to ensure that $F_{ant} = \overset{*}{F}_{ant}$. The historical scaling of F_{ant} holds if the following conditions are met: the impacts of climate change are small, ocean chemistry is relatively unchanged, and emissions continue at a exponential rate. Over the historical period, changes in these conditions are small, therefore $F_{ant} \approx \overset{*}{F}_{ant}$. In Figure 3a from Devries (2014), it is evident that observational estimates of the increase in F_{ant}, is nearly proportional to the increase in pCO_2^{atm} , therefore we can assert that variability in k_M doesn't make the long term change in F_{ant} inconsistent with the historical scaling. We refer to the scaling as the historical scaling because the necessary conditions are only satisfied over the historical period, in the RCPs these conditions are not all satisfied.

4) Choice of time periods:

The time period 1920-2006 is not the "historical period". From a CMIP5 perspective this would be 1850-2006. Why do the authors choose 1920 as a starting year? Likewise, why do the authors not use the last 20 years of the scenarios, which would be most interesting period in the mitigation scenarios?

The choice of time period was set by the length of NCAR's simulations. The historical period of the CESM ensembles begins in 1920, which differs from the CMIP5 protocol (1850 starting year). The CESM ensemble for RCP4.5 ends in 2080, while the RCP8.5 and 1.5C scenarios end in 2100. From 2080-2100 in the 1.5C scenario, the air-sea flux remains close to 0, thus 2080 is a natural cut off.

5) 1d-model evaluation:

To me it is not given that the 1d diffusion model has skill in reproducing the CESM global mean C_ant profiles for all scenarios, but this is the basis of the analysis of the "gradient effect".

We will update our manuscript to emphasize that our diagnosis of the gradient effect is not based on the 1D diffusion model. Changes to downward anthropogenic carbon transport are either due to changes in the circulation or the gradient of C_{ant}. In experiments with ocean GCMs, changes to downward anthropogenic carbon transport due to changing ocean circulation are small (Winton et al. 2013, Bronselaer and Zana 2020). Thus, regardless of whether the circulation is parameterized as diffusive as in the 1D model, or a mix of diffusive and advective processes as in the CESM, the change in vertical transport is largely due to changes in the vertical gradient of C_{ant}. In conclusion, we can diagnose the gradient effect directly from the CESM simulations.

The space saved by shortening Section 2 could be invested in presenting a brief evaluation of the (full) 1d-model compared to CESM. How well does the fitted 1d-model reproduce F_ant in the 3 different scenarios?

The 1D-model's representation of F_{ant} is shown in Figure 2d. F_{ant} as simulated by the 1D model is almost identical to the CESM simulations of F_{ant}, as a result of the tuning process.

More important, how well are vertical profiles of C_ant simulated? To me it is not given that the 1d diffusion model has skill in reproducing the CESM global mean C_ant profiles for all scenarios, but this is the basis of the analysis of the "gradient effect".

In order to clarify any potential confusion, we have updated the manuscript to make it more clear that the profiles of $C_{ant}(z)$ shown are only from the CESM. In fact, with the pulse response form of the model, we do not simulate the 1D model's vertical profile.

6) Section 3.4:

This section is not easy to follow. What is the main point here? I guess it is the fact that the ocean uptake in the strong mitigation scenario after 2040 is maintained by the ocean through continuous downward mixing (otherwise the

surface ocean would start outgassing because pCO2atm declines already). Could the authors please add some easy to understand explanations here?

Your interpretation is correct. This section quantifies how the atmospheric growth rate supports the air-sea flux in the RCP4.5 and RCP8.5 scenarios, and acts to decrease the air-sea flux in the 1.5C scenario.

Further, related to my point 5) above, how realistic is this process simulated by a 1d-model?

This is the most simplistic way to view the ocean anthropogenic carbon air-sea uptake, but has been shown by many authors to be very useful in the study of the carbon cycle (Joos et al., 2013; Raupach et al., 2014; cited).

In reality we would have upwelling of waters that have been last in contact with the atmosphere in preindustrial times, that can potentially sustain ocean uptake even under declining CO2, but this is not the case in the 1d-diffusive model. Here the processes must be different. Could the authors please comment on this?

Advection and diffusion both act to mix anthropogenic carbon downwards in the ocean. Although the upward and downward advective fluxes are not necessarily colocated (e.g. Southern Ocean upwelling, subtropical downwelling), we are only interested in the integrated effect of advection on the global air-sea flux. Thus, the vertical mixing of anthropogenic carbon can be conceptualized as a 1D diffusion process.

Secondly, the HILDA model, which the mixed layer response function is derived from, includes a representation of advection and diffusion. As the impacts of climate change on ocean circulation increase, the advective and diffusive processes respond differently; however, over the next 100 years changes in uptake related to these transport processes are small (Winton et al., 2013, Bronselaer and Zana 2020).

References

Bronselaer, B., Zanna, L. Heat and carbon coupling reveals ocean warming due to circulation changes. Nature 584, 227–233 (2020). doi:10.1038/s41586-020-2573-5

DeVries, T. The oceanic anthropogenic CO2 sink: Storage, air-sea fluxes, and transports over the industrial era. Global Biogeochem. Cycles 28, 631– 647 (2014), doi:10.1002/2013GB004739.

Joos, F., Roth, R., Fuglestvedt, J. S., Peters, et al. Carbon dioxide and climate impulse response functions for the computation of greenhouse gas metrics: a multi-model analysis. Atmos. Chem. Phys. 13, 2793–2825 (2013), doi:10.5194/acp-13-2793-2013

Winton, M., S. M. Griffies, B. L. Samuels, J. L. Sarmiento, and T. L. Frölicher Connecting Changing Ocean Circulation with Changing Climate. J. Climate, 26, 2268–2278 (2020), doi:10.1175/JCLI-D-12-00296.1