We are grateful to the referee for their constructive feedback and the time they spent reviewing our manuscript. This helped us to improve the presentation, while results and conclusions remain unchanged. Below are our responses (in **bold**) to the referee comments (in *italics*).

Referee #1

Fundamental remarks:

1. Weathering vs sedimentation: It is said that when sedimentation is included in the model the atmospheric $\Delta 14C$ strongly decreases in comparison to an atmosphere-ocean model version only. I believe this is naming the wrong process. My understanding of the model description is, that weathering is the process that brings 14C-free C into the system, so it is carbonate weathering, that is fundamental for the 14C cycle. It is clear that once weathering input (of alkalinity and DIC to the ocean) is considered also sedimentation as sink needs to be implemented (otherwise the carbon cycle would run away with an ocean accumulating alkalinity, and subsequent changes to atmospheric CO2 levels), but sedimentation is not the important process here that changes $\Delta 14C$. This might then also lead to a different name of the model configuration now called OCN-SED.

Weathering fluxes are an important component of the global carbon cycle and input to the ocean from terrestrial weathering of carbonate and silicate rocks and volcanic emissions is included in our model to balance material loss by burial of particulate organic matter, calcium carbonate, and opal at the sea floor. See also our response to the following comment #2. Potential further changes in weathering fluxes may have been very important for past changes in δ^{13} C but are largely irrelevant for the present study with its focus on Δ^{14} C. What we are interested in demonstrating here is that a change in the global ocean carbon inventory linked with the weathering/sedimentation balance is a potentially important factor affecting atmospheric Δ^{14} C levels. Also note that the nomenclature of the model configurations is meant to describe the global carbon reservoirs (model components) under consideration, so, e.g., model configuration OCN-SED includes the ocean model and the sediment model. We will modify the text to make this clearer.

2. No details on weathering are given, but since it is said, that 14C-free C is entered via weathering I have to assume, this implies carbonate weathering. However, it need to be clarified (and maybe corrected?), that in carbonate weathering, 50% of the carbon that enters the ocean as weathering product (bicarbonate ion, HCO-3, which changes DIC and alkalinity in the ocean) comes from rocks (14C-free), and 50% has its origin in atmospheric CO2 with its atmospheric 14C-signature. For silicate weathering, also bringing HCO-3 to the ocean, 100% of the carbon has its origin in the atmosphere. Is silicate weathering considered? For details see, for example, Colbourn et al. (2013). Without checking on recent updates, I believe both silicate and carbonate weathering contributed about a similar amount of HCO-3 input into

the ocean. At least in a study some years ago (Hartmann et al., 2009) in present day weathering the CO2 consumption is twice as big in silicate than in carbonate weathering, but since in carbonate weathering 50% of the C has its origin from rocks, both processes should contribute about the same. Since weathering is the relevant process for this paper more details on its implementation in the model should be included. From the Appendix I understood, that weathering rates are constant in time, but please give their numbers, which would be especially of interest to other modellers doing similar things. Also consider in a discussion, that missing temporal changes in weathering rate might be one reason why reconstructed $\Delta 14C$ (and CO2) is not met with simulations. Having found, that the input of 14C-free carbon to the system is so important for an understanding of Δ 14C brings me also to the quesion if 14C-free CO2 outgassing from volcanos is considered, which might have similar effects on 14C. I understand that this has been investigated previously with the Bern3D model (Roth and Joos, 2012), but with focus on 13C. Maybe some more insights from previous simulations are possible here, at least in a discussion. At least please mention the applied CO2 volcanic outgassing rates. Note, that there is a fundamental, analytical derived solution from the steady state assumption on volcanic CO2 input being 50% of the CO2 consumption by silicate weathering, which is of relevance for times longer than 100 kyr (briefly mentioned in Munhoven and François (1996) or in depth discussed on pages 80-81 of Munhoven (1997), http://www.astro.ulg.ac.be/~munhoven/en/PhDIndex.html). For shorter periods such as the last 50 kyr considered here, differences from this numbers are certainly possible, but this relationship gives a rough guideline, and might explain long-term drifts in the C cycle, if not obeyed. Taken together, I have the impression, that no silicate weathering, and also no volcanic outgassing of CO2 is considered here, which would indicate according to this theory no drift in the system, but also the missing of two important processes. If so, I am not saying, these should be implemented in the revision, but it needs to be stated clearly if and how they are (not) included. How does your weathering flux compare to others, e.g. Fig 7 in Brovkin et al. (2012) or Colbourn et al. (2013)?

We are afraid that there has been a misunderstanding with regard to the model representation of terrestrial weathering fluxes. Apparently the referee missed the description of the sediment model and weathering fluxes given in Appendix A (lines 756-771). We will include the preindustrial steady-state values for weathering rates in Sect. 2.2 and provide additional details on the representation of terrestrial weathering in Appendix A. For convenience, the approach is described here in detail.

The Bern3D model simulates net ocean-sediment exchange, sediment stocks, and burial fluxes from the ocean sediments to the lithosphere of P, Si, Alk, DIC, DI¹⁴C, and DI¹³C using a 10-layer ocean sediment model. In steady state, the net loss fluxes from the ocean to the sediments and lithosphere are compensated by corresponding input fluxes to the ocean, termed "weathering fluxes". These input fluxes are thought to represent the fluxes from weathering (dissolution) of carbonate and silicate rocks on land, of phosphorous release by rock weathering, and from CO₂ emissions from volcanic activity. The ocean inventories of P, Si, and Alk and atmospheric CO₂ and its isotopic signature are prescribed during the atmosphere-ocean-sediment spin-up. Any loss of P, Si, Alk, DIC and DI¹³C by net fluxes to the sediments is compensated by a corresponding input flux during the spin-up phase. Input fluxes are added uniformly to the coastal surface ocean. At the beginning of transient simulations, the global input fluxes of P, Si, Alk, DIC and DI¹³C are set equal to the burial fluxes diagnosed at the end of the model spin-up. These input fluxes are jointly denoted as "weathering fluxes". Radiocarbon is transferred from the ocean to the sediments and lithosphere where it decays, but no radiocarbon is added to the ocean from "weathering" as old rocks are radiocarbon free (or "radiocarbon dead").

The preindustrial spin-up results in steady-state values for weathering-derived inputs (and hence steady-state burial rates) of DIC, Alk, P, and Si of 0.46 Gt C per year, 34.37 Tmol HCO₃⁻ per year, 0.17 Tmol P per year, and 6.67 Tmol Si per year, respectively. These values are within the range of observational estimates (see, e.g., Jeltsch-Thömmes et al., 2019, Table 1). The weathering input of Alk as HCO₃⁻ is also comparable to the global riverine bicarbonate flux presented in Brovkin et al. (2012), very close to their interglacial estimates (36 to 38 Tmol per year) but lower than their LGM estimates (almost 50 Tmol per year).

These input fluxes may be further attributed to the weathering of organic material, CaCO₃, and CaSiO₃ on land, and to volcanic CO₂ outgassing. The flux of phosphorus (P) is assigned to weathering of organic material, and the related carbon (C) and Alk fluxes are computed by multiplication of the P flux with the Redfield ratio for organic matter stoichiometry (C:P:Alk = 117:1:-17). Similarly, the silicon (Si) flux is assigned to CaSiO₃ weathering, and the related Alk flux is computed using Si:Alk = 1:2 based on the simplified CaSiO₃ weathering reaction: $2CO_2 + CaSiO_3 + H_2O \rightarrow 2HCO_3^- + Ca^{2+} + SiO_2$ (Colbourn et al., 2013). As Colbourn et al., we sidestep the carbon flux from the atmosphere to the ocean. The remaining Alk flux is attributed to CaCO₃ weathering with the stoichiometric ratio C:Alk = 1:2 based on the CaCO₃ dissolution reaction: CO₂ + H₂O + CaCO₃ \rightarrow Ca²⁺ + 2HCO₃⁻. The volcanic outgassing flux is the remaining flux needed to balance the C input flux. The diagnosed fluxes at the end of the spin-up are 0.24 Gt C per year for terrestrial weathering of organic material, 0.13 Gt C per year for terrestrial CaCO₃ weathering, 0.09 Gt C per year for volcanic CO₂ outgassing, and 6.67 Tmol Si per year for terrestrial CaSiO₃ weathering.

For simplicity, we kept weathering fluxes constant during transient simulations, and weathering feedbacks (Jeltsch-Thömmes and Joos, 2020; Colbourne et al., 2013) were not enabled in this study. This does not affect our results and conclusions. First, changes in weathering fluxes have no influence on the ocean radiocarbon inventory as weathering fluxes are "radiocarbon dead". Second, the impact of potential changes in weathering fluxes on atmospheric CO_2 and the ocean and sediment carbon inventories, which would influence atmospheric $\Delta^{14}C$, is implicitly considered in our sensitivity experiments where CO_2 and carbon inventories are forced to vary. Third, we note that there is a large uncertainty in the dissolution rates of carbonate and silicate rocks on land (terrestrial weathering) over time, and that these weathering reactions represent a very long-term sink of atmospheric CO₂. In particular, weathering of silicate rocks on land is occurring too slowly (on a time scale of hundreds of thousands of years) to be important on the time scale relevant for this study (~50,000 years). Furthermore, Roth and Joos (2013) demonstrated that even massive changes in volcanic emissions cause changes in atmospheric Δ^{14} C that are much smaller than the discrepancies between reconstructed and modelled Δ^{14} C.

Finally, we would like to point out that because DIC and Alk are conservative with respect to changes in state (temperature, salinity, and pressure) during mixing, both are carried as tracers in ocean carbon cycle models like the Bern3D. Together they completely determine the CO₂ system in seawater (H^+ , pCO_2 , $H_2CO_3^*$, HCO_3^- , and CO_3^{2-}), using the well-known carbonate chemistry routines. These parameters can be used to compute, e.g., air-sea CO₂ and ¹⁴CO₂ fluxes or the saturation state of seawater with respect to CaCO₃.

3. Earlier simulation studies have shown, that to get the 14C cycle right, one needs to have the C cycle right as well. Köhler et al. (2006) has shown that previous studies (Beck et al., 2001; Hughen et al., 2004) focusing only on 14C, but showing no simulated CO2, they therefore have very likely some deficits. For atmospheric Δ 14C especially the air-sea gas exchange is important, which depends similarily on the gas exchange velocity (kw, which is considered here in sensitivity experiments), but also on the CO2 gradient between atmosphere and surface ocean. This implies that whenever simulated CO2 differs from reconstructions there will also be an offset in simulated Δ 14C from data. In a recent simulated (Heaton et al., submitted). There, the importance of time-dependent changes in CO2 has been as important for the simulated surface ocean Δ 14C as that of climate change (temperature change, ocean circulation change etc), which via gas exchange would also feedback to atmosheric Δ 14C. This is unpublished so far, but since it is submitted and will probably be available in due time I nevertheless mention it here.

We thank the referee for bringing this very interesting-sounding work to our attention. We look forward to reading it once it becomes publicly available.

It is well known that gross isotopic air-sea fluxes scale with atmospheric CO_2 and the ¹⁴C/C ratio. In 6 of our 8 model carbon cycle scenarios, parameter values were selected to reproduce low glacial atmospheric CO_2 concentrations. Note also that atmospheric CO_2 was prescribed in the model runs where the ¹⁴C production rate is deconvolved from the model results (i.e., Sect. 3.4). Thus, temporal changes in atmospheric CO_2 are taken into account by our model simulations.

4. The coauthor Florian Adolphi is also coauthor of the nows ubmitted IntCal20 effort (updating the atmospheric Δ 14C record), (Reimer et al., submitted) and should therefore be aware of the large changes which occur between IntCal13 and IntCal20, namely the amplitude of the Δ 14C maxima around 40 kyr increases in IntCal20 towards the Hulu Cave numbers. Maybe this should be briefly discussed in an outlook.

The backbone of the new IntCal20 calibration curve is the Hulu Cave Δ^{14} C dataset from Cheng et al. (2018) that we use in this work, since IntCal20 is not yet published. Essentially all datasets underlying the IntCal20 curve are tied to the Hulu Cave record, either via time scales (Lake Suigetsu plant macrofossil data) or marine reservoir corrections (marine records). Hence, IntCal20 looks more or less like the Hulu record. Since we are not discussing the fine-scale structure of the record but rather the largescale changes in Δ^{14} C, using IntCal20 would not significantly impact our conclusions.

Minor issues in chronological order:

1. The decay constant of 14C used here is based on a halflife of 14C of 5700 yr (here) but of 5730 yr in Intcal13 (and IntCal20) which produced the atmospheric Δ 14C record. If you consider the decay of 14C over 40 kyr (the time of the maximum in atm Δ 14C) with either 5700 vs 5730 yr you get a 2.5% smaller number when based on 5700 yr, although the halflife time differed only by 0.5%. This difference is small when compared to the difference of IntCal13 and the Hule Cave data, but should nevertheless be mentioned.

As mentioned in Fig. 1 caption, reconstructed Δ^{14} C values taken from the IntCal13 calibration curve and the Hulu Cave dataset were adjusted to the presently accepted value of the radiocarbon half-life (5700 years), allowing comparison with our modelled Δ^{14} C values.

2. line 173: Please be specific, how 14C is fractionation corrected.

Radiocarbon measurements are generally reported as Δ^{14} C which includes a correction for fractionation effects. The measured δ^{13} C value is used to remove the effects of isotopic fractionation. As indicated in Sect. 2.2, this model study simulates a ¹⁴C concentration that is "fractionation corrected". What we mean by this is as follows. To model ¹⁴C, the Bern3D neglects effects due to fractionation during gas exchange and photosynthesis, which means that model results for Δ^{14} C are directly comparable to measurements reported as Δ^{14} C. If we were interested in dealing with absolute values of the ¹⁴C concentration over time, then a correction is needed to account for fractionation effects, using the following equation (see Orr et al., 2017, Eq. A3 and associated discussion in the Appendix):

$${}^{14}C = \left({}^{14}C_{model} / \left[1 - 2\left(\frac{\delta^{13}C + 25}{1000}\right)\right]\right){}^{14}r_{std}$$
(1)

For ¹³C, the Bern3D includes fractionation effects during gas exchange and photosynthesis. Eq. 1 together with modelled δ^{13} C values could be used to compute corrections for atmospheric and oceanic ¹⁴C, were we interested in looking at absolute values of the ¹⁴C concentration.

3. line 180-181: Weathering is prescribed as constant input of DIC, but no 14C. See fundamental comment above, but no matter where the C of weathering comes from the input in the ocean should be a HCO-3, changing both DIC and alkalinity in the ocean. I hope this is only a too simplified description here, but has implemented correctly in the model. Please revise.

This comment has already been addressed in our response to comment #2 of the referee's "fundamental remarks".

4. line 184: For the preindustrial spin-up CO2 is fixed to 278.05 ppm. Why this number, would not 278 ppm do the same job? Is this OCMIP protocol? Also: What would be the internally calculated preindustrial CO2? By prescribing CO2 concentration during spin-up C is added or extracted from the simulated system, which might be a potential source of bias. How long is the spin-up time?

This is irrelevant. The reason for this atmospheric CO_2 value is that it is the nominal value for year 1750 in one of our in-house CO_2 data compilations, but yes, holding atmospheric CO_2 constant at 278 ppm would do the same job. In the preindustrial spin-up simulation, the model is run to equilibrium over a ~50,000-year integration, as mentioned in line 186. During this spin-up, atmospheric CO_2 is held constant at 278.05 ppm and $\Delta^{14}C$ at 0 permil. These are the values that the atmospheric carbon and ^{14}C tracers see. The ocean carbon and radiocarbon inventories change in response to this forcing. After spin-up the ocean is in equilibrium with the atmosphere.

5. line 229: Consider citing the underlying ice core paper, from which the greenhouse gases splines provided by Köhler et al (2017) have been calculated.

Agreed. We will cite the underlying ice core papers and Enting (1987) for the spline smoothing method.

6. line 232: Global benthic δ 18O is not only a global ice volume proxy, but has also a considerable contribution from deep ocean temperature, see papers of the van de Wal group from Utrecht Universitiy on the deconvolution (e.g. Bintanja and van de Wal, 2008). Taken this knowledge into consideration, would this change your approach how sea level is changing? You might also discuss how different sea level reconstructions vary, e.g. see three different sea level reconstructions in Fig 1f of Hasenclever et al. (2017), and what this uncertainty in sea level might introduce into your approach.

No, this would not significantly change our approach or our results. As explained in the manuscript, we use the global benthic δ^{18} O stack to scale the ice sheet size for periods where no reconstructions are available. The tie points of this scaling, i.e., the LGM and preindustrial ice sheet reconstructions, remain unaffected by the scaling method, i.e. whether the scaling is done with δ^{18} O corrected for deep ocean temperature or not. Ice sheet size is important for albedo, salinity and latent heat fluxes, but has little influence on modelled atmospheric Δ^{14} C (e.g., the small difference between scenarios MOD and PAL in Fig. 8).

7. lines 223-236: Model description says that greenhouse gas radiative forcing has been taken from data, thus I assume that CO2 seen by the carbon cycle is never prescribed, but always model-internally calculated. Please state this explicitly (or the correct version of this sentence, if this was not the case). However, you might also consider one scenario in which CO2 is prescribed for the C cycle from data (similar as in Butzin et al., 2017, in which atmospheric CO2 and Δ 14C has been prescribed by data), since this would bring your simulated C cycle as close to observations as probably possible, which might further reduce the bias in 14C (see also fundamental remarks above).

Although the radiative forcing for CO_2 is prescribed, the atmospheric CO_2 concentration is allowed to evolve freely, except in the simulations described in Sect. 2.5. We will clarify this point in Sect. 2.4.

As discussed in lines 302-315, our approach to estimating the ¹⁴C production rate over the last 50 kyr relies on model simulations forced by reconstructed changes in atmospheric Δ^{14} C and CO₂ as well as 7 different carbon cycle scenarios. None of the model runs are able to reproduce the reconstructed variations in ¹⁴C production during the last glacial, especially between 32 and 22 kyr BP (see Sect. 3.4 and Fig. 10 and 11). Thus, the discrepancy between reconstructions and model results remains even when prescribing atmospheric CO₂.

8. Obtained surface reservoir ages (Fig 8c) might be compared with data and other models, e.g. see Butzin et al. (2017); Skinner et al. (2019). Benthic-atmospheric offsets (Fig 8d) might be compared for the LGM with the data compilation of Skinner et al. (2017). Note, surface reservoir ages might vary a lot as function of latitude, so this Fig 8c needs more information on averaging; even better: might be revised and thus restricted to sea ice-free areas only.

We agree comparison with measurement- and model-based estimates of radiocarbon reservoir age offsets from, e.g., Skinner et al. (2017) and Butzin et al. (2017), is a missed opportunity. It was a sacrifice made to reduce the length of an already very lengthy manuscript. Nonetheless, some intriguing points can be made by such a comparison, so we will incorporate it into Sect. 3.3 and Fig. 8.

Comparison of our LGM B-Atm age offset estimates from runs CIRC, VENT, and VENTx (range of 3682 to 3962 ¹⁴C years) with the compiled LGM marine radiocarbon data of Skinner et al. (2017) demonstrate that the carbon cycle scenarios are extreme, although it should be noted that they consider a wider depth range (~500 to 5000 m) of the ocean than we do. Skinner et al. (2017) predict a global average LGM B-Atm value of \sim 2048 ¹⁴C years, an increase of \sim 689 ¹⁴C years relative to preindustrial. Turning our comparison to surface reservoir ages, we note that our global average LGM surface reservoir age of ~1132 ¹⁴C years from runs VENT and VENTx is comparable to the ~1241 ¹⁴C years obtained by Skinner et al. (2017) for the LGM. The model-based estimates of surface reservoir age from Butzin et al. (2017) indicate a much lower LGM value of ~780 ¹⁴C years, and values ranging from 540 to 1250 ¹⁴C years between 50 and 25 kyr BP. Note that these estimates are based on model-simulated values between 50°N and 50°S. If the polar regions are included in the calculation (see Fig. 8c), their surface reservoir age estimates become comparable to our glacial values (range of 911 to 1354 ¹⁴C years), and between about 34 and 22 kyr BP can exceed them, including even those from model runs VENT and VENTx, unless atmospheric Δ^{14} C and CO₂ are prescribed (dashed colored lines in Fig. 8c). Interestingly, this is also roughly the time period where our deconvolutions of the IntCal13 and Hulu Cave Δ^{14} C records give production rate estimates that are about 17.5 percent higher than the reconstructions, which indicates at the very least this is an important piece of the puzzle of the glacial-interglacial Δ^{14} C problem, given that the effect of upper ocean stratification and/or sea ice on air-sea gas exchange is particularly important for surface reservoir ages.

Comparison of our surface reservoir ages with estimates from Skinner et al. (2019) will have to await a future study. A clear picture of the spatiotemporal evolution of the global average surface reservoir age has yet to emerge, but the regionally distinct patterns as demonstrated by Skinner et al. (2019) have important implications for the calibration of marine radiocarbon samples. These results need to be scrutinized more carefully and investigated in more detail with models, and with experiments specifically designed with this question in mind.



Fig. 8. Modelled records of atmospheric (a) Δ^{14} C and (b) CO₂, compared with their reconstructed histories (black and dark blue lines). Also shown are modelled records of the global average (c) surface reservoir age and (d) B-Atm ¹⁴C age offset, compared with a recent compilation of LGM marine radiocarbon data (dark blue squares) by

Skinner et al. (2017) and model-based surface reservoir age estimates between 50°N and 50°S (solid black line) and across all latitudes (dashed black line) from Butzin et al. (2017), as well as (e) ideal age and (f) apparent oxygen utilization (AOU). Colored lines show the results of model runs using the mean paleointensity-based ¹⁴C production rate and the eight different carbon cycle scenarios described in Sect. 2.4 and Table 1. The gray envelope in (a) shows the uncertainty (2 σ) from all production rate reconstructions and carbon cycle scenarios, providing a bounded estimate of Δ^{14} C change. The dashed colored lines in (c) show the surface reservoir age results from VENT and VENTx where atmospheric Δ^{14} C and CO₂ are prescribed. Radiocarbon ventilation ages are expressed here as radiocarbon reservoir age offsets following Soulet et al. (2016) which are used extensively by the radiocarbon dating community.

9. Please state somewhere the absolute (PI) values of those parameters which are changed in your sensitivity experiments, maybe in Table 1?

We would like to direct the referee, and the reader, to the appendix of Roth et al. (2014) for the Bern3D model parameter set.

10. Table 2: 14C production rates is given in relative units, relative to what? Probably preindustrial state. I also do not remember if the 14C production rate in absolute numbers is once given in the manuscript for preindustrial state, please insert somewhere.

Our model-based records of the global production rate of ¹⁴C are in units relative to the preindustrial value, as mentioned in lines 416-417. We will include the preindustrial steady-state absolute value of 443.9 mol ¹⁴C per year (1.66 atoms cm⁻² s⁻¹) in Sect. 2.2.

11. Fig 4 captions does not need a description of the different colors of the lines, since a legend is given in the figures themselves.

Here we reference the colored lines in order to remind the reader that their labels refer to the model configurations representing different combinations of global carbon reservoirs, which is important when comparing the response of Δ^{14} C to the step changes.