Reply to the Referee #2

We would like to thank the reviewer for the thoughtful and constructive comments.

Q1: Page 1 Line 16: I consider this sentence as a bit unclear or vague. Please state more explicitly which models might show a bias, and which results of those models might therefore be less reliable than previously thought.

A1: First, we will replace the word 'biases' with 'uncertainty' since we intended to address the general shortcomings of such models. Second, we will add the following sentence to the abstract. "For example, an ocean model that does not treat sedimentary processes depending on the chemical composition of the ambient water can overestimate the amount of remineralization of organic matter in the upper sediment under an anoxic environment, which would lead to lighter $\delta 13C_{DIC}$ in the bottom water."

Q2: Page 2 Line 24-25: "Up to the present, no fully-coupled comprehensive climate model has been coupled with a sediment diagenesis model for longer time-scale applications (e.g., the glacial-interglacial variations)." -> Has this approach been used on shorter (e.g. centennial) time scales? Can you give examples, and how does your approach differ from them?

A2: We will revise the sentence as follows: "To our knowledge, a fully-coupled comprehensive climate model including a sediment diagenesis model has been applied to millennial time scales only (Jungclaus et al., 2010). Here we aim at applying such a model to time scales of tens of millennia, having glacial-interglacial variations in mind."

Q3: Page 5 Line 24-31: Can you show a figure, possibly in the online supplement, that proves that your time step was sufficiently small and your integration period sufficiently long to show something like a convergence of the sediment-water fluxes in the end (for all but the 14C of course)?

A3: In terms of DIC flux back to the ocean, the difference between a run with the original time step and a run with only 1/10 of the original time step is smaller than 0.5% for most grid cells. We will add figures to show this in the supplementary material. Time steps shorter than 1 year do not make sense because the input from CESM is annually averaged.

It should be noted that the CESM-MEDUSA coupled simulation (EXCPL) is not a "steady-state" run but a "transient" run where the model state evolves. Therefore, the length of the MEDUSA runs is not determined by the convergence of a model state but by the coupling interval.

Q4: Page 7 Line 3-4: "which would lead to the overestimate of biological production" -> "which would lead to an overestimation of biological production"?

A4: The sentence will be corrected.

Q5: Page 9 Line 15-17: "Otherwise, one would need to translate records obtained from sediments into corresponding variables of the ocean model, which would introduce an additional source of uncertainty to the model--data comparison." -> You have an opposite translation by the MEDUSA model: ocean model variables are translated to sediment records. Why is this less uncertain than the other way around?

A5: The "forward modeling" by MEDUSA is beneficial because it provides a process-based translation rather than an empirical translation that would be inevitable without such a sediment model. Therefore, the point is not the direction of translation but the way of translation. We will revise the sentence as follows. "Otherwise, one would need to translate records obtained from sediments in an empirical way to corresponding variables of the ocean model, which would introduce an additional source of uncertainty to the model-data comparison."

Q6: Table 1: Would it make sense to add a third column for the values in the EXORG run?

A6: We will add the following two tables to compare EXCPL and EXORG in Section 3.2 to show that the two experiments are comparable in general in terms of globally-integrated quantities (except for the burial flux). In Section 3.1, we would rather focus on the comparison between EXCPL and observations as in the current version of the manuscript.

Table 1. Globally-integrated annual mean deposition flux of particulate matter to the sediment and their burial flux (in parentheses) at the end of EXCPL and EXORG.

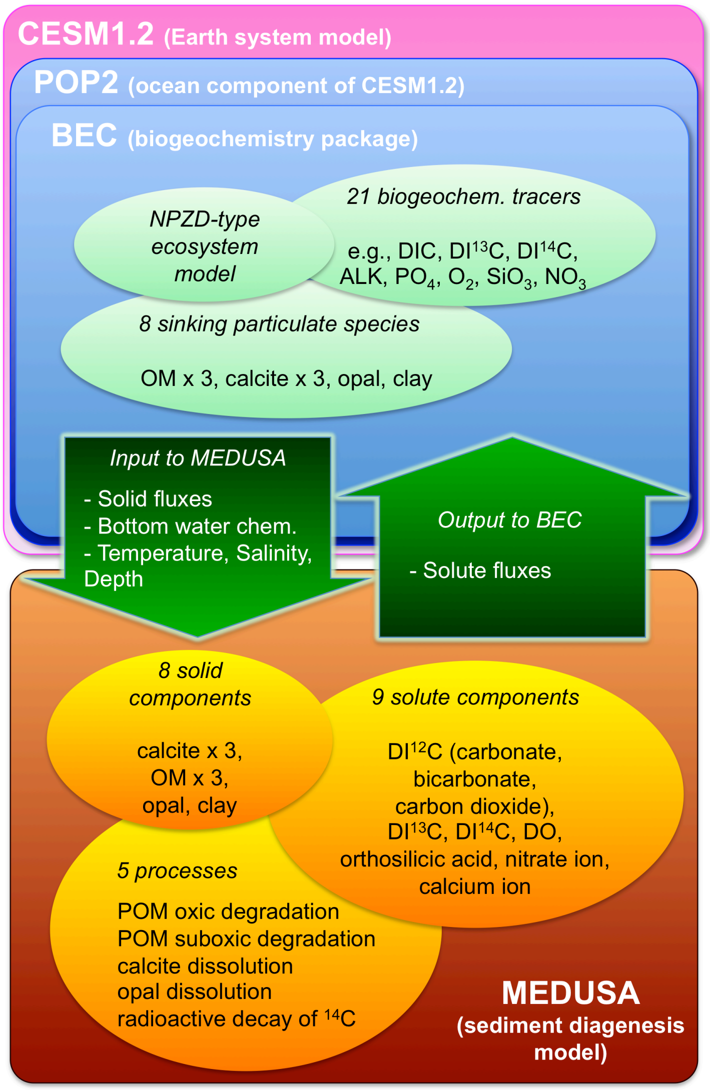
	EXCPL	EXORG
POC (GtC/y)	0.57 (0.091)	0.51 (0.12)
$CaCO_3$ (GtC/y)	0.39 (0.082)	0.38 (0.14)
Opal (Tmol/y)	46 (0.72)	45 (3.4)

Table 2. Total inventories in the global ocean of DIC, ALK, and PO_4 in EXCPL and EXORG. Values averaged over the last CESM run (10 surface years) are shown.

	EXCPL	EXORG
DIC (GtC)	3.660×10^4	3.657×10^4
ALK (Peq)	3.201×10^3	3.201×10^3
PO_4 (Pmol)	2.948	2.923

Q7: Figure 1: Why are the state variables only listed for the MEDUSA model and not for the BEC model? Probably the list of processes might be too long, but at least the state variables would give an indication of the model complexity for those not familiar with BEC.

A7: We will update the figure as follows.



Q8: Figure 6: Having this figure separate from Fig. 5 and using changed color scales makes the comparison very hard. And the improved behaviour of the model using the coupling is the main point of your manuscript. If you think the subfigures become too small if you put all three weight fractions for EXORG, EXCPL and OBS into one figure, you might consider one figure for each weight fraction but then containing EXORG, EXCPL and OBS?

A8: We will re-arrange the figures according to the reviewer's suggestion.