Thank you for the constructive review of our manuscript. We answered your questions and followed your advices whenever possible. In the following lines we copied your text as bold and gave our remarks in detail.

GENERIC COMMENT the paper presents an interesting implementation of a vertically structured benthic model to estimate the alkalinity fluxes from the Southern North Sea sediments. This is a challenging topic that needs to be addressed and authors are commended for this. The methods are generally sound, with some clarification needed and some suggestions for improvements provided. The results are presented clearly, the discussion could benefit of some more in-depth analysis, particularly on the role of pelagic primary productivity and on the relevance of the alkalinity fluxes for the entire ecosystems.

Thank you for the hints regarding primary production and the role of alkalinity fluxes. During revising the text and answering your specific comments primary production was highlighted. One example is the positive feedback of pelagic production when enhanced nitrogen effluxes occur. In our model the direct effect of alkalinity effluxes on the whole ecosystem is restricted to the pelagic carbonate system and the air-sea flux of CO_2. The indirect effect of connected effluxes of nutrients and oxygen is discussed now in a more comprehensive manner.

SPECIFIC COMMENTS: section 2.2.5 and table 1: why I appreciate that turnover rate for deep ocean would not be suitable for shelf seas, authors did not specify how they define the new values, if via calibration (against which observations?) or with literature (which references?).

As for the North Sea the annual budget of carbon export into the sediment and the efflux of DIC is nearly in balance, the changes of the reaction rates aimed at this target. In a first step we replaced the spatial uniform porosity used in the open ocean sediment module (0.85 first layer) to observation-based values for the southern North Sea (between 0.3 – 0.51). Additionally, the constant diffusion rate was replaced by a porosity, temperature and substrate dependent formulation. These changes made it necessary to adjust the turnover rates. As shelf areas are hardly resolved in the coarse resolution of the global model seasonality of organic matter fluxes and DIC effluxes were no tuning criteria. We changed the text in section 2.3.5 accordingly.

Lines 144-146: authors assumed that in advective sediments the coefficient for diffusion is increased tenfold. They state that this has been determined by several sensitivity analysis, but they did not state what were the criteria of the sensitivity analysis (stability? calibration? Something else?) more details are needed

As said before the target was to achieve a more or less equal annual DIC efflux and POC influx. Furthermore, the seasonality of the DIC efflux should resemble the known temporal evolution. An upper constraint for the diffusion coefficients was the penetration depth of significant oxygen concentrations. Below 1 cm depth almost no oxygen should be detected. We changed the text accordingly.

lines 175-178: authors claims that a reduction of 10% of riverine input of nitrogen corresponds to a "pristine" scenario without anthropogenic influence. Authors cite a paper to corroborate this

assumption. However it seems to me that 10% is a bit of an underestimation for such and industrialised area.

This is a misunderstanding. We reduced the riverine input to 10 % of the anthropogenic value.

Line 179-183: do the rates in the "plate run" scenario are comparable with those of table 1? I appreciate that the equation will be different and therefore the value of the parameter, but reporting these for a comparison would help in understanding how much of the difference is due to model structure and how much to simple parameter values

The carbon degradation rate of the plate run is defined as a time constant (2.8E-2/d), whereas the aerobic rate of the reference run is an oxygen dependent rate (r1 =2 E-10 m3/(mmol [O2] s). For an off-shore station (54.4 °N, 7.4 °E) with a typical oxygen concentration of 100 mmol/m3 r1 results in a comparable rate (1.7 E-3 /d). In contrast to the plate run where nearly all POC is dissolved over one year POC concentrations in the upper most sediment level of about 0.35 E6 mmol C/m3 in winter and 0.48 E6 mmol C/m3 in summer are found in the reference run. This results in a winter flux of 1.2 mmol C/(d m2) and a summer flux of 1.6 mmol C/(d m2). The corresponding fluxes of the plate run were 0.36 mmol C/(d m2) in winter and 2.3 mmol C/(d m2) in summer. We added some text in section 4.4 accordingly.

line 220 and following. I'm not sure that providing a point-to-point comparison on a single day is the more effective way to assess the model. Small shift in phenology (not rare in coupled biogeochemical models) could result in a significant error that could not be related to the benthic model rather to error in the physics or in the forcing. I would suggest to compare the observations with a longer temporal means (monthly?) and to discuss the uncertainty. Also, while visual comparison can be appealing, they are not much informative: I suggest to provide also measures of the actual fit. For example, in relative error term, I'm not sure figure 6 shows a much better fit.

We switched from the analysis of 15 September to September means. But this did not change much because especially in September no big changes were expected (compare Fig 11b). We added an error analysis (definition in section 2.6).

Section 4.2: authors seem to suggest that the strong undersaturation of pCO2 in the German bight is driven by the strong alkalinity fluxes from benthos. I'm not entirely sure that the simple co-location of the two is enough to establish a causal link. For instance what's the role of pelagic primary production (PP)? A high PP could explain both the strong undersaturation (DIC is fixed into plankton) and the alkalinity fluxes (due to strong POM settling and associated processes in the benthic environment). Have the author tested to turn off the benthic fluxes of TA and check the consequences in the delta pCO2 signal?

This was a very helpful hint: in a further sensitivity run we switched off the benthic TA fluxes in the reference run. Consequently the pCO2 values increased. But the coastal low pCO2 did not fully vanish. We concluded that both the primary production and the benthic TA effluxes were responsible for the near-coast low pCO2 values. We added some text accordingly.

Section 4.3: authors said that it's astonishing that the model simulates higher benthic pelagic fluxes under higher porosity, when the diffusion coefficient is lower. Do authors have any suggestion on what are the mechanisms driving this?

To understand this contra-intuitive dynamics we compared the model results of the high porosity run (HP with 0.51) with the low porosity run (LP with 0.3) in the first and second spinup year. At the beginning of the first spinup year all conditions are the same. Until spring the flux of oxygen into the sediment was lower in HP because there the effective diffusivity was lower than in LP. The lower oxygen content in HP stimulated the benthic anaerobic processes. At the end of the first year this resulted into a higher efflux of NH4 from the sediment in the HP scenario. The higher NH4 efflux of the HP scenario was not compensated by the higher NO3 flux into the sediment. At the end of the year more DIN was in the pelagic water column in the HP scenario than in the LP scenario. In the second year this surplus of DIN stimulated higher primary production for the HP scenario. The corresponding enhanced particle export additionally increased the benthic-pelagic fluxes. The loss of N2 due to enhanced denitrification was compensated by the larger NH4 efflux. These deviating dynamics are even stronger at stations with lower pelagic DIN concentrations. We added some text accordingly.

Section 4.4: this section is important to understand the need for detailed model. However authors simply state the difference between the two models, without trying to tease out the reason behind that, particularly in regard to the difference in the seasonal signal

There are several reasons for the deviating seasonal cycle of DIC efflux. In general the less pronounced cycle comes about

- the structure of the 3d-sediment model which leads to a combination of multiple time scales acting on the reaction rates due to diffusive processes between the layers,
- the fact that the remineralization fluxes do not produce immediately effluxes. In case of the 3d-sediment model the dissolved compounds have to be transferred via diffusion into the pelagic system,
- the reservoir effect in the 3d-sediment model: Whereas the 2d-plate model more or less all POC is degraded after winter, in the 3d-sediment model a relative high POC concentration remains.

We added text accordingly.

technical comment: please translate "gedankenexperiment" in English.

done