Detailed responses to review by A. Dale (reviewer comments are included in black, responses in blue font)

General comments

This is an interesting study that focuses on coupling benthic processes in biogeochemical circulation models of the water column. I like the optimization approach and can see how it would be useful in my own work. I agree completely with the authors that smart ways are needed to parameterize benthic process in a computationally efficient manner, and this is a topic close to my own heart. I have some questions and comments that the authors should attend to, but there are no major flaws as far as I can see. This is a sound paper that meets the scope of the journal and I recommend publication. I hope that my comments are fair and that they will add some value to the paper.

Response: Thank you for your positive assessment. We appreciate the constructive comments and provide detailed responses below.

Specific comments

Comment:

1. The only major critique I would like to raise, and would like the authors to respond carefully to, is that the steady state meta models are applied to a highly dynamic environment with huge intra-annual variability in POM flux and bottom water O2 (Fig. S1). The benthic functions are not dynamic i.e. they do not account for the storage of POM and the decoupling of solute fluxes from POM fluxes. POM deposited on the seafloor will degrade over a wide range of time scales, rather than instantaneously as the present functions assume. This approach would be permissible in a setting exhibiting less temporal variability, but not here. The authors are aware of the potential error incurred (cited Soetaert et al., 2000) but this aspect is barely discussed.

For instance, the NH4 flux responds almost instantaneously with POM flux in their model, but since OM does not degrade instantaneously, the NH4 flux in reality would lag behind POM flux and be more attenuated and without the pronounced flux spikes. It would really enhance the appeal of this paper if these ideas were incorporated into a couple of discussion paragraphs, with an estimate of the error in not accounting for the temporal. This result of this exercise would be a great value to both the pelagic and benthic modelling community alike, me included.

One assumption is to make the return flux of oxygen and nutrients a function of the instantaneous organic matter (OM) deposition flux, whereas in reality nutrients are produced and oxygen consumed as a result of OM mineralization. Instantaneous deposition flux is a good proxy for mineralization only if the OM decay rate is very high. Yet the model settings are such that 26% of OM is rather refractory, which is inconsistent with this assumption. It also means that the time-lag induced between deposition and sediment-water nutrient and oxygen fluxes is not taken into account, i.e. the memory of the sediments is ignored. Hence it is not surprising that the modeled deposition fluxes are not well suited to reproduce the measured oxygen and nutrient fluxes. It is also the reason why the modeled O2 flux follows the POM deposition so closely (P16 L12). In the recommended procedure of S2000, the sediment model dynamically describes two OM fractions (vertically integrated), and the meta model therefore prescribes the dissolved fluxes are a function of OM mineralization. The reason for choosing a reflective boundary condition, and the implications, are however not discussed in this manuscript.

Response:

We agree that the steady state assumption does not represent the memory of the sediment and thus may not be realistic in this dynamic environment. Reviewer 2 (Karline Soetaert) raised a similar concern in her comments. In order to address this problem, we decided to modify our approach using time-dependent diagenetic simulations for the derivation of the metamodel. Figure 1, where we compare sediment-water fluxes derived from the metamodel with those predicted by a time-varying diagenetic simulation, illustrates that the new approach to deriving the metamodel can capture time-dependencies in a reasonable way. The new metamodel tracks the fluxes from the diagenetic model fairly well and is able to reproduce time-lagged responses to changes in bottom water conditions and PON deposition. Figure 2 compares the old and the new metamodels with the diagenetic model results. In the new parameterization the maxima in O2 uptake, NH4 efflux and NO3 efflux all occur after the main deposition event in May, and the dependency on depositional flux is much weaker than in the old metamodel (see also the new Figure 3 below). The new metamodel is overall in much better agreement with the time-dependent diagenetic simulation.

In deriving the old parameterization we randomly sampled (in time and space) 10^5 sets of bottom water conditions from the output of a coupled circulation-biological model simulation and used these as forcing without any time dependence. In our new approach for deriving the metamodel we use the time series at each grid cell on the Louisiana Shelf from the same model simulation (3791 grid cells, 1824 days each) to force the diagenetic model. The parameterization is then calculated using the resulting $\sim 7\times 10^6$ data vectors (consisting of daily bottom water concentrations and PON deposition, as well as simulated sediment-water fluxes) from the diagenetic model simulations. Figure 1 illustrates how well the metamodel reproduces the time dependent simulation from the diagenetic model for one grid cell. We calculate the goodness-of-fit and the correlation coefficient between the diagenetic model and the metamodel parameterization for all grid cells on the Louisiana Shelf. The spatially-resolved correlation coefficients (Figure 4) are above 0.8 for most of the Louisiana Shelf for O_2 and NH_4 fluxes and above 0.6 for NO_3 fluxes. The parameterization fails in some limited regions near the offshore limit of the shelf. We will discuss this in detail in the revised manuscript. We will also include a more detailed discussion on the limitations of the metamodel, as suggested by the reviewer.

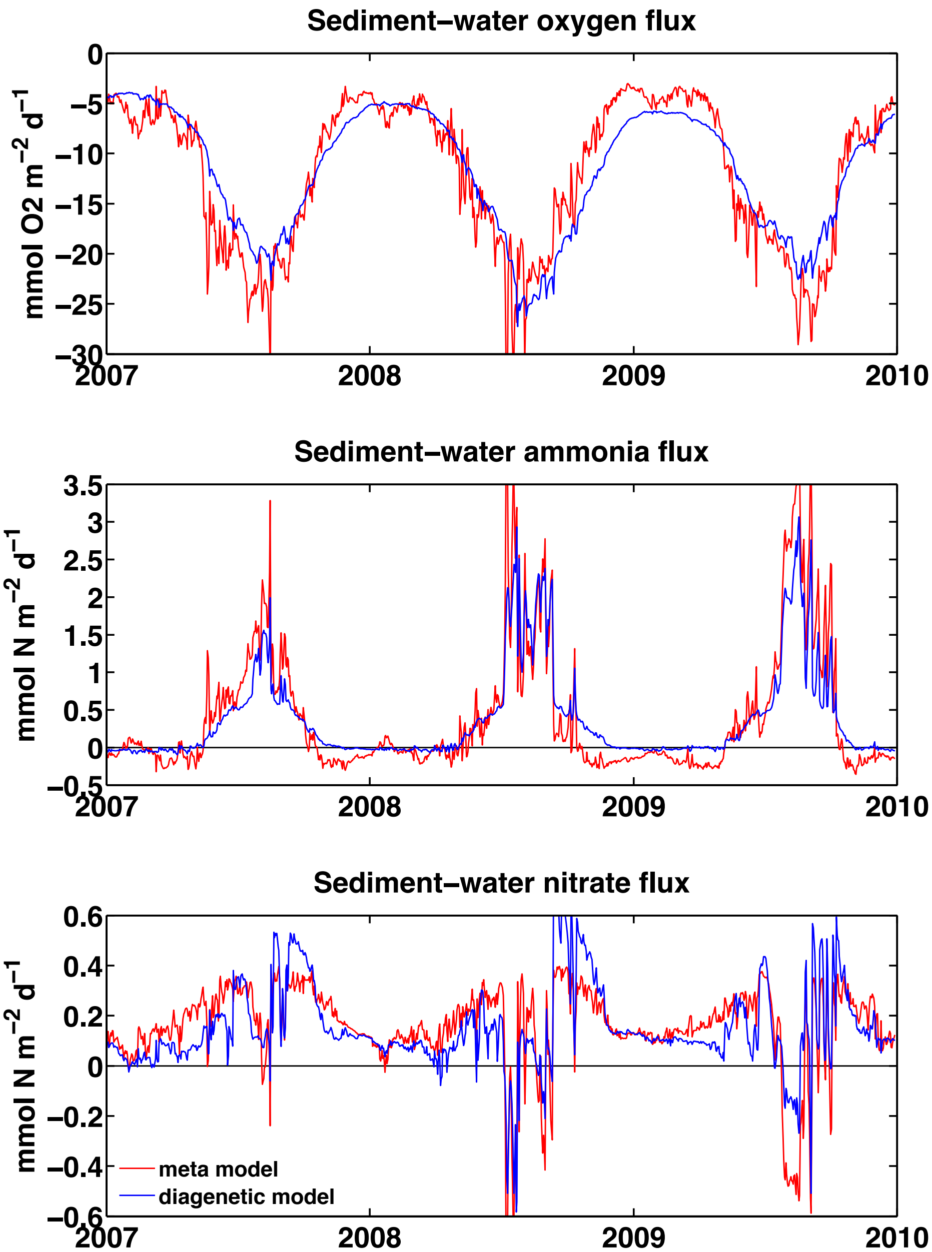

Figure 1. Time series of sediment water fluxes from a time-dependent simulation of the diagenetic model and from the new metamodel at a mid-shelf location (91.5°W, z=20 m). The time series are used to calculate goodness-of-fit and correlation coefficients at this location, as shown in Figure 4 below. In this case correlation coefficients for O_2, NH_4 and NO_3 fluxes are 0.90, 0.93 and 0.76, respectively.

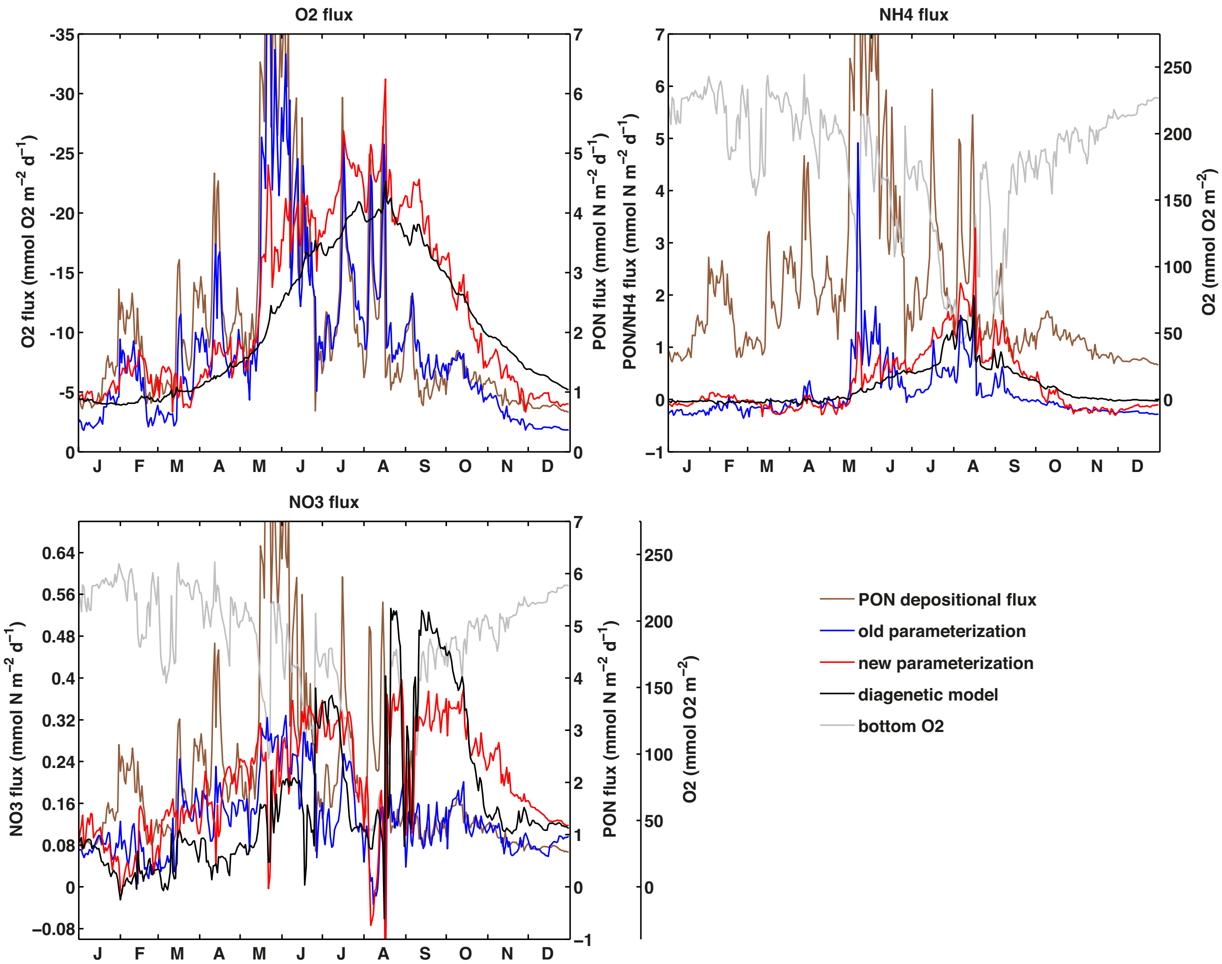

Figure 2. Comparison of sediment-water fluxes from the old metamodel (blue), the new metamodel (red) and the diagenetic model (black) at a mid-shelf location (91.5°W, z=20 m).

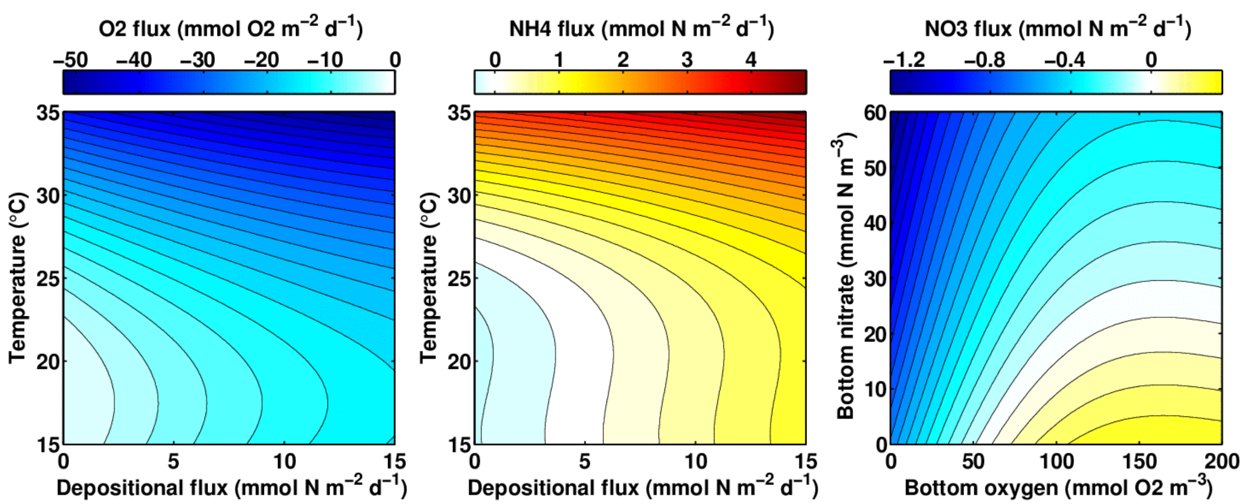

Figure 3. Updated metamodel as a function of the three dominant drivers.

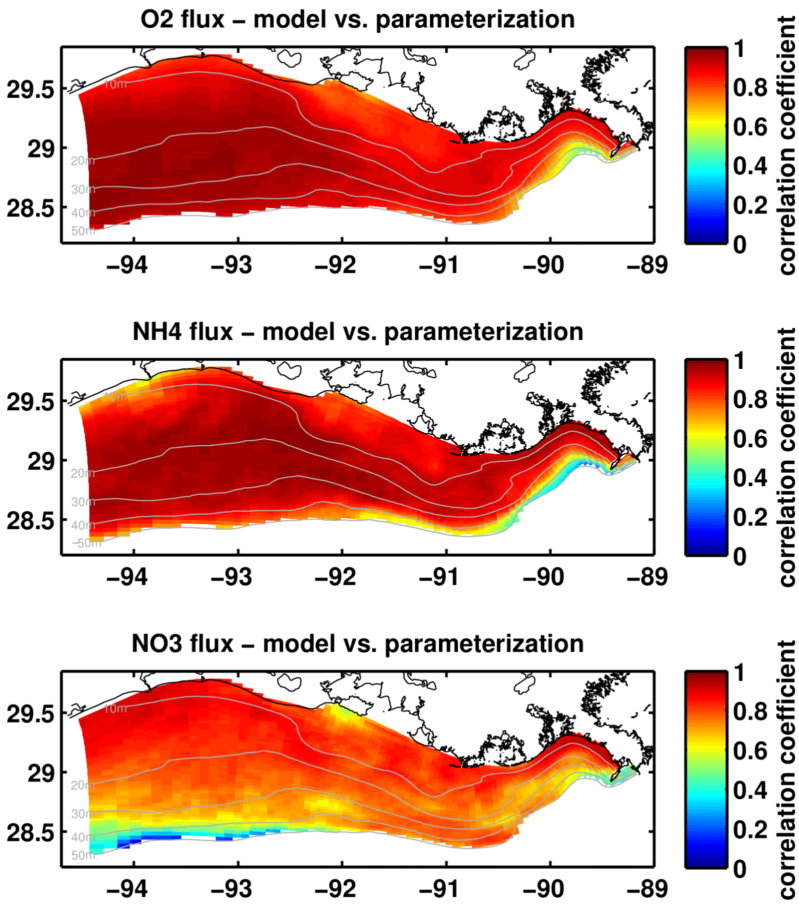

Figure 4. Correlation coefficients between time-dependent diagenetic model simulations and the parameterized fluxes for each location on the Louisiana Shelf.

Comment:
3. On similar level, on p7553 line 16+ the authors claim that using a time-varying forcing for the optimization would not have changed the results significantly given the constraint of the dataset on the optimization. This is a firm statement based on belief rather than fact. I do not agree with the authors, given the previous comments, and would prefer it if they would be more open about the possibility that a transient optimization routine, combined with a dynamic sediment component, would improve their model.

Response:
We agree and will remove this statement. Unfortunately, given the limited data that we have available, we can't run the optimization in a time-dependent mode. This was not stated in the previous version of the manuscript. In the revised manuscript we will better explain our choice of using steady state conditions for the optimization and its implications for the metamodel.

Comment:
4. The POM fluxes (actually PON fluxes) in Table 1 used in the diagenetic model are taken from the pelagic model (Fig. S1). However, the tabulated fluxes in June and September for St. Z02 are up to a factor of 5 lower than in the pelagic model. Could this explain why the meta-model predictions are improved in June when the POM flux is used as an additional parameter to optimize (p7549,L10, Fig. 3)? In fact, there are other indications that the POM flux is too low. The optimized model underestimates NH_4 and NO_3 fluxes in June at Z02 (Fig. 2). Also, it requires minimal infaunal mixing, no burial of reduced compounds, and an increased T dependence on mineralization (p7548,L15+). This suggests to me that the optimized model is trying to mineralize as much POM as possible in order to fit the NH_4 concentrations and fluxes.

Response:
The monthly PON depositional fluxes in Table 1 used to force the steady state diagenetic model are a climatology. They represent average conditions and therefore underestimate the short-lived, high PON deposition events that occur at station Z02, as shown in Figure S1. Given the lack of PON depositional flux data, and therefore the high uncertainty on PON deposition, we chose to use climatological fluxes to force the steady state diagenetic model. This is a limitation to our optimization that we will discuss more extensively in the revised manuscript.

Comment:
5. The optimization procedure works by firstly selecting n sets of parameters whose values are determined by adding random noise to the original parameterization of Soetaert et al. (1996). The next n sets of parameters in the following generation are based upon the best half (n/2) parameter sets from the previous generation using the cost function, and so on, for 200 generations. The evolutionary trajectory leading to the final 'fittest' optimized set may depend more on the best fit parameterization in the early generations rather than the later ones. In other words, small differences in the parameterization at the start could lead to large differences at the end, resulting in a different set of optimized parameters that explain the observations equally well, analogous to following a tree trunk upwards and coming out at a different twig every time. Have the authors run through the whole procedure repeatedly to double check that the same final parameter set is predicted each time?

Response:

The evolutionary algorithm explores the parameter space by adding random noise and is therefore less prone to finding local minima than the widely used gradient-descend algorithms (see also response to comment 3 by reviewer 2). Nonetheless the evolutionary algorithm technique may be sensitive to the initial set of parameter values (Schartau and Oschlies, 2003). We repeated the optimization procedure with different initial parameter sets and found that the end result was not sensitive to the initial parameter choice.

Comment:
6. POC fluxes in this system are high (>100 mmol/m2/d in summer) at station Z02, in combination with severe hypoxia (Fig. S2). I would expect high sediment porewater sulfide concentrations under these conditions and the development of sulphur oxidizing bacteria communities on the sediment surface that carry out DNRA ($NO_3 + H_2S \rightarrow NH_4 + SO_4$). Can the authors justify why this process was omitted in their approach? It is not enough to reply that this process was not included in the original Soetaert model (which, incidentally, is a deep-sea application). Perhaps DNRA explains why the model does not simulate the high NH_4 porewater concentrations in September at Z02 when DNRA rates would be expected to be highest. I would speculate that during the summer period, DNRA would become an important contributor to the N cycle, as observed in seasonally hypoxic settings elsewhere (Dale et al., 2013, Biogeosciences, 10, 629-651, doi:10.5194/bg-10-629-2013). Enhanced NO_3 uptake by bacteria may also explain why the diagenetic and meta-models are unable to simulate high NO_3 fluxes (Fig. 2 and Fig. 5c).

Response:
We would first like to note that hypoxia at station Z02 is not severe for extended periods of time. Fig. S1 shows that hypoxia is mild in July and severe only for a brief period in August at station Z02. Figure S2, which was cited by the reviewer, does not show oxygen concentrations. High porewater sulfide concentrations near the sediment-water interface are not reported for sediments on the Louisiana Shelf (Morse and Eldridge, 2007; Morse and Lin, 1991).

We will add a paragraph in the revised manuscript to discuss the limitations of the diagenetic model and the potential effects of the omission of DNRA on our results. Overall, there is a relatively poor understanding of the importance of DNRA on the Louisiana shelf due to a lack of observations in the region (Dagg et al., 2007). Nunnally et al. (2013) suggested the occurrence of DNRA given observed nitrate depletion in bottom water samples. In a recent study, McCarthy et al. (2015) didn't find DNRA to be a consistent N pathway on the Louisiana Shelf; however, they recommend further investigation of this process. Given the lack of consensus on DNRA in our region we didn't include this process in the diagenetic model recognizing that this is a potential shortcoming of the diagenetic model. We will discuss this in the revised manuscript. See also our response to comment 5 by reviewer 2.

Comment:
7. Similarly, given the severe depletion of O_2 in late summer, one could expect infaunal mixing by bioturbation and bioirrigation to be dependent on O_2, in line with other observations and models (Dale et al., 2013). Please comment.

Response:
The optimization tends to minimize the influence of bioturbation in the diagenetic model (depth of bioturbated layer, non-local mixing). This likely reflects the effect of hypoxia on the sediment biota and is in line with observations from the Louisiana Shelf, which show that bacteria tend to

dominate the sediment community. Likely bioturbating macrobiota does not re-establish itself in the regions affected by recurring seasonal hypoxia, thus we don't expect a strong dependence of bioturbation and bioirrigation in this system. We discuss this in the Discussion section, and will comment on this assumption in the description of the diagenetic model (section 2.2.1) in the revised manuscript.

Comment:
8. The rate of organic matter mineralization is temperature dependent, but other microbially mediated reactions are not. Please explain.

Response:
Temperature influences the solute diffusivity and the degradability of reactive and refractory organic matter in the diagenetic model. Bioturbation diffusivity is also temperature-dependent. We will clarify the role of temperature in the revised description of the diagenetic model (section 2.2.1).

Comment:
9. The meta model predicts a high O2 flux at zero O2 bottom water concentrations (Fig. 9a). This is strange and must be clarified.

Response:
O2 flux at zero O2 bottom waters represents the production of ODUs under anaerobic conditions; we assume that ODUs are oxidize instantaneously in the water column and therefore the total O2 flux represents an O2 consumption by the sediment. We will clarify this point in the revised manuscript.

Comment:
10. Can the authors explain why the NO3 flux does not depend at all on POM flux (Table 4). After all, no POM = no diagenesis.

Response:
The dependence of NO3 flux on POM deposition has changed in the updated metamodel, which is now able to capture time-lagged effects on NO3 fluxes. NO3 flux is strongly dependent on the presence of O2 in bottom waters, as shown in Figures 2 and 3.

Comment:
11. General: font size on the figures is really, really small. Please correct this.

Response:
The figures will be updated with larger font sizes.

Comment:
12. Table 1 and section 2.1. Please provide the water depth of the two stations.

Response:
In section 2.1 we mention that the two stations are along the 20m isobath, as shown in Figure 1. We feel that including the water depths in Table 1 would be redundant.

Comment:

13. Table 1 header. I believe that the fluxes and NH4 profiles are used to optimize the diagenetic model via Eq. 3, not the boundary conditions listed in this table.

Response:
The boundary conditions listed in Table 1 are used as boundary conditions for the diagenetic model. Fluxes and NH4 profiles are used to calculate the cost in the optimization procedure. We will clarify the header in Table 1 as follows: "*These data are used as boundary conditions during the optimization of the diagenetic model*".

Comment:
14. Table 4 header: Unclear 'the direction of its effect'.

Response:
We will clarify that point by adding the following sentence to Table 4 header: "*A positive effect promotes a release to the water column whereas a negative effect leads to a sink into the sediment.*"

Comment:
15. P7538,L13 and P7540,L40. ...O2, NO3 and NH4 fluxes. P7539,L19 nitrate/nitrite.

Response:
The text will be modified accordingly.

Comment:
16. P7539,L28. Please add reference Bohlen et al. 2012 after Fennel et al., 2009. (Bohlen, L., Dale, A. W., Wallmann, K. (2012) Simple transfer functions for calculating benthic fixed nitrogen losses and C:N:P regeneration ratios in global biogeochemical models. Global Biogeochemical Cycles 26, GB3029, doi:10.1029/2011GB004198).

Response:
The text and references will be modified accordingly.

Comment:
17. P7540,L9. Vertically integrated and depth resolved models are not the same thing. The context of this paragraph makes me believe that the authors are referring to the latter type only. Vertically integrated (to my mind) would be a sediment-transfer function or a single layer model (see Soetaert et al., 2000).

Response:
For clarity we will modify the sentence to "*Vertically integrated and depth-resolved mechanistic models of diagenesis…*" Vertically integrated model refers to a single layer model of diagenesis.

Comment:
18. I would personally like to see, for the sake of correctness, charges assigned to the anions, (e.g. NH4+ instead of NH4). But that's just my own preference.

Response:
The text will be modified accordingly.

Comment:
19. Section 2.1. Please provide some more information on how the fluxes were determined (e.g. ex situ versus in situ, no. replicates etc). The authors report a standard deviation on the measured data, which only makes sense if a reasonable large number of observations were made.

Response:
We will add more methodological information on sediment-water flux observations in Section 2.1.

Comment:
22. P7541,L17: Suggest change 'observations' to 'data'. P7541,L19: Suggest delete 'process leg'. P7541,L21+24. Suggest delete or clarify 'near shelf survey stations X'. This may mean something to the authors, but will mean nothing to most readers.

Response:
The text will be modified accordingly. We will also replace "*near shelf survey stations X*" by "*see Murrell et al., 2013, for details on sampling design*".

Comment:
23. P7543,L21-22. Please briefly clarify 'Given the lack of observations on the labile and refractory fraction of OC'. Does this mean the rates constants? Please briefly explain how Wilson et al. constrained these values. And anyway, why are constraints needed if these parameters are optimized?

Response:
The diagenetic model divides OC into labile and refractory pools. There are no direct observations available for our sites. We assumed that the fraction of deposited OC that is labile is constant in our experiments. Since deposited OC mainly originates from local primary production, we assumed that labile OC makes up 74% (OC_{frac2} = 0.74, see Table 2), as in Wilson et al (2013). The fraction of labile and refractory material in deposited OC is part of the model forcing and we chose to not make it part of the optimized parameter set. In the revised manuscript we will exclude OC_{frac2} from the parameter list in Table 2 and instead describe it with the diagenetic model (section 2.2.1). We will also clarify our choice and the underlying assumptions.

Comment:
24. Are the OM degradation rate constants listed in Table 2? I see only R1opt and R2opt, which have units of 1/time but are described as 'rates', which has a unit of concentration/time. Please clarify this, both in the table and next to Eq. 1. Whilst on this subject, it would help the reader if units were included next to all rates/parameters in the model description (section 2.2.1).

Response:
In Table 2 and in the text we now refer to "*remineralization*" rather than "*remineralization rate*" since R_1^{Topt} and R_2^{Topt} have units of 1/time. We also provide units next to the parameters in section 2.2.1, as suggested by the reviewer.

Comment:

25. P7545,L14. How did the cost function (Eq. 3) account for the NH4 profile? Was every data point (Xmodel – Xobs) considered, or some integration of all the points together?

Response:
In order to avoid a biasing toward profiles in the cost calculation we compute an average cost per profile. This information will be added to the revised manuscript.

Comment:
26. P7545, L17-18. I don't follow the weighting approach, please clarify. Why was the initial parameter set used?

Response:
Without weighting, the contribution to the total cost from O_2, NH_4 and NO_3 sediment-water fluxes and NH_4 profiles would be quite different despite the presence of the standard deviation in the denominator of the cost function terms. This is common when different data types are combined into one total cost and reflects the fact that models have an easier time fitting some data types than others. We chose the weights in order to ensure that all data types contribute about equally to the cost function initially. The initial parameter set is used to estimate the weights. We will clarify the weighting approach in the revised manuscript.

Comment:
27. P7545,L21-27. The authors summarize here the sensitivity analysis, but all too briefly. There are several steps mashed together in only one sentence. Please take care to explain these steps in more detail so that others can follow the logic.

Response:
We will extend and clarify the description of the sensitivity analysis in the revised manuscript describing each step.

Comment:
28. P7546,L14. Please clarify that O2, NO3 and NH4 refer to bottom water concentrations. Please also provide the range of values used from the pelagic model in the met-model procedure. In Fig. S1, only POM flux and O2 concentrations are shown, but presumably NO3 and NH4 concentrations were also simulated, so please show them.

Response:
The sentence will be modified by including "*namely combinations of bottom water temperature,* ...". In the description of the meta-modelling procedure (section 2.2.3) we will provide the range of values for each variable used in the metamodel. NH_4 and NO_3 will also be added to the revised supplementary Figure S1.

Comment:
29. P7546,L15. Suggest delete 'for each flux variable'. P7546,L19. ...to an explanatory variable i, and... P7547,L8. Should 'there' be 'three'?

Response:
The text will be modified accordingly.

Comment:
30. P7547, section 2.3. The authors should show mathematically these other different approaches, otherwise the reader has no means to judge the current model and interpret Fig. 9 and 10 (without going back to the original sources).

Response:
The formulations of each parameterization will be added to section 2.3.

Comment:
31. P7545 P7548,L20. Please explain in a bit more detail (in the model description) what permanent burial of ODU refers to.

Response:
The model description will be modified accordingly.

Comment:
32. P7550,L25. I take it that bottom water NO3 and NH4 concentrations are also available from the pelagic model to drive the meta-models? Please show them.

Response:
Bottom water NO_3 and NH_4 at station Z02 and Z03 will now be included in Figure S1.

Comment:
33. P7550,L28. 'LUMCON' means nothing to most readers. Please write out the acronym (if it is one) and add a reference if possible.

Response:
The text will be modified accordingly and a reference included.

Comment:
34. Finally, given the importance of the pelagic model results to this study, i suggest shifting Fig S1 (bottom) into the main text along with NO3 and NH4 concentrations which must also be available.

Response:
We will move Figure S1 to the main text.

References:

Dagg, M., Ammerman, J., Amon, R., Gardner, W., Green, R. and Lohrenz, S.: A Review of Water Column Processes Influencing Hypoxia in the Northern Gulf of Mexico, Estuaries Coasts, 30(5), 735–752, 2007.

McCarthy, M. J., Newell, S. E., Carini, S. a. and Gardner, W. S.: Denitrification Dominates Sediment Nitrogen Removal and Is Enhanced by Bottom-Water Hypoxia in the Northern Gulf of Mexico, Estuaries and Coasts, doi:10.1007/s12237-015-9964-0, 2015.

Morse, J. W. and Eldridge, P. M.: A non-steady state diagenetic model for changes in sediment biogeochemistry in response to seasonally hypoxic/anoxic conditions in the “dead zone” of the Louisiana shelf, Mar. Chem., 106(1-2), 239–255, doi:10.1016/j.marchem.2006.02.003, 2007.

Morse, J. W. and Lin, S.: Sulfate reduction and iron sulfide mineral formation in Gulf of Mexico anoxic sediments, Am. J. Sci., 291, 55–89, 1991.

Nunnally, C. C., Rowe, G. T., Thornton, D. C. O. and Quigg, A.: Sedimentary Oxygen Consumption and Nutrient Regeneration in the Northern Gulf of Mexico Hypoxic Zone, J. Coast. Res., 63, 84–96, doi:10.2112/SI63-008.1, 2013.

Schartau, M. and Oschlies, A.: Simultaneous data-based optimization of a 1D-ecosystem model at three locations in the North Atlantic: Part I—Method and parameter estimates, J. Mar. Res., 61(6), 765–793, doi:10.1357/002224003322981147, 2003.