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

General comments

As the method is of use for a broad audience, and it is suitable for the journal, I would recommend publication, albeit after major revision. One of the flaws of the current manuscript is that the reasons and consequences of the deviations of the approach from S2000 are not at all discussed.

Response: We appreciate the careful and critical comments. We have made an effort to improve upon the assumption of steady state, which was raised as a shortcoming by both reviewers. We will also better discuss limitations of our approach in the revised manuscript.

Specific comments

Comment:
1. 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:
In response to this concern (also Comment 1 by Reviewer 1, Andy Dale), we modified the diagenetic model used in the metamodel derivation. Instead of steady state we now use dynamic forcing for the diagenetic model. Figure 2 (given in our response to Reviewer 1) shows how our original metamodel (based on steady state forcing) fails to capture the temporal evolution of sediment-water fluxes from a time-varying diagenetic simulation (consistent with the concern raised by both reviewers). Figure 2 also shows how our modified procedure using time-dependent simulations in the derivation of the meta-model results in a parameterization that is much better able to reproduce the temporal evolution of the fluxes. The updated diagenetic model accounts for memory of the sediment, although not in a mechanistic way. The decoupling between PON deposition and O_2 flux in the revised metamodel is obvious in Figure 2. The updated metamodel remains simpler than the procedure recommended in S2000. We will discuss the implications and potential limitations of our choice in the revised manuscript.

Comment:

2. The ‘novelty’ of the method is that the metamodel directly fits the oxygen, nitrate and ammonium fluxes as a function of the water-column conditions. In contrast, in the S2000 paper, the fraction nitrified, denitrified and anoxic mineralized is fitted instead, and the fluxes derived from the mineralization rates. The reason for this choice was that this is mass conservative. Deviating from this as in the current manuscript, it is well possible that, due to statistical fitting, for instance more nitrogen is returned than is originally deposited in the sediment – which means that mass is numerically created. This may not be the case, but at least the authors should spend a paragraph as to why they decided to do the meta-modeling differently and whether or not mass is (numerically) created by doing this.

Response:

We agree that our parameterization is not mass conservative. An accounting of organic matter in the sediment would be necessary to implement the mass-conserving option recommended in S2000. In this case we chose the simplest approach to provide more realistic sediment-water fluxes than have been used in the past in regional biogeochemical models for the Louisiana Shelf. As long as the metamodel is used for the system for which it was developed and within the range of conditions that were used for the parameterization, violation of mass conservation should be minor. We agree that these limitations and caveats have to be clearly stated. We will provide more discussion of the assumptions, the mass balance issue and the limitations of our method in the revised manuscript.

Comment:

3. I have strong doubts on the statistical validity of the fitting procedure. In my experience, it is not possible to find robust estimates of the 20 parameters that were selected, given the small amount of data available. To be able to fit all these parameters, it is necessary that they are ‘identifiable’, i.e. the value of one parameter does not depend to a large extent on the values of (a set of) the other parameters. Parameters that are not identifiable by the data, have very large uncertainty. Unfortunately, the genetic algorithm does not return a measure of parameter uncertainty. A very old paper of mine deals with the identifiability of a diagenetic model (http://dx.doi.org/10.1357/002224098321822401); a more recent paper presents software to do more robust parameter fitting that includes a.o. estimating parameter identifiability, parameter sensitivity, and to evaluate the uncertainty of the derived parameters. (www.jstatsoft.org/v33/i03/paper). I realize that it would be too much of an effort to require all this now, but at least the lack of statistical rigor should be mentioned in the manuscript and the results discussed in section P12-L13 and on P18-L25 should be phrased less strongly.

Response:

We will expand the presentation of the fitting procedure and mention the identifiability issue (including the two references provided by the reviewer). However, we note that the issue of poor identifiability of certain parameters would not be alleviated by using a different optimization approach (assuming that different approaches find the same optimum). The only real solution to the issue would be to have a more complete set of observations to optimize against. We will discuss this in the revised manuscript. The genetic algorithm is a well-accepted method for optimization problems in the statistical literature (Chatterjee et al., 1996; Fogel, 1994; Hibbert, 1993; Kolda et al., 2003). This technique has been increasingly used to optimize parameters in biogeochemical models of the water column (Schartau and Oschlies, 2003; Robson et al., 2008;

Ward et al., 2010; Kuhn et al., 2015) and the sediment (Wilson et al., 2013; Wood et al., 2013). An advantage of the evolutionary algorithm over the traditionally more common gradient-descend algorithms is that it explores the parameters space with an element of randomness and therefore is less prone to converging on a local minimum. In addition to the optimization, we carried out a sensitivity analysis of the parameter set to explore the effect of parameter change ("uncertainty") on the model results. We will better discuss this in the revised manuscript.

Comment:
4. From table 2, it is clear that the modeled sediment depth is only 10 cm. Given the boundary conditions that are imposed on the diagenetic model, this entails that gradients vanish at that depth (which is seen on the modeled ammonium profiles in Fig. 2). However, the observations, especially in April and one profile in September still show a large gradient at 8 cm; this makes me suspect that restricting the model to the upper 10 cm of sediment is not adequate for these data.

Response:
We acknowledge that restricting the diagenetic model to the upper 10 cm is a limitation and will mention this in the revised manuscript.

Comment:
5. The OMEXDIA model would need a few other additions to make it better suitable for these high-flux sediments (e.g. a reaction of ODU and NO3).

Response:
While further modification of the diagenetic model is beyond the scope of the present manuscript, we would be most appreciative of further guidance on how the model can be improved for the region. The main focus of this manuscript is to present a method for parameterizing sediment-water fluxes and apply it to the Louisiana Shelf.

Comment:
6. Salinity is a parameter inputted to the meta-model, but it is unclear how it influences the diagenetic model.

Response:
Salinity is not used in the diagenetic model so the values in Table 1 are only informative of bottom water conditions. We removed salinity from Table 1 to avoid confusion.

Comment:
7. Fig. 2. The original OMEXDIA model would never be able to generate the steady-state ammonium profiles from April. As ammonium peaks at 3 cm, it means that there is a sink of ammonium below that depth, although there is no oxygen. What causes this decline?

Response:
We reviewed the optimized steady state model results from April and ruled out the presence of a deep ammonium sink at steady state. However we noticed that the optimized simulations for April didn't reach a full steady state at depth, which resulted in the negative gradient in the deep layers. We will correct this issue in the revised manuscript.

Comment:
8. Fig. 9. The oxygen flux is negative everywhere except in this figure. Should be made consistent. Equation (1) Topt is not a good name for this parameter; the term 'optimal' suggest that the function peaks at the temperature, which it does not. Topt is the "base" temperature, i.e. the temperature for which the rate is defined. Usually the base temperature is taken as 0 degrees or 20 degrees. It would be easier to compare the derived rates with those from other models if using a more standard base temperature (rather than 30 degrees).

Response:
For consistency we will make the O_2 flux negative in Figure 9. We will change T_{opt} to T_b for base temperature throughout the revised manuscript.

Comment:
9. Equation (2). Give the units of I(z).

Response:
The units of I(z), µmol L^{-1} y^{-1}, will be added to the revised manuscript. In the revised manuscript we will provide the units next to each parameter in the text (see comment 24 by reviewer 1)

Comment:
10. Equation (3):

*why is it sigma (s,i) and not sigma(s,t,i).

*I do not understand the meaning of 1/wi – this weighing is not standard and seemingly dependent on an –arbitrary- initial parameter set? Due to the division by the standard deviation, the fact that units are different is already taken into account, so I so not see the need to have an extra weighing term.

*How does the variable cost in Table 3 relate to formula (3)-does it include the 1/wi term?

Response:
We corrected to $\sigma^2_{s,t,i}$.

We will also clarify the weighting approach in the revised manuscript (see also response to comment 26 by Reviewer 1). A weight was included in the cost function to prevent that some variables have more influence on the overall cost than others, at least initially. The weighting is common in parameter optimization studies (see, e.g., Friedrichs, 2001; Friedrichs et al., 2007; Kane et al., 2011; Schartau and Oschlies, 2003). By using the initial parameter set in estimating the weights prior to optimization we ensure that all data types initially contribute equally to the optimization.

The values in Table 3 correspond to the cost function *F* given in Equation 3 (including the weights).

Comment:
11. Table 2.

*Units of aNH3, aO2, etc.. and the unit of kinanox are wrong.

*Unit of Dbcoeff and PB is lacking; also the NC ratio has a unit: is it gram/gram or mol/mol?

*A lot of the optimized parameter values are rather round numbers, which make me assume that they are located somewhere near the edge of the allowed range? What was the range imposed?

Response:
In the revised Table 2 we modified the units of a_{XX} ($cm^2\ d^{-1}\ (°C)^{-1}$), kin_{anox} (µmol NO3 L^{-1}), Db_{coeff} (cm), PB (d^{-1}) and N:C (mol N (mol C)$^{-1}$).

Some of the optimized parameters reach the lower or upper edge of the allowed range. This will be stated explicitly in the revised manuscript We will also add a column to Table 2 indicating the range of values allowed for each parameter during the optimization to make clear for which parameters this happens.

Comment:
12. Table 3. Why are FNO3 and FNH4 so high in comparison with the other data?

Response:
NH_4 and NO_3 fluxes represent a more difficult problem for the optimization and therefore their cost is larger. Note that, as per our response to comment 10, all fluxes contribute about equally at the outset of the optimization.

Comment:
13. Page4-Line7: 'a proportion of the deposited nitrogen is lost as N2 gas'. If the denitrification would work like this, it would mean that the sediments would never be a sink of nitrogen, which is in contradiction with the previous sentence (P4-L6). 'A proportion of organic matter is mineralized with nitrate' is a correct description of denitrification.

Response:
We will modify the sentence "*a proportion of the deposited nitrogen is lost as biologically unavailable N gas (N_2)*" to "*a proportion of the deposited organic matter is remineralized via denitrification which produces biologically unavailable N_2 gas*".

Comment:
14. P14-L14: what is RMSE ? How does this relate to the cost (equation 3)?

Response:
We use the root mean square error (rather than the cost) to show that even though the total cost improves, the agreement between observed and modeled NH_4^+ profiles gets worse when the profiles are not included in the cost calculation. To clarify this point we will modify the paragraph as follows: "*However, when NH_4^+ profiles are not included in the optimization there is a large deviation between observed and modeled sediment NH_4^+ concentrations (not included in the cost). The root mean square error for the sediment profiles increases from 87.59 mmol N m^{-2} d^{-1} for the baseline case to 174.45 mmol N m^{-2} d^{-1} (Site-specific, flux only) and 111.86 mmol N m^{-2} d^{-1} (Site-specific, flux only + F_{POM}).*"

Comment:

15. P20L22 What does the metamodel do when there is an O2-debt ? I assume this means that negative oxygen concentrations arise in the pelagic model – were negative oxygen concentrations used for fitting the metamodel?

Is the consumption of oxygen in the absence of a PON flux caused by the nitrification of ammonium that fluxes into the sediment or is it a statistical artifact (it appears to be quite high)?

Response:
For anoxic bottom waters, the O_2-debt is represented by negative O_2 concentrations in the biogeochemical models cited in L23. We did not use negative O_2 concentrations to fit the meta-model. These waters are anoxic and we used $O_2 = 0$ in these cases.

The occurrence of nitrification in the sediment explains the O_2 consumption in the absence of PON flux. This is clearer in the revised Figure 6 (see Figure 2), which shows that O_2 and NH_4 uptake by the sediment occur at low PON flux in the meta models.

References:

Chatterjee, S., Laudato, M. and Lynch, L. A.: Genetic algorithms and their statistical applications: an introduction, Comput. Stat. Data Anal., 22(6), 633–651, doi:10.1016/0167-9473(96)00011-4, 1996.

Fogel, D. B.: An introduction to simulated evolutionary optimization., IEEE Trans. neural networks, 5(1), 3–14, doi:10.1109/72.265956, 1994.

Friedrichs, M. A. M.: A data assimilative marine ecosystem model of the central equatorial Pacific: Numerical twin experiments, J. Mar. Res., 59(6), 859–894, doi:10.1357/00222400160497544, 2001.

Friedrichs, M. A. M., Dusenberry, J. A., Anderson, L. A., Armstrong, R. A., Chai, F., Christian, J. R., Doney, S. C., Dunne, J., Fujii, M., Hood, R., McGillicuddy, D. J., Moore, J. K., Schartau, M., Spitz, Y. H. and Wiggert, J. D.: Assessment of skill and portability in regional marine biogeochemical models: Role of multiple planktonic groups, J. Geophys. Res., 112(C8), C08001, doi:10.1029/2006JC003852, 2007.

Hibbert, D. B.: Genetic algorithms in chemistry, Chemom. Intell. Lab. Syst., 19(3), 277–293, doi:10.1016/0169-7439(93)80028-G, 1993.

Kane, A., Moulin, C., Thiria, S., Bopp, L., Berrada, M., Tagliabue, A., Crépon, M., Aumont, O. and Badran, F.: Improving the parameters of a global ocean biogeochemical model via variational assimilation of in situ data at five time series stations, J. Geophys. Res., 116(C6), C06011, doi:10.1029/2009JC006005, 2011.

Kolda, T. G., Lewis, R. M. and Torczon, V.: Optimization by direct search: New perspectives on some classical and modern methods, SIAM Rev., 45(3), 385–482, doi:Doi 10.1137/S003614450242889, 2003.

Kuhn, A. M., Fennel, K. and Mattern, J. P.: Model investigations of the North Atlantic spring bloom initiation, Prog. Oceanogr., doi:10.1016/j.pocean.2015.07.004, 2015.

Robson, B., Hamilton, D., Webster, I. and Chan, T.: Ten steps applied to development and evaluation of process-based biogeochemical models of estuaries, Environ. Model. Softw., 23(4), 369–384, doi:10.1016/j.envsoft.2007.05.019, 2008.

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.

Ward, B. a., Friedrichs, M. a. M., Anderson, T. R. and Oschlies, A.: Parameter optimisation techniques and the problem of underdetermination in marine biogeochemical models, J. Mar. Syst., 81(1-2), 34–43, doi:10.1016/j.jmarsys.2009.12.005, 2010.

Wilson, R. F., Fennel, K. and Paul Mattern, J.: Simulating sediment–water exchange of nutrients and oxygen: A comparative assessment of models against mesocosm observations, Cont. Shelf Res., 63, 69–84, doi:10.1016/j.csr.2013.05.003, 2013.

Wood, C. C., Statham, P. J., Kelly-Gerreyn, B. a. and Martin, A. P.: Modelling macronutrients in shelf sea sediments: fitting model output to experimental data using a genetic algorithm, J. Soils Sediments, 14(1), 218–229, doi:10.1007/s11368-013-0793-0, 2013.