Biogeosciences Discussion

Dr Felipe Sales de Freitas
School of Earth Sciences
University of Bristol
Wills Memorial Building, Queen’s Road
BS8 1RJ Bristol
United Kingdom

Email: felipe.salesdefreitas@bristol.ac.uk

Date: March 15th, 2021

Subject: Response to Reviewer’s comments on BG-2020-435

Manuscript ID: BG-2020-435 “Advancing on large-scale trends on apparent organic matter reactivity in marine sediments and patterns of benthic carbon transformation”

Dear Reviewer 2,

On behalf of the co-authors, I want to thank you for the detailed and constructive review of our manuscript. In the following, we reply to each of the issues raised and explain how these will be addressed in the revised manuscript.

Sincerely yours

Felipe Sales de Freitas

Felipe Sales de Freitas and co-authors

General comments

“the text is not always easy to read. This is especially the case for its first paragraph, which is close to unreadable. Readers have to search for actual sentence parts among a plethora of citations—there are nearly sixty of them in thirty lines of text (are these really all necessary?). The rest of the paper is more readable but in general so wordy that one often loses the focus on what is actually important.”

We will shorten the manuscript to help keep the reader’s focus on what is important, but without losing the essence of our work and findings.

We assume that Reviewer 2’s comment on citations refers to line 65-75. This section discusses the interplay of different environmental controls on organic matter (OM) reactivity, and we prefer to refer to the original literature and topic specific review papers instead of merely citing a couple of more general review papers. However, to increase the readability, we will create a table summarizing the main environmental factors controlling OM reactivity in marine sediments and specific references to the original work that explores their respective influence on OM reactivity.

“In fact, OM reactivity is determined by the dynamic interplay of a plethora of different factors (see Table 1 and Arndt et al., 2013; LaRowe et al., 2020 and references therein).”

Despite its central role in the study, the inversion method is described in a completely inadequate way. In its current form, that description does not allow to reproduce the results as critical details on the adopted algorithm are missing.

We will extend the description of the inverse modelling approach in section 3.2.3. See detailed answer to specific comments below.

The notation in the paper is not entirely consistent (the same symbols have different meanings or different symbols used for the same types of variables, or symbols that even change for one and the same variable) and at times unconventional which makes reading again unnecessarily difficult.

We apologise for inconsistencies in notation and will carefully revise the manuscript accordingly. To keep consistency and guide readers, we will include a summary table of used symbols and their meaning. See detailed answer to specific comments below.

The conclusions are overdone, both in content and tone. They would benefit from a fair share nuance and precise contextualization (they are for an RC model with a particular choice of reactivity distribution, which is not necessarily the only possible choice, even if it is the most commonly adopted one) would do them good. It would be desirable to let the results speak for themselves so that there is no need to emphasize how important they are.

See detailed answer to specific comments below.

Finally, there are problems of more fundamental nature: there are no such things as Log(yr) units, central concepts are misnamed. More unfortunately still, the all too critical treatment of previously published results make the rationale adopted here end up in incoherences in some instances. I think it is not a referee's duty to raise these types of issues, which ought to be purged or fixed at an earlier stage.

We will correct the units in the figures and revise where concepts are incorrectly stated. Also see detailed answers to specific comments below.

We certainly did not intend to be overly critical with respect to previous results (see answer to Reviewer 1). We would like to clarify that the highlighted issues are inherent to all inverse modeling/model fitting approaches, including the one adopted here. Hence, we advocate for including additional constraints that can help alleviate some of these limitations. We will clarify this point. Afterall, the findings of this study corroborate the findings of previous continuum model studies. Also see detailed answer to specific comments below.

In conclusion I cannot recommend this paper for publication in Biogeosciences unless it undergoes a major revision. Here are a series of points to consider for such a revision.

- *The actually used inversion algorithm must be described in detail. As the study is presented as an inversion study, this is obviously a conditio sine qua non.*

- *Meaningful uncertainties and correlations between the determined values must be reported. This may possibly require to use another inversion method if the currently used one cannot provide this information.*

- *The paper would benefit from a substantial shortening (by one third to one half).*

- *Attention should be paid to precise language, consistent notation and fair treatment of the existing literature.*

See detailed answers to specific comments below.

Specific comments

Model Description and Inverse Method

The description of the Biogeochemical Reaction Network Simulator (BRNS) comes with a great deal of details. Most of these can, however, be found elsewhere. This part could essentially be reduced to one half to one page at most and, for the readers' convenience, the current model description transferred to the Supplement.

We think that, in the context of this manuscript, it is important to provide a detailed model description, but agree that this could be moved to the supplementary information. Therefore, we will provide a short, more general model description in the main manuscript and move the detailed description (including a summary table of used symbols and their meaning) into the Supplements.

On the other hand, the inverse method used is presented only very summarily and above all in very general terms. Nothing is said about the optimization algorithm or the inversion procedure, which makes me wonder if it does not reduce to a trial and error approach. This would then also explain why there are no uncertainties or correlations reported for the derived n and a values. We are only told that the adjustment is carried out with the TOC and pore water SO4 concentration profiles (and CH_4 *where available) and that there is an exploration stage on a coarser grid of values (by the way a values above* 10^4 *are not evenconsidered—see conclusions) However, nothing is said about the applied metric (*χ^2*?) to quantify the misfit etc. nor about the criterion to select the results.*

For an inversion study—and this study is announced as such—this is clearly insufficient.

We will extend the description of the inverse modelling approach in section 3.2.3 and will also provide a table summarizing the statistical information for each of the best fit a—v couples (see below).

In short, we determined the optimal parameter set by assuming that the rank of a parameter set depends on the similarity between simulated and measured data. The optimal parameter set was thus found by minimizing the misfit between model results and simulations. As illustrated in the supplementary information, we quantified the similarity between the simulated and observed OM and SO_4^{2-} depth profiles in terms of correlation coefficient (r), centred root mean square difference (RMSD), and standard deviations (SD). All these measures can be graphically represented in a single Taylor diagram (Taylor, 2001) that statistically quantifies the misfit between the observed and simulated data and visualizes how closely the simulated depth profile resembles the observed one. If available, observed dissolved methane (CH_4) profiles are also used to evaluate the model-data fit, albeit in a non-quantitative way due to common sampling and measurement uncertainties in ex-situ measurements of CH_4 from deeper sediment layers.

As stated in section 3.2.3, the RCM parameters were inversely determined by first running the model for each site with a set of a—v couples over the entire range of previously published a (a = 10^{-3}–10^{7} yr) and v (v = 10^{-2}–10^{0}) values. The approach adopted for the sampling of the parameter space is a critical step in inverse modeling since the chosen statistical measures may reveal in, addition to a global optimum, multiple local optima. Because of the comparably low computational cost of the forward modeling, the low number of unknown parameters (two parameters in our case), as well as the weak nonlinearity of the problem, we here used a simple two-step nested regular sampling of the two-dimensional parameter space. First, the entire plausible a—v space (see above) was sampled on a coarse regular grid ($\Delta \log(a) = 1$, $\Delta\ v = 0.02$) and, for each of these runs, r, RMSD, and SD were calculated for both OM and SO_4^{2-} profile. Based on these measures, a new, more finely resolved a—v grid (10x10) was defined around the a—v couples that showed the smallest combined misfit between simulated and observed OM and SO_4^{2-} profiles. On this grid, a—v couples were again sampled regularly and the optimized a—v couple was determined by the lowest combined misfit (Table 1).

Because both a—v exert an influence on the apparent OM reactivity and its evolution with depth, they are not completely independent parameters. For instance, a decrease in v (decreasing reactivity) can be compensated by a simultaneous decrease in parameter a (increasing reactivity), thus producing a few equally statistically satisfying fits between the simulated and observed profile (e.g., Fig. S1). However, because changes in a—v exert different effects on OM and SO_4^{2-} depth profiles (see Fig. S1 & S2), including two species reduces the impact of the a—v correlation on the uniqueness of fit.

Table 1. Summary of statistical tests for the best fit a—v pair at each site derived from Taylor Diagrams.

	Reactivity parameters			**Organic matter**				**Sulfate**			
Site	a	v		**Mean**	**SD**	**RMSD**	***r***	**Mean**	**SD**	**RSMD**	***r***
A	$5.0\cdot10^{1}$	0.100	data	2.97	0.03			18.49	3.57		
			model	4.90	1.62	1.64	0.513	17.51	3.44	0.82	0.947
B	$3.0\cdot10^{0}$	0.125	data	1.96	0.02			14.10	8.07		
			model	1.92	0.16	0.14	0.640	12.59	9.18	1.79	0.974
C	$1.0\cdot10^{2}$	0.125	data	1.28	0.01			27.51	0.80		
			model	1.29	0.01	0.005	0.9998	27.76	0.89	0.37	0.823
D	$1.0\cdot10^{0}$	0.060	data	2.41	0.11			6.00	6.18		
			model	1.63	0.86	0.77	0.738	5.51	5.83	0.57	0.994
E	$1.0\cdot10^{0}$	0.155	data	2.17	0.60			1.54	1.83		
			model	1.79	0.46	0.40	0.561	1.35	1.84	0.23	0.984
F	$3.0\cdot10^{1}$	0.125	data	0.89	0.22			1.71	4.93		
			model	0.89	0.08	0.25	0.014	1.75	5.18	0.76	0.980
G	$1.8\cdot10^{3}$	0.100	data	1.12	0.45			9.54	9.15		
			model	1.34	0.00	0.45	0.597	9.70	10.08	1.52	0.984
H	$1.0\cdot10^{1}$	0.130	data	0.47	0.18			4.40	7.10		
			model	0.38	0.04	0.20	0.240	5.88	8.84	2.73	0.931
I	$3.5\cdot10^{3}$	0.110	data	1.69	0.34			3.51	5.84		
			model	1.84	0.00	0.34	0.002	3.35	6.23	0.64	0.993
J	$1.0\cdot10^{0}$	0.250	data	0.67	0.13			3.26	5.67		
			model	0.78	0.22	0.19	0.289	2.81	5.21	0.50	0.999
K	$1.0\cdot10^{1}$	0.390	data	0.56	0.12			5.30	7.79		
			model	0.48	0.24	0.22	0.154	4.73	6.97	0.90	0.998
L	$1.0\cdot10^{1}$	0.330	data	0.49	0.15			17.67	7.45		
			model	0.78	0.42	0.33	0.544	9.62	6.14	2.74	0.878
M	$5.0\cdot10^{3}$	0.270	data	0.82	0.34			10.38	9.93		
			model	0.94	0.17	0.24	0.566	7.52	7.74	2.79	0.961
N	$2.0\cdot10^{3}$	0.125	data	0.84	0.07			7.56	8.89		
			model	0.78	0.12	0.10	0.298	7.32	8.72	0.62	0.995

The Reactive Continuum Model

The presentation of the Reactive Continuum (RC) model is rather one-sided, especially when it comes to comparing it to n-G models with all their disadvantages.

Are reactive continuum models really that superior to the discrete models? Although they allow to describe the same phenomenon with fewer free parameters than, say, 3-G models, I do not think this comparison is completely fair. The shape of the chosen probability distribution easily substitutes for a number of parameters. The number of parameters for a 3-G model could also be reduced by prescribing the ratios of successive class sizes (i. e., by prescribing the shape of their distribution). RC models are *certainly attractive, but is their predictive ability really superior to few-G models that are easier to understand? Mathematical complexity is not a compelling argument. Since the main purpose of this study is the derivation of a and n parameter values for the RC model in different settings, that RC model is actually a centerpiece of the paper and it should be discussed in a more adequate way,*

We do not mean to discard discrete OM degradation models and fully acknowledge that they have been instrumental in making significant advances in our quantitative understanding of biogeochemical cycling in marine sediments. We also generally agree that RC and multi-G models (generally 2G or 3G models) perform equally well in reproducing comprehensive local porewater and sediment depth profiles, as well as estimating biogeochemical reaction rates and recycling fluxes in the upper sediment layers (tens of centimeters).

However, RC are characterized by several advantages over discrete solutions:

- RC models are in better agreement with our current conceptual understanding of OM reactivity and its evolution during burial.
  They emphasize the notion that OM is continuously distributed over a wide spectrum of reactivities and resolve the widely observed continuous decrease in OM reactivity with burial depth. They are fully consistent with widely observed, continuous decrease in OM reactivity in soils, lacustrine and marine sediments (Middelburg, 1989; Boudreau and Ruddick, 1991; Forney and Rothman, 2012; Katsev and Crowe, 2015). In contrast, multi-G models divide the complex and dynamic OM mixture into a number (generally one to two and less often three) distinct pools and OM reactivity than converge to a constant, fixed reactivity at depth. Therefore, RC models have a higher predictive ability and are a better choice for the representation of OM degradation dynamics in deeper sediments or on longer timescales.

- They are less sensitive to noise in the data.
  RC only require constraining two unknown parameters or, as indicated by previously published results and further corroborated by the results presented in this manuscript, one parameter (parameter a). Thus, they are simpler than multi-G representations as the dynamic information is contained in only two parameters. In contrast, discrete nG-Models require identifying n*2 unknown parameters. Consequently, multi-G models are not robust to noise in the data and are over-parametrized (e.g., Jørgensen, 1978; Forney and Rothman, 2012; Middelburg, 1989, 2018). Therefore, RC approaches facilitate model parametrization in data poor area.

- They are the most suitable degradation model type to explore links between reactivity parameters and environmental controls.

  Multi-G models are too over-parametrized to identify robust quantitative links between the plethora of environmental controls and OM reactivity (Forney and Rothman, 2012). Therefore, RC represent the most suitable choice for the study presented here.

Although it is beyond the scope of our work to provide an extensive discussion on the advantages and disadvantages of RCM and G-type models and this topic has already been extensively discussed in a number of previous articles (e.g., Manzoni et al., 2009; Forney and Rothman, 2012; Arndt et al., 2013; LaRowe et al., 2020), we will revise the manuscript to clarify the points discussed above.

also addressing shortcomings (only parallel reactions, no interaction between the different k-components, . . .).

We will address these points shortly in the revised manuscript.

“OM is composed of a complex and dynamic mixture of compounds that are distributed over a wide, continuous spectrum of reactivities. This heterogeneity indicates that degradation of OM proceeds at different rates in parallel. Interactions between different compounds or transformations of compounds can change the reactivity of a given compound. While the RC model does not explicitly account for such interactions/transformations, the fitted parameters take these interactions implicitly into account.”

The results for a are critically dependent on the completeness of the sediment data used for the inversion and it is therefore important to ask how reliable the results obtained here can be. From eqn. (39A) in Boudreau and Ruddick (1991), we have for two instants in time, t1 and t2,

$$G(t_1) = \frac{g_0}{(a+t_1)^\nu} \quad \text{and} \quad G(t_2) = \frac{g_0}{(a+t_2)^\nu} \tag{1}$$

and thus, assuming that t2 > t1

$$G(t_2) = G(t_1)\left(\frac{a+t_1}{a+t_1+\Delta t}\right)^{v} \quad \text{with } \Delta t = t_2 - t_1 \tag{2}$$

which is a generalization of eqn. (39B) from Boudreau and Ruddick (1991), i. e., eqn. (6) in the manuscript. Hence due to "[. . .] the common loss of the upper few centimetres during sampling [. . .]" (lines 438–439), the core top may be missing and thus the critical fresh POC(0) (eqn. (6)) cannot be recovered. Instead, the organic matter at the assumed sediment water interface actually has already (an unknown) age t1 and not 0 as supposed.
The inversion procedure will thus provide an estimate for a+t1 instead of the expected a. Depending on t1, these two may be significantly different from each other. k(0) at the sediment water interface (eqn. (23)) incorrect.

As discussed in the manuscript and correctly pointed out here, core loss and/or the definition of the sediment-water interface in very dynamic settings is a common limitation of all inverse modeling/ model fitting studies. Including additional species depth profiles provides additional constraints and can alleviate this limitation. However, given that macrobenthic activity often efficiently mixes the upper layer (<10 cm) of the sediment, we expect that potential core top loss exerts a limited impact on bioturbated sediments. Its impact might become important at sites characterized by high sedimentation rates and/or an enhanced delivery of initially very reactive OM. In these cases, the inversely determined initial apparent reactivity might not be representative for the reactivity at the sediment-water interface (although eqn. 23 remains correct). We will further emphasize these points in the revised manuscript.

Imprecisions, errors and incoherences

The text is very long, the language not always precise, and it unfortunately even contains a number of errors.

We will carefully revise and address these issues. Also see detailed answer to specific comments below.

In the model description, we read that the sequence of terminal electron acceptors (TEAs) is "described by a combination of Michaelis-Menten and inhibition terms" (line 296). The description that follows suggests that these rate law formulations are the same as in Van Cappellen and Wang (1996). This gets confirmed later at lines 380ff, where part of the information is repeated and completed—it would be good to avoid such repetitions and give one complete description in a single place. Van Cappellen and

Wang (1996) call their approach a modified Monod formulation or modified Monod kinetic formulation. Notwithstanding the debate whether Michaelis-Menten or Monod is the more appropriate denomination (see Boudreau and Westrich (1984) and references therein on this issue), the expressions used here correspond neither to the former nor the latter as both would imply a hyperbolic shape. Keeping the denomination from Van Cappellen and Wang (1996) would be just fine.

We will avoid repetitions and follow the definition by Van Cappellen and Wang (1996).

The description of the multi-G approximation of the RC model in the bioturbated layer is impossible to understand.

We will carefully revise the description of the multi-G approximation.
In short, the PDF, which determines the fraction of OM with reactivity k at time t=0 is:

$$f(k,0) = \frac{a^{\nu}\, k^{\nu-1}\, e^{-ak}}{\Gamma(\nu)}$$

The corresponding CDF then gives the fraction of total OM with a reactivity less or equal than k at t=0:

$$F(k,0) = \frac{\Gamma(\upsilon, 0, a\, k)}{\Gamma(\upsilon)}$$

We then separate the total OM into 200 different discrete fractions by calculating the fraction of OM with reactivity $k_i = 10^{i+1/2}$ yr^{-1} at t=0 (for a given a and υ couple) with:

$$F_i(k_i,0) = \frac{\Gamma(\upsilon, 0, a * k_{i+1}) - \Gamma(\upsilon, 0, a * k_i)}{\Gamma(\upsilon)}$$

Again, the notation is misleading, as the symbol f was previously used to denote inhibition terms. (...) Besides these, there are too many other errors in the details.

The use of symbols and descriptions will be carefully revised. A summary table of symbols and their meaning will be included to the extended model description in the Supplements. This will assure consistency, as well as better guide the readers.

The incoherences between the treatment of previous studies and the approach adopted here are truly unfortunate. I am deeply annoyed to read on one hand that previous estimates of the RC model parameters were "compromised by several factors" (line 437) among which the steady-state assumption of OM deposition fluxes, and then see that same steady-state assumption being made here.

This is a misunderstanding and will carefully revise the first paragraph to avoid this misunderstanding.

The limitations discussed in the manuscript apply to all inverse model/model fitting approaches, including the one applied here. The intention of this section (p. 14, l. 437 onwards) was to emphasize that adding further constraints can alleviate some of these limitations. However, we also highlight that this is not always the case and discuss the potential effect of some of the mentioned limitations (i.e., core loss, non-steady state and uniqueness of fit) on our results. For instance, as discussed in the manuscript, results for the Bering Sea site are likely affected by non-steady state conditions, while results from the OMZ of the Arabian Sea might be influenced by H_2S oxidation processes that are not explicitly resolved due to the lack of data (p. 17, l. 515-522, p. 18, l. 554- 557).

Furthermore, we certainly do not discard previous findings. Previously published parameter values (also including those from multi-G model studies) are an integral part of our discussion and we highlight how our model results further corroborate previous findings throughout the manuscript (e.g., all throughout section 4.1.2, specifically p. 17, l. 535 onwards, p. 19, l. 585-587, but also section 4.1.4). We will carefully revise these sections to see how this can be further emphasized.

One furthermore reads that "multiple a and n pairs could potentially fit a given TOC profile" (lines 440–441), i. e., that the fitted values for these parameters are more or less strongly correlated, and then not even find a hint of a check for possible correlations between the values determined here. The extra constraint from the pore water sulfate profile may alleviate the degree of correlation between the two parameter values, but it is highly unlikely that the correlation between their values would go away.

As discussed above, the adopted inverse modeling approach (four statistical measures, 2 species) alleviates the limitation induced by the correlation between parameter a and v (see discussion above). In addition, our best-fit parameter estimates do not reveal a correlation between a and v.

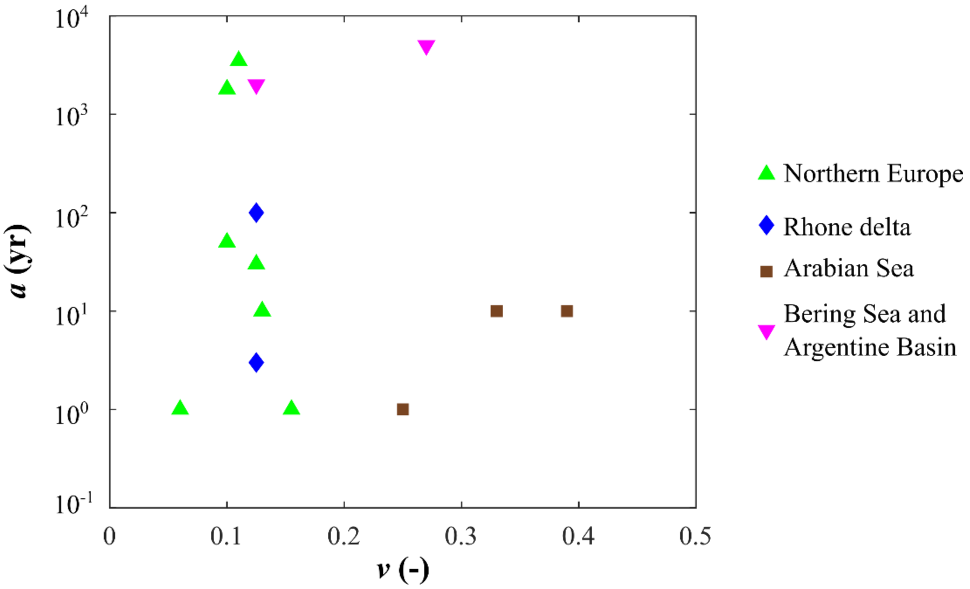

Figure 1. Model-derived a and v values for best fits determined in this study (see Table 1).

We will add a short discussion of this point in the section describing the inverse modeling approach.

Accordingly, it does not make any sense to conclude that the whole story boils down in the end to determine solely a. In the absence of information about the correlation between a and n values, such a conclusion is simply unwarranted. Similarly, problems related to the difficulties "to quantify OM contents at the sediment-water interface (SWI) due to the common loss of the upper few centimetres during sampling [. . .]" (lines 438–439) are emphasized for the results of previous studies, but not even mentioned regarding the results presented here. The uncertainties on the a values derived here may thus possibly be significantly larger than one might think, as illustrated above.

While we agree that model results are subject to the discussed limitations (non-steady state, core top loss, uniqueness of fit); however, we reiterate that adding additional constraints alleviates these limitations. In addition, as discussed in the manuscript, our conclusions are fully supported by:

- Several different, previously published RC model studies, each using a different model structure, different observational data (degradation rates, pore water depth profiles, OM depth profiles) and spanning a wide range of different environments and timescales (Middelburg, 1989; Boudreau and Ruddick, 1991; Forney and Rothman, 2012; Katsev and Crowe, 2015; Jørgensen and Parkes, 2010; compilation by Arndt et al., 2013).
- Previously published, additional observational data from the respective sites (e.g., direct rate measurements, other model studies).
- Their consistency with our current understanding of the main controls of OM degradation in marine sediments (see discussion in section 4.1.4).

We will further emphasize this in the revised manuscript.

Finally, in the conclusions readers are "[. . .] caution[ed] against the use of single environmental master variable such as water depth, sedimentation rate or organic matter deposition fluxes beyond the local scale" (lines 912–914) and yet the model simulations rely upon exactly this kind of master variable when it comes to determining the solids' burial rate, w(z), which is above all quite critical for the outcome of the paper. It is furthermore not entirely clear what is meant by "beyond the local scale." Does "local" stand for regional or for a given environmental setting, or does it stand for a given depth, deposition rate or some other chosen independent master variable? This should be stated more unambiguously.

We argue that the complex and dynamic set of environmental factors that control OM reactivity cannot be easily reduced to one single master variable such as water depth or sedimentation rate. This is supported by the lack of statistically significant relationships between the inversely determined OM reactivity parameters established here (as well as in previous studies, e.g., Arndt et al, 2013) and those single environmental master variables. However, we do observe – albeit weak – relationships between OM reactivity and water depths or sedimentation rate on the regional scale (i.e., under broadly similar environmental conditions). By *local scale* we thus refer to a specific, local depositional environment (see discussion p. 22, l. 674 onwards), in contrast to multiple depositional environments or the global scale. We will clarify this.

We do make use of empirically derived global relationships that link water depth with sedimentation rates and bioturbation coefficients (Middelburg et al., 1997). However, in contrast to OM reactivity data, the compiled global observational data sets in Middelburg et al. (1997) reveal statistically significant relationships between sedimentation rate/ bioturbation coefficient and water depths that are further confirmed by additional global data (e.g., Burwicz et al., 2011; Solan et al., 2019).

In addition, the empirical water depth–sedimentation rate relationship is only applied to estimate sedimentation rates at sites for which no direct observational data was available (e.g., Severn estuary). For the majority of our sites, we apply observed sedimentation rates. Unfortunately, bioturbation rate estimations are often not available for the studied sites, thus we mostly rely on the empirical global relationship to estimate bioturbation coefficients.

Conclusions

As mentioned in the general appreciation of manuscript, I find that the conclusions are overdone. The authors claim that they "[. . .] narrowed the most plausible range of a to 10^0*–*10^4 *years." (line 909), from a range of* 10^{-3}*–*10^7 *derived from a global compilation. This reads impressive. However, upon inspection of the results on Fig. 4, one realizes that the size of "This study" is only about one quarter of the "Global compilation", and that more than 90% of the values for a from the "Global compilation" were already within* 10^0*–*10^4*. As models with complex reaction networks are known to generally present a stable rather than a dynamic behavior when it comes to the sensitivity with respect to parameter value perturbations, the clustering of the model results around the mean or the median was thus to be expected. It is therefore not clear how significant and important this narrowing is in the end.*
A reformulation of the conclusions in more modest terms would be appropriate.

The significance of our findings, alongside previous publications, is the consolidation of this information. As pointed out by the authors of previous studies, more data and consistent model approaches would be needed to confirm the observed parameter trends (e.g., Boudreau and Ruddick, 1991; Arndt et al., 2013). Previous model results have been obtained using different model structures and observational data sets of different comprehensiveness, questioning the direct comparability of these results. By gathering these data and confirming this window, we offer means for model parametrisation when dealing with data poor areas and timescales based on a robust and transferable dataset.

In addition, the model applied here, and certainly some of these applied previously, do not include complex reaction networks. We focus on OM dynamics mainly driven by organic carbon, sulfate, and methane interplay (see Table S1).

Technical comments

Throughout the paper: as far as I know, the symbol used for "year" in Biogeosciences is not y but yr.

Agree.

Throughout the paper: watch out for spurious points after equation numbers

Agree

Line 247: "Findley" should read "Findlay"

Agree

Line 264: "a Gamma function" should read "a gamma distribution" (the name is not capitalized; the gamma function and the gamma distribution are two completely different things— here it is the distribution that is relevant)

Agree

Lines 267 and 269: please use the standard notation for the gamma function which has been Γ *for more than two centuries (and not* τ*; also notice that "the Gamma function" should read "the gamma function"*

Agree. The error must have occurred during editing of the manuscript.

Lines 267ff: before this line, the RC parameters were denoted by a and n. Now they become a and n and on line 477, it is switched back to a and n . . . Eq. (5) here has both. Please chose one notation and use it throughout. Be aware that the irrigation parameter is also denoted by the symbol a.

Agree. The error must have occurred during editing of the manuscript.

Line 286: "the Gamma distribution" should read "the gamma distribution"

Agree

Line 286: "the burial age (z = z – zbio)" is ambiguous/unclear

We will clarify here to avoid confusion.

Line 296: "Michalis-Menten" should be "Michaelis-Menten" (but should be discarded anyway)

Agree

Lines 296–305: the style of this whole paragraph is very confusing and cannot be understood without a priori knowledge. Please rewrite and use precise language, starting by making a clear distinction between TEA and TEA concentrations.

We will revise this paragraph accordingly.

Lines 303 and 305 (eqns. (8) and (9)): something is wrong here – neither $j > Kj$ nor $j \leq Kj$ make sense. Also: there is a spurious point after the equation number (9).

We will revise Eq. 8 and 9.

Line 309: "biomolecular" should read "bi-molecular"

Agree

Lines 319ff and eqn. (10): is this correct? What is ki? Is the numerator at the right-hand side of eqn. (10) correct? There is a spurious point after the equation number.

This will be changed with improved description of the multi-G approximation within the bioturbated zone.

Line 331: "incomplete Gamma function" should read "incomplete gamma function"

Agree

Lines 333 and 335 (eqns. (12) and (13)): τ should read Γ (four times)

Agree

Line 361: φ^2 *is certainly not tortuosity (symbol φ was previously used for porosity). . .*

Agree

Lines 370ff: please specify the units of the parameters and of the resulting Dbio

zbio is given in cm; Dbio is given in cm^2 yr^{-1}

Line 411: should read "quadratically distributed for ξn >> ξc and linearly distributed for ξn << ξc"

Agree

Line 418 (eqn. (21):

Yes, the run time is dependent on the length of the model domain L, thus $run\ time = L/\omega$.

Line 418: "where concentrations exceed mM": something must be missing here

It means 'concentrations on the order or above millimolar'. We will clarify that.

Lines 486ff: here, F is used as a symbol for fluxes; elsewhere (section 3.2.1.3), F denotes fractions; a few lines below, the symbol for fluxes becomes by J. Please revise the notation scheme for the whole paper and chose a consistent one to improve readability. As far as possible, please stay close to schemes used elsewhere in the literature

We will revise the use of notations throughout the manuscript to avoid inconsistencies and double uses.

Line 584: Not sure what the citation of Westrich and Berner (1984) has to do here as it is about (multi) G models, not about RC models.

The mentioned a value is derived from a fit to the data provided in *Westrich and Berner (1984)*. This might be misleading, and we will thus remove the reference.

Line 1068: 137Cs should read ^{137}*Cs*

Corrected.

Table 1: this table runs over five (!) pages, with some pages having only two (!) lines. This is completely impractical. It would be better to present this information, if required, in an alternative form.

We will reduce this table by removing the more descriptive column *Biological Parameters* and including this information to the Supplements. This should significantly improve the visualisation of the table and keep it to a reasonable size.

Table 4: character size too small (exponents essentially unreadable)

Font size will be increased to improve readability.

Figure 4: the x-axis annotation for part b) is incorrect. There is no unit such as Log y. What is shown on the x axis is log(a/yr), i.e., the logarithm in base 10 of a expressed in years. Logarithms never have units and neither have their arguments.
Figs. 5, 6, 7, 8: same as for Fig. 4 (all units related to logarithmic axes are wrong).

We will correct the axis notations on the figures and express them using scientific notation (10^x).

References

Arndt, S., Jørgensen, B. B., LaRowe, D. E., Middelburg, J., Pancost, R., and Regnier, P.: Quantifying the degradation of organic matter in marine sediments: A review and synthesis, Earth-Sci. Rev., 123, 53–86, 2013.

Boudreau, B. P. and Ruddick, B. R.: On a reactive continuum representation of organic matter diagenesis, Am. J. Sci., 291, 507–538, https://doi.org/10.2475/ajs.291.5.507, 1991.

Burwicz, E. B., Rüpke, L. H., and Wallmann, K.: Estimation of the global amount of submarine gas hydrates formed via microbial methane formation based on numerical reaction-transport modeling and a novel parameterization of Holocene sedimentation, Geochim. Cosmochim. Acta, 75, 4562–4576, https://doi.org/10.1016/j.gca.2011.05.029, 2011.

Forney, D. C. and Rothman, D. H.: Inverse method for estimating respiration rates from decay time series, Biogeosciences, 9, 3601–3612, https://doi.org/10.5194/bg-9-3601-2012, 2012.

Jørgensen, B. B.: A comparison of methods for the quantification of bacterial sulfate reduction in coastal marine sediments: II. Calculation from mathematical models, Geomicrobiol. J., 1, 29–47, https://doi.org/10.1080/01490457809377722, 1978.

Jørgensen, B. B. and Parkes, R. J.: Role of sulfate reduction and methane production by organic carbon degradation in eutrophic fjord sediments (Limfjorden, Denmark), Limnol. Oceanogr., 55, 1338–1352, https://doi.org/10.4319/lo.2010.55.3.1338, 2010.

Katsev, S. and Crowe, S. A.: Organic carbon burial efficiencies in sediments: The power law of mineralization revisited, Geology, 43, 607–610, https://doi.org/10.1130/G36626.1, 2015.

LaRowe, D. E., Arndt, S., Bradley, J. A., Estes, E. R., Hoarfrost, A., Lang, S. Q., Lloyd, K. G., Mahmoudi, N., Orsi, W. D., Shah Walter, S. R., Steen, A. D., and Zhao, R.: The fate of organic carbon in marine sediments - New insights from recent data and analysis, Earth-Sci. Rev., 204, 103146, https://doi.org/10.1016/j.earscirev.2020.103146, 2020.

Manzoni, S., Katul, G. G., and Porporato, A.: Analysis of soil carbon transit times and age distributions using network theories, J. Geophys. Res. Biogeosciences, 114, https://doi.org/10.1029/2009JG001070, 2009.

Middelburg, J. J.: A simple rate model for organic matter decomposition in marine sediments, Geochim. Cosmochim. Acta, 53, 1577–1581, https://doi.org/10.1016/0016-7037(89)90239-1, 1989.

Middelburg, J. J.: Reviews and syntheses: to the bottom of carbon processing at the seafloor, Biogeosciences, 15, 413–427, https://doi.org/10.5194/bg-15-413-2018, 2018.

Middelburg, J. J., Soetaert, K., and Herman, P. M. J.: Empirical relationships for use in global diagenetic models, Deep Sea Res. Part Oceanogr. Res. Pap., 44, 327–344, https://doi.org/10.1016/S0967-0637(96)00101-X, 1997.

Solan, M., Ward, E. R., White, E. L., Hibberd, E. E., Cassidy, C., Schuster, J. M., Hale, R., and Godbold, J. A.: Worldwide measurements of bioturbation intensity, ventilation rate, and the mixing depth of marine sediments, Sci. Data, 6, 58, https://doi.org/10.1038/s41597-019-0069-7, 2019.

Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, J. Geophys. Res. Atmospheres, 106, 7183–7192, https://doi.org/10.1029/2000JD900719, 2001.

Van Cappellen, P. and Wang, Y.: Cycling of iron and manganese in surface sediments; a general theory for the coupled transport and reaction of carbon, oxygen, nitrogen, sulfur, iron, and manganese, Am. J. Sci., 296, 197–243, https://doi.org/10.2475/ajs.296.3.197, 1996.