Responses to Comments by Referee #2

We thank Referee#2 for the detailed reviews he/she provided. Please find our detailed response below *(referee comments in Italic).*

General comments:

This manuscript by Ou et al. utilized a coupled physical-biogeochemical model to investigate the controlling factor of bottom hypoxia on the northern Gulf of Mexico and Louisiana-Texas Shelf. The authors added the phosphorus cycle and modified the sediment oxygen consumption module in an existing biogeochemical model NEMURO and coupled it with ROMS model. The coupled model was validated with observational data and then used to implement a 15-year hindcast simulation during 2006-2020. Then the authors explored the spatial variation of hypoxia development in the study area and found sediment oxygen consumption (SOC) and water column stratification are main factors to control the bottom oxygen in nearshore and offshore area respectively. Their model results also indicated separate hypoxia development schemes on the west and east Louisiana-Texas Shelf. Coastal deoxygenation is one of the most prominent environmental issues with important implications for marine ecosystem services. Although this paper made efforts to adopt a more sophisticated biogeochemical model with added phosphorus cycle and improved sediment oxygen consumption module, making contributions to investigate the spatial differences of dominant processes on hypoxia, it lacks original and novel aspects to explore the well-studied topic in this region, as well as comprehensive comparison with previous modeling study on the model performance, simulation results and conclusions, and address the question that how this new model stands out. The manuscript missed an advanced understanding and deep insight on the research topic of coastal hypoxia in a well-organized discussion section, thus this paper is a little thin on content. Although I see the value of this work, I perceive that the publication is premature at this time.

Response: Compared with existing modeling efforts, our model, for the first time, included a silicate cycle as well as multiple plankton functional groups, the importance of which has already been demonstrated in previous studies yet not included in hypoxia modeling efforts. We plan to include more results and discussion of the sensitivity tests focusing on the contribution of silicate limitation and the benefits of incorporating multiple plankton groups in the revision. The extensive model validation against nutrient and dissolved oxygen profiles confirmed the robustness of our model. In addition, another purpose of this model study is to provide the needed numerical solution for the development of a novel statistical model presented in paper Part II.

Major comments:

The hypoxia at the northern Gulf of Mexico has been well studied since the 1990s with increasing model studies in recent years. It ranged from a simple oxygen respiration model (Hetland&DiMarco, 2008) to a sophisticated coupled biogeochemical model (Laurent et al. 2012; Fennel et al. 2013). Including this study, they all generated similar conclusions that SOC is the controlling factor for hypoxia. In this sense, the improvement of complexity in the biogeochemical model does not make much sense. Also, the authors mentioned the additional work done on the NEMURO-based model filled gaps in phosphorus cycling and improved SOC representation. It's better to prove the advancement of the new model by validating with important variables, such as DO, Chla, PO4, NO3, with other model simulation studies.

Response: We carried out a series of sensitivity experiments to demonstrate the advance of incorporating the silicate cycle and will also provide more results to demonstrate the benefits of a more complicated plankton functional groups in DO simulation. We performed extensive model valuations against the DO (spatial distribution and vertical profiles) and nutrients (vertical profiles), and we will perform more validation for Chl-a.

We added some sensitivity tests to address the importance of silicate limitation. we plan to discuss silicate's impacts on hypoxia in two aspects: 1) the contribution of large phytoplankton (Pl, diatom); and 2) riverine silicate inputs. We were able to present some new model results with different riverine silicate inputs. The control run is symbolized as exp0 with the same setups as that in the first submission but with an updated sinking velocity changed to 5 m day^{-1} based on the above discussion (also see response below to comments "*L195: SOC/THKbot is basically the oxygen consumption rate in the sediment. Why notintegrate SOC in the hypoxic area and get an overall integrated SOC?Any observational data validation on the newly added sediment and phosphorus module? In addition to the oxygen concentration validation?*").

Please note that the following contents regarding riverine silicate inputs are also included in our response to reviwer #1.

We changed the riverine SiOH4 concentration from 0.2 to 2.0 with an increment of 0.2 in each experiment (exp1 through exp 9). For exp10 (ongoing), river silicate inputs were the same as that of the exp0, however, we removed the silicate limitation on large phytoplankton growth. More detail on these new experiment tests is listed in Table 1.

Table 1. Model setups of different sensitivity tests. Simulation of exp10 has not yet finished and thus updates will focus on results by testing exp0 through exp9.

	Simulation	Sinking velocity (m day^{-1})	Scale of Riverine SiOH4 concentration	Silicate limitation on Pl	# phytoplankton group	# zooplankton group
exp0 (control)	1 Aug 2017–25 Aug 2020	5	1	Yes	2	3
exp1	1 Aug 2017–25 Aug 2020	5	0.2	Yes	2	3
exp2	1 Aug 2017–25 Aug 2020	5	0.4	Yes	2	3
exp3	1 Aug 2017–25 Aug 2020	5	0.6	Yes	2	3
exp4	1 Aug 2017–25 Aug 2020	5	0.8	Yes	2	3
exp5	1 Aug 2017–25 Aug 2020	5	1.2	Yes	2	3
exp6	1 Aug 2017–25 Aug 2020	5	1.4	Yes	2	3
exp7	1 Aug 2017–25 Aug 2020	5	1.6	Yes	2	3
exp8	1 Aug 2017–25 Aug 2020	5	1.8	Yes	2	3
exp9	1 Aug 2017–25 Aug 2020	5	2.0	Yes	2	3
exp10	1 Aug 2017–25 Aug 2020	5	1.0	No	2	3

The hypoxic area is estimated as the sum of the area of model grids when the bottom DO is less than 2 mg l^{-1}. Percentage changes (Fig. 1) of the May–September hypoxic area are calculated between the sensitivity run (i.e., exp1–9) and the control run (exp0). The results of exp1 through exp9 indicate that the hypoxia area is mostly positively correlated with riverine silicate inputs. It is worth noting that the impact on the hypoxic area due to the changing riverine loads is not linear. The hypoxic area is more sensitive to elevated silicate supply especially when riverine input increases by more than 1.2. The average percentage increment of the hypoxic area ranges from ~35% to ~72% as the riverine silicate supply increase by 40% to 80%. In contrast, the hypoxic area decreases by ~19% to ~53% as the silicate supply decreases by 40% to 80%. When the changes in riverine silicate supply are less than 20%, the changes in the hypoxic area would be expected to be greater when riverine silicate is reduced (-19%) than increased (+12%).

We found three points that are new to coastal managers. Firstly, decreases in riverine silicate loads by 20% and 40% do not lead to significant differences in terms of hypoxic area reduction (by ~19% for both cases). Secondly, it is hard to meet the Gulf Hypoxia action plan goal of a 5,000 km^2 hypoxic area by reducing the riverine silicate loads solely. The average summer hypoxic area from 2018 to 2020 is 15000 km^2 which is comparable to that of 1985–2010 (14000 km^2). Thus, to meet the Gulf Hypoxia action plan goal, the average percentage reduction of the hypoxic area should be ~67%. More discussion of the combined silicate and nitrogen reduction is needed and will be provided in the revision. Thirdly, as the range of 25th–75th percentile of hypoxic area changes enlarges as the riverine silicate load increases, an elevated riverine silicate input is likely to introduce much worse hypoxia events.

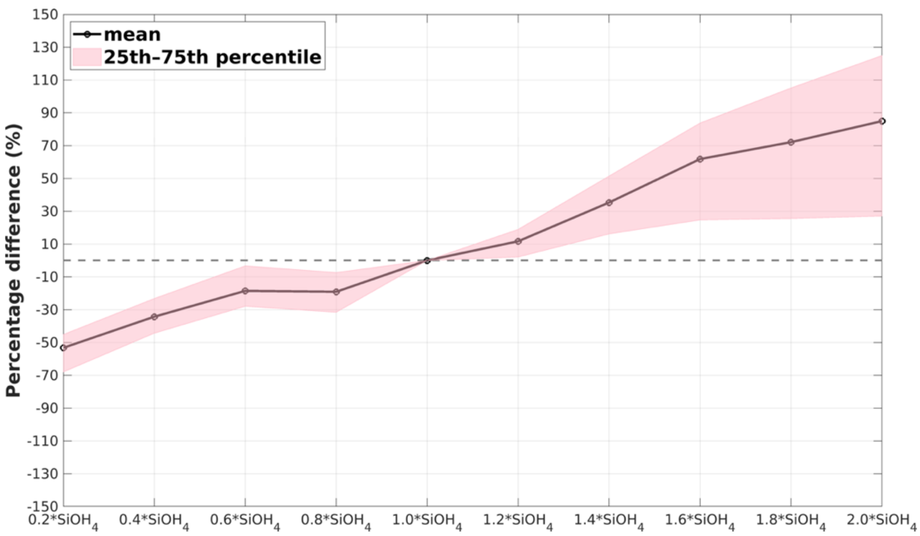

Fig. 1 Percentage differences of the simulated hypoxic area between sensitivity tests (exp1–9) and the control run (exp0). Statistics are based on the simulations in May–September from 2008 to 2020.

The differences in the spatial DO distribution between the riverine inputs sensitivity test (i.e., exp1–9) and the control run in August (2018–2020 average) are shown in Fig. 2. When riverine silicate is reduced, the low slope west shelf is more sensitive to the changing silicate supply than the east shelf. In the cross-shelf direction, bottom DO between 10–50 m isobaths is more sensitive to the reduction of riverine silicate inputs than the rest regions. A slight decrease (by 20%; Fig. 2d) of silicate supply would lead to a maximum bottom DO increase of 2 mg O_2 l^{-1} in this region. A 20% decrease in silicate supply can therefore easily induce a change from hypoxic to normoxic bottom waters in such regions. When riverine silicate inputs are increased, not much difference in the spatial distribution of DO reduction between the west and east part of the La-Tex shelf until the increase is more than 80%—then DO drops more in the west part of the shelf than the east part.

The above results show the impacts of riverine silicate loads on the bottom hypoxic area and bottom hypoxic water extent. We found that 1) the distribution of the hypoxic waters is more sensitive to the elevated riverine silicate loads with greater uncertainties than the reduced inputs; 2) a dual or triple silicate reduction is needed to meet the goal of the Gulf Hypoxia action plan; 3) the responses of bottom DO concentration is not spatially homogeneous along the shelf when riverine silicate loads are adjusted, and 4) the west shelf will suffer more from hypoxia conditions when riverine silicate is increased by more than 60%. We will provide further analysis and discussion on this topic in the revision including a recommendation of combined nitrogen and silicate reduction to meet the goal of the Gulf Action Plan.

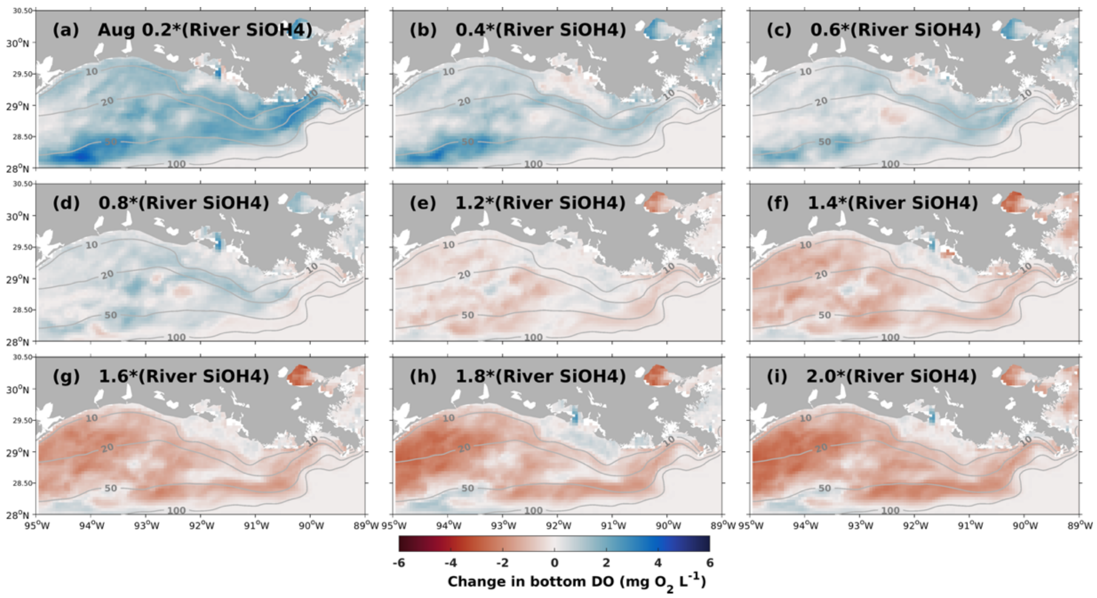

Fig. 2 Differences in bottom DO (in mg O_2 l^{-1}) between experiments with different riverine inputs and the control run (August mean of 2018–2020).

The oxygen balance analysis is confusing and questionable. Although SOC is the dominant process in the bottom hypoxia generation (You et al. 2015), water column respiration (WCR) should not be orders of magnitude smaller than SOC, especially in the whole water column, as shown in Figure 15 and L455-456. Observational studies still showed varying evidence on SOC contribution (Murrell&Lehrter, 2011; Quiñones-Rivera, et al. 2010). More importantly, the reviewer has a sense that the authors did not understand and explain the oxygen dynamics well (Figure 10 and 15, section 4.2).

What is oxygen balance in the text? Based on L450-452, it should be water column respiration plus phytoplankton photosynthesis. This is a very confusing term and the physical transport of oxygen was totally missing. A lot of oxygen studies utilized standard oxygen budget analysis to separate dynamic terms in oxygen change (Li et al. 2014; Scully 2013; Yu et al. 2015). Please refer to those studies on the analysis and consider recalculating/rewriting this part.

Response: We agree with Reviewer#2 that the impact of WCR should not be neglected. In those figures, we mainly present the monthly climatology of model results which might make WCR less important. We plan to output all terms related to oxygen budget, including diffusivity, and advection, and calculate oxygen budget following the literature suggested by Reviewer#2.

Although this study employed sophisticated machine learning techniques to determine the controlling mechanisms on hypoxia in different regions. It could be actually achieved by oxygen budget analysis, with much clear representation in physical terms (advection and diffusion), rather than relying on stratification indicators. In addition, compared to the manipulating force on DO variability on a seasonal scale, the interannual variability is more of interest and worthy to look into.

Response: Following our responses to the above comments, we will provide a more comprehensive comparison of the DO balance including the local rate of changes, advection, and diffusion. According to both referees' comments, DO seasonality seems to be relatively well studied. We will look into interannual variability as well as the mechanism behind that in the revision.

The manuscript missed a comprehensive discussion section of advanced understanding of the study topic in-depth and in breadth. The overview of previous observational and model studies in this region, comparison with the current study, what are the agreements and differences, what are the causes, what are the defective aspects of this study, etc. are all important points to include. Expanding the implication to the global context is also valuable to discuss.

Response: We agree with reviewer #2 on this point. Our plan is to incorporate a series of sensitivity tests and more quantitative analysis regarding the contribution of silicate and the benefit of utilizing more complicated plankton functional groups in DO dynamics than existing modeling studies. We will also include more analysis of interannual variability.

Detailed comments:
Method
L105-106: are the new features of this biogeochemical model suitable for NGoM?

Response: For zooplankton modeling, we followed Shropshire's et al. (2020) and focus on the zooplankton dynamics in the Gulf of Mexico, which use a similar model setup. According to that study, the zooplankton community could substantially affect the primary production in the study area.

Reference:
Shropshire, T. A., Morey, S. L., Chassignet, E. P., Bozec, A., Coles, V. J., Landry, M. R., ... & Stukel, M. R. (2020). Quantifying spatiotemporal variability in zooplankton dynamics in the Gulf of Mexico with a physical–biogeochemical model. Biogeosciences, 17(13), 3385-3407.

L108: what is PL? should it be LP (large phytoplankton)?
Response: The PL is a typo. It should be LP.

L120-122: no reactive, labile and refractory category in organic matter pool? In other words, is a single reaction rate enough?

Response: We included a burial PON pool in the conceptual sedimental layer. The PON settled down at the conceptual sedimental layer fuels the PONsed pool, which is a reactive labile organic sediment pool. After a portion of PONsed is decomposed during aerobic and anaerobic processes in sediment (see the SOC scheme), a certain portion of PONsed is burial and fueling the PONburial pool, which will be removed from the system.

L156: What are ExcZS, ExcZL and ExcZP represented (I could not find those in the Appendix, and guess they should be zooplankton excretion rate to NH4?)? Why not include the zooplankton respiration term?

Response: ExcZS, ExcZL and ExcZP are typos in the equation. They should be ExcZSn, ExcZLn, and ExcZPn, respectively as shown in Table B2. In our model, we combined zooplankton excretion and respiration. Thus, during excretion, zooplankton consumes oxygen.

L158-159: How did oxygen inhibition on nitrification and aerobic decomposition rates were calculated? Using Michaelis–Menten formula?

Response: The oxygen inhibition (Fennel et al., 2006; 2013) is considered as the maximum of 0 and an oxygen-dependent unitless term. It uses Michaelis–Menten formula. The inhibition term ($\hat{r}$) is described in A2 and A3 with the relative parameter description in Table B4.

$$\hat{r} = max\left[\frac{max(0, Oxyg - Oxyg_{th})}{K_{Oxyg} + Oxyg - Oxyg_{th}}, 0\right]$$

Where

Parameter	Description	Units	Values

K_{Oxyg}	Oxygen concentration at which inhibition of nitrification and aerobic respiration are half-saturated	mmolO_2 m^{-3}	3.0
$Oxyg_{th}$	Oxygen concentration threshold below which no aerobic respiration or nitrification occurs	mmolO_2 m^{-3}	6.0

Reference:
Fennel, K., Wilkin, J., Levin, J., Moisan, J., O'Reilly, J., and Haidvogel, D.: Nitrogen cycling in the Middle Atlantic Bight: Results from a three-dimensional model and implications for the North Atlantic nitrogen budget, Global Biogeochem. Cycles, 20, 1–14, https://doi.org/10.1029/2005GB002456, 2006.

Fennel, K., Hu, J., Laurent, A., Marta-Almeida, M., and Hetland, R.: Sensitivity of hypoxia predictions for the northern Gulf of Mexico to sediment oxygen consumption and model nesting, J. Geophys. Res. Ocean., 118, 990–1002, https://doi.org/10.1002/jgrc.20077, 2013.

L164-166: how was the portion of sinking PON buried (PONburial) determined? How the initial sediment PON pool was calculated? Is there also an anaerobic layer? Is there any exchange between PONburial and PONsed?

Response: The burial faction is determined using the scheme embedded in the original NEMURO model, where the burial faction is a function of the vertical flux of particulate organic matter (POM). As the organic matter is buried, it will leave the system without returning to PONsed.

The PONsed is initialized as 0 due to a lack of available data to initialize the model. Our model does not include a sediment module, thus, the sediment PON pool is in an imaginary or conceptual sediment layer. In this layer, aerobic decomposition, nitrification, and denitrification occur simultaneously following the linear relationship between denitrification rate and total oxygen consumption rate (Eq. (5)). So, there is no specified anaerobic layer.

L193: the description of THKbot is confusing. Is it the thickness of overlying water, or sediment layer?

Response: THKbot is the thickness of overlying water or the thickness of the bottom layer of the ocean model. In our model, we do not separate overlaying water and bottom water. We consider THKbot since we assumed that oxygen consumption at the conceptual sediment layer directly contributes to decreases in oxygen concentration in the overlying water.

L195: SOC/THKbot is basically the oxygen consumption rate in the sediment. Why not integrate SOC in the hypoxic area and get an overall integrated SOC?
Any observational data validation on the newly added sediment and phosphorus module? In addition to the oxygen concentration validation?

Note: both Reviewer #1 and Reviwer#2 share similar comments on this, so the following response is partially copied from our responses to Reviewer #1.

Response: We assumed that oxygen consumption at the conceptual sediment layer directly contributes to decreases in oxygen concentration (only) in the overlaying water. It also implies that the oxygen consumed in the sediment is from the overlying water. The rate of oxygen removal due to sedimental biochemical processes is thus assumed the same in the overlying water. We transform the area concentration of SOC (mmolO_2 m^{-2}) by dividing the THKbot to the volume concentration of DO.

We updated the validation for SOC and the ratio of SOC/overlaying water respiration. Such validation is conducted due to questionable sinking velocity (previously set to be 15 m day^{-1}) pointed out by the other reviewer. We, therefore, added two sensitivity tests with different sinking velocities as 1 m/day and 5 m/day, respectively. We chose the best set-up by validation of SOC, the ratio of SOC and overlaying water respiration, bottom hypoxic area, and bottom hypoxic extent. Model set-ups are the same in all the tests except the sinking velocity of PON. Measured SOC and overlaying water respiration are derived from McCarthy et al., (2013), while the measured hypoxic area and extents are from the Shelf-wide cruises. Following McCarthy et al., (2013), we extract the daily SOC, and overlaying water respiration at sites F5, C6, B7, and MRM (Fig. 3 below) and applied the monthly average for months compared.

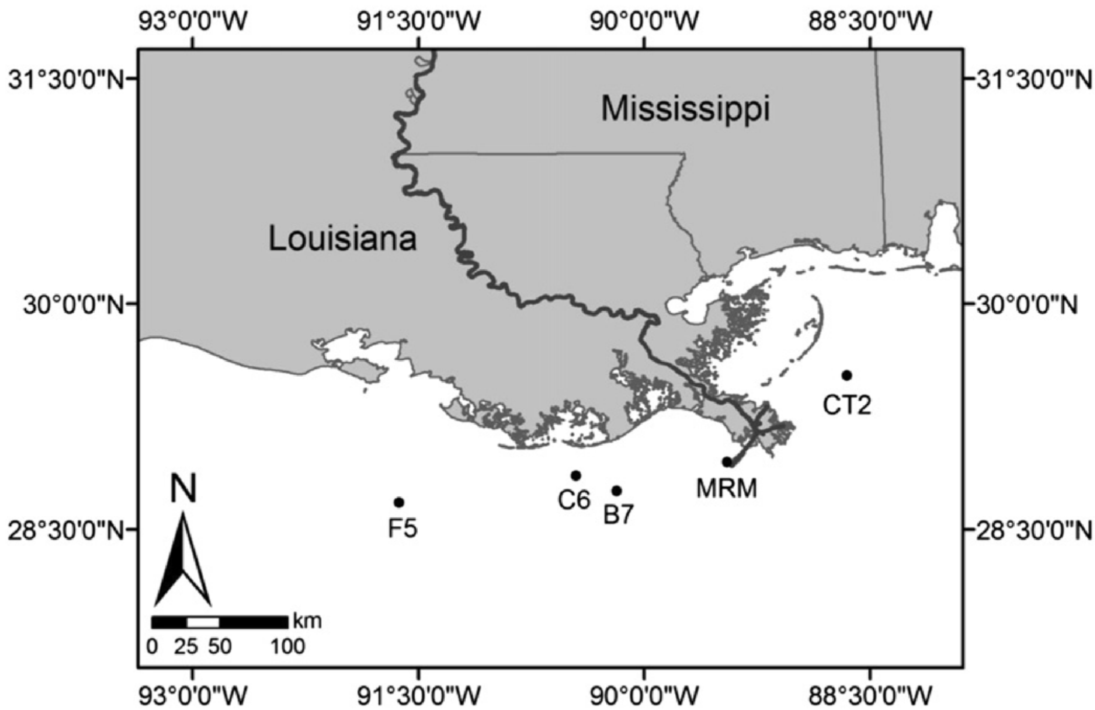

Fig. 3 Map showing the location of sampling site in the northern Gulf of Mexico (McCarthy et al., 2013).

Fig.4 indicates that a sinking velocity of 5 m/day provides the best estimate of SOC. The root-mean-squared errors (RMSEs) are 567 μmol m^{-2} h^{-1}, 713 μmol m^{-2} h^{-1}, and 452 μmol m^{-2} h^{-1} for sensitivity tests with a sinking velocity of 15 m/day (used in the first draft), 1 m/day, and 5 m/day, respectively. The simulated (5 m/day) and observed SOC are generally in the same order of magnitude. The model results in general overestimate the SOC at sites F5 and C6 except for January 2009 and May 2010 at site C6, and underestimate SOC at sites B7 and MRM. Times series also reveals that the magnitude of simulated SOC by tests with a sinking velocity of 5 m/day is generally within the measured range (Fig. 5) over the entire year. The magnitude of simulated SOC by tests with a sinking velocity of 15 m/day is out of the upper measured bound especially in summers. Modeled SOC by the test with a sinking velocity of 1 m/day always yields a SOC below the measured ones.

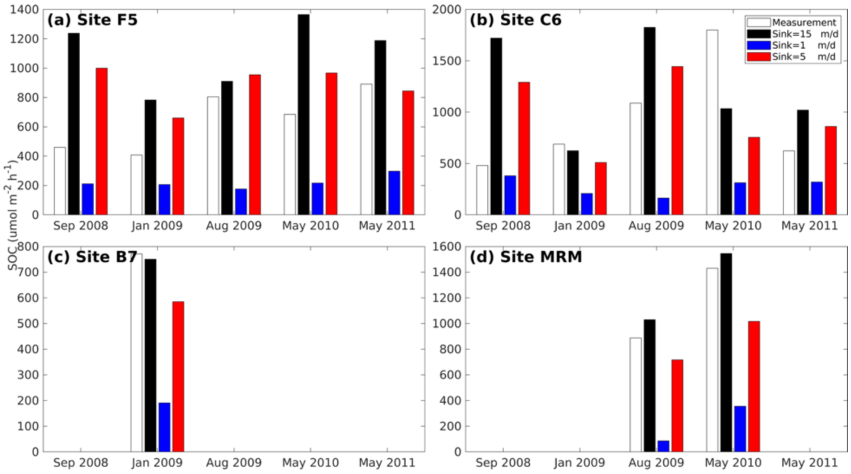

Fig. 4 Comparison of observed SOC (in μmol m^{-2} h^{-1}) by McCarthy et al., (2013) and simulated SOC by different sensitivity tests.

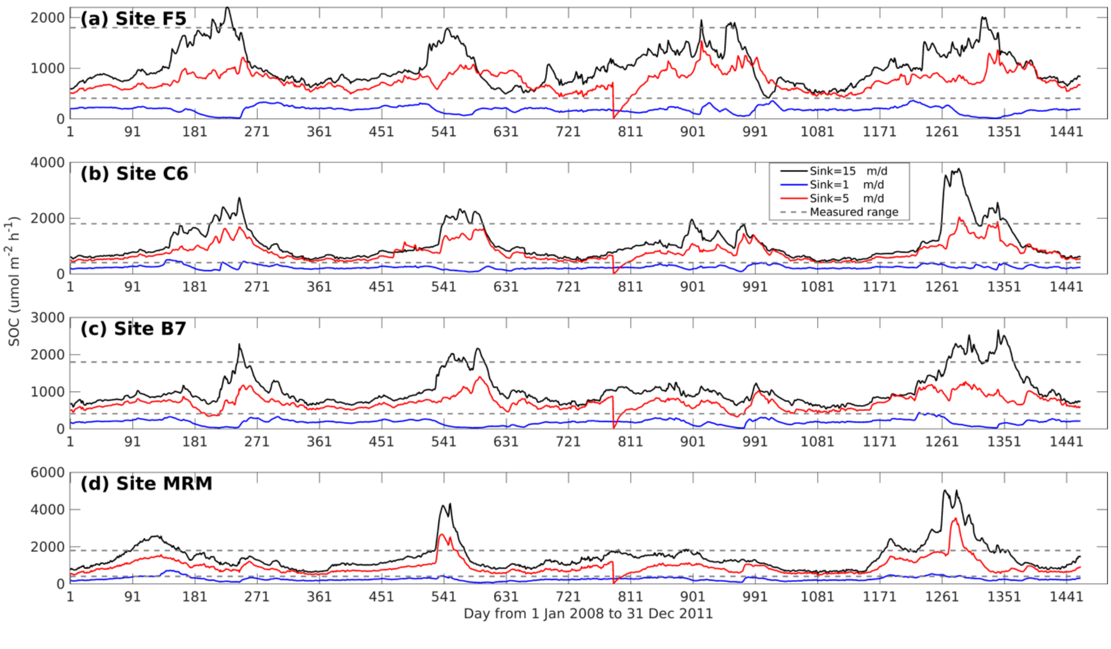

Fig. 5 Daily average of simulated SOC with different PON sinking velocities

We further compared the model-simulated ratio of SOC/overlaying water respiration against that based on available measurements (Fig. 6). The test run with a sinking velocity of 5 m/day shows most agreement with observations with a low averaged RMSE of 4.23 over site F5, C6, and B7, compared with an RMSE value of 4.58 generated by experiment using a sinking velocity of 15 m/day and a value of 6.51 by experiment using a sinking velocity of 1 m/day. At the site near the Mississippi river month (MRM), the two experiments with a sinking velocity of 5 and 15 m/day highly overestimate the ratio observed in August 2009. A possible reason for such bias is that point sources are applied in the model for diverting momentum and concentration tracers from the river to the rest of the computational grid cells. The scheme can lead to an overshot of river water at the near-mouth grid cells, which, may further result in a shorter residence time for organic matter and plankton.

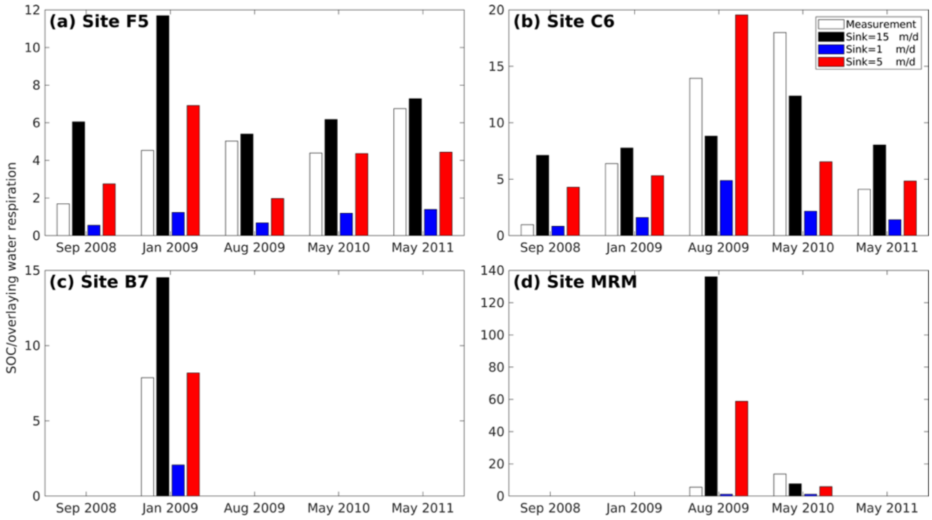

Fig. 6 Comparison of the observed ratio of SOC/overlaying water by McCarthy et al., (2013) and simulated ratio using different settling velocity

Following this comment, we changed the coverage of the model grids used for hypoxic area estimation (Fig. 7). Since the Shelf-wide cruise surveys did not reach the west of 95°W during most of the summers, and the surveys could reach the water with a depth of around 6 m near the Atchafalaya River mouth, we restricted the region from the west side of the Mississippi Delta to 95°W with a water depth ranging from 6 to 50 m. We then compared the model-estimated hypoxic area with different sinking velocities against the Shelf-wide cruise in Fig. 8. Estimations by the two tests with faster sinking velocity (5 and 15 m/day) are close to each other during the cruise periods, while the estimation by the other test (1 m/day) is generally greater than the former two. A sinking velocity of either 5 and 15 m/day can reproduce the magnitude and interannual variability of the measured hypoxic area. Compared to the Shelf-wide observations, the simulated bottom hypoxic extent (Fig. 9–11) by the test run with a 5 m/day sinking rate seems to produce less bias.

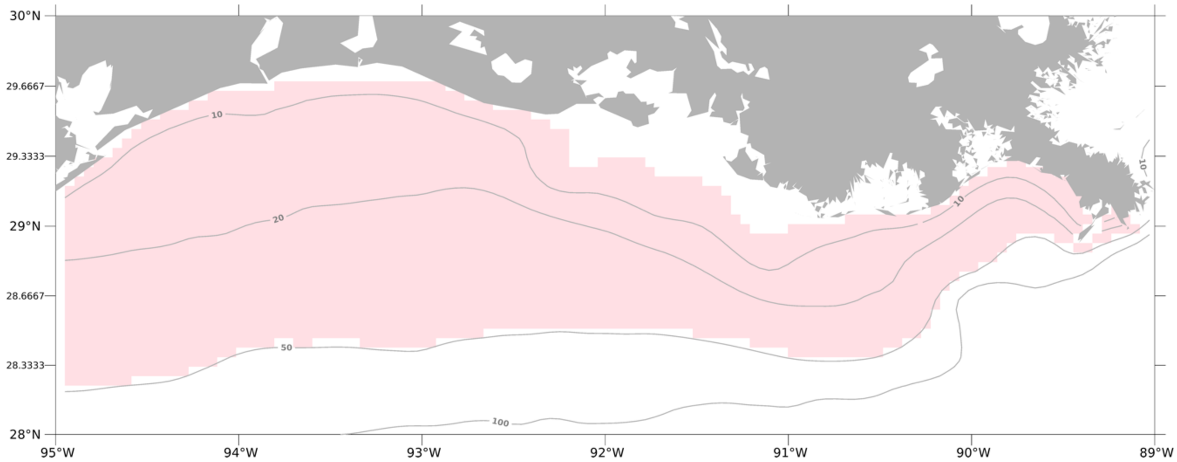

Fig. 7 A distribution of model grids used for hypoxic area estimation.

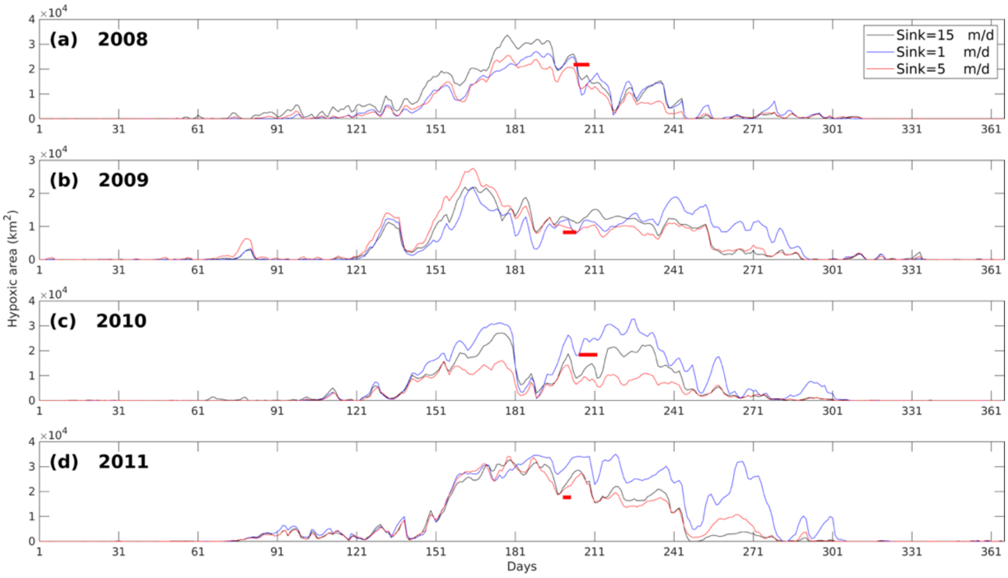

Fig. 8 Comparison of observed and simulated hypoxic area. Note that the horizontal red bars denote the magnitude and temporal coverage of the Shelf-wide cruise measurements.

According to the above comparisons, we will change the sinking velocity of PON from 15 m day^{-1} to 5 m day^{-1} in all experiments and will update the relevant results and discussion.

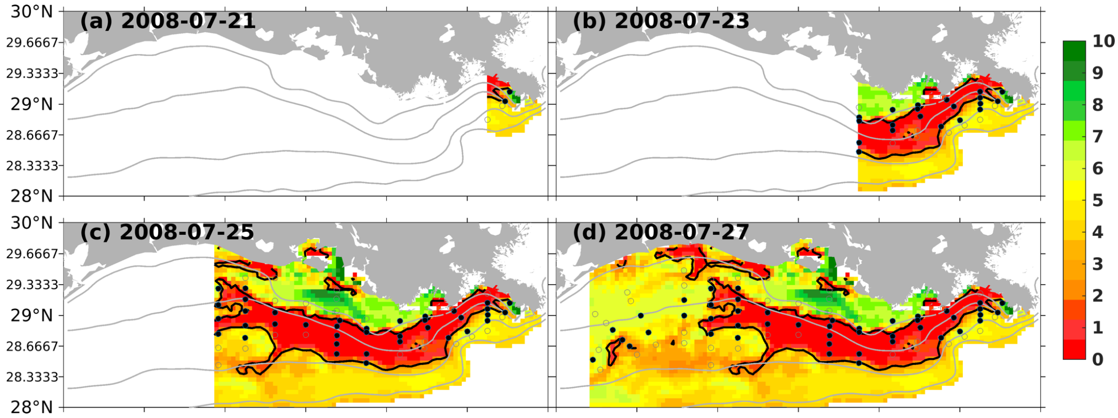

Fig. 9 Evolution of simulated bottom water dissolved oxygen concentration (unit mg l^{-1}) with a sinking velocity of 15 m day^{-1}. The black-filled circles and open circles indicate the hypoxic site and non-hypoxic site, respectively, according to the Shelf-wide cruise observations. The grey curves denote bathymetry of 5, 10, 20, and 50 m.

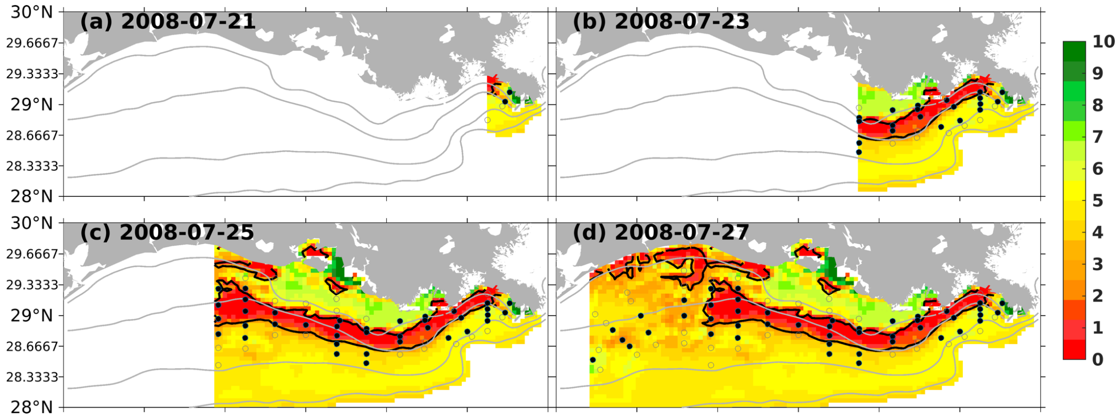

Fig. 10 Same as Fig. 7 but for the test run with a sinking velocity of 1 m day^{-1}.

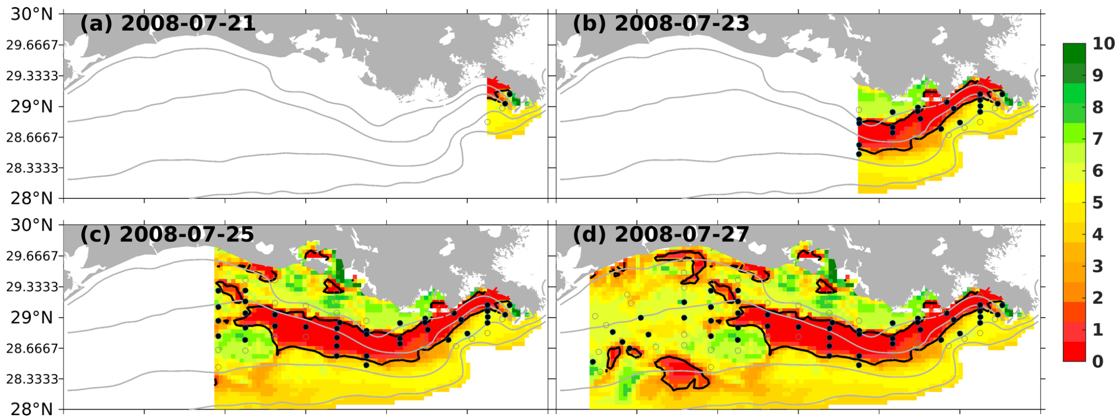

Fig. 11 Same as Fig. 7 but for the test run with a sinking velocity of 5 m day^{-1}.

L211: is 5 months enough for spin-up in this area? What is the initial condition (cold start or hot start)?
Response: We will provide more evidence to demonstrate the model is stable after 2007, e.g., time series of water level, temperature, salinity, PONsed, bottom DO, etc.

Biogeochemical model validations
The entire validation is qualitative rather than quantitative. Need statistic metrics to assess the overall model performance, i.e. taylor and target diagram.
Response: We performed extensive model validation and will add more quantitative validation with more statistic metrics. We cannot agree with Reviewer#2 that our model validation lacks quantitative evaluation. For instance, in Figures 3 and 4, we compared model results against thousands of nutrient and DO profiles and shows good statistics between model and in-situ data.

Figure 3: which cross-section was compared in Figure 2b? The difference histogram in(c)(f)(i) is vertically averaged or bottom value?
Response: The profile comparisons in Figures 3 and 4 are not in a manner of cross-section comparison, instead, we re-arranged the sequence of profiles by their maximum depth. The two histograms are derived based on point-to-point comparisons between simulation and observation of vertical profiles. To make such point-to-point comparisons available, we interpolated the simulated profiles to the observed layers.

L287-288: both NO3 and PO4 were overestimated
Response: We will provide more detailed and quantitative comparisons for the profiles.

L295-296: this statement is a bit questionable that the high riverine nutrient concentration may not be the cause for the model-observation bias. Because the high concentration of PO4 and Si(OH)4 is at the bottom which indicates that it is nutrient regeneration, rather than the allochthonous source.
What are the causes for the hot points (with bottom high nutrient concentration) of PO4 and Si(OH)4?
Response: We did not attribute the nutrient bias to the high riverine nutrient concentration, instead, we want to emphasize that such bias is acceptable if we considered the strong influences of high riverine nutrient loads. We will provide more model-data comparisons for nutrients including surface nutrients in different seasons.

There is a lack of discussion of the higher bottom PO_4 and $Si(OH)_4$ concentrations in our manuscript. We did not include the sediment module in our model, therefore, the high PO_4 and $Si(OH)_4$ concentration in the bottom layers could be a result of the recycled DOP, POP, sedimental organic matter (measured as PONsed), and opal.

L303-304: model overestimates DO while also overestimating the recycled nutrient concentration. Usually, it is the opposite case since nutrient remineralization is associated with oxygen consumption. Any explanations?
Response: The nutrient recycling was overestimated. However, bottom DO variability is more related to sedimental biogeochemical processes. We compared SOC and water column DO balance attributed to water column biochemical processes and found that the impact from the former is much greater than the latter. Please see the above validation for SOC/overlaying water respiration for further justification.

L331-332: in section 3.4 model validation of oxygen, the result suggested that the model

overestimated DO and hypoxia was more frequent in observed WOD profiles. Why here the modeled hypoxia thickness (<=4m) is greater than observed profiles?
Response: We note that the DO profile measurements provided by WOD and Shelf-wide cruise are not always consistent due to different instruments and sampling protocols. The reason we conducted comparisons with SEAMAP and WOD observations was that we could not find any Shelf-wide cruise data prior to 2012 or after 2018 when we prepared the first draft. We now have such "missing" observations available and will conduct the DO profile comparison based on the Shelf-wide observations to assure consistency. Our model slightly underestimated the Shelf-wide observed DO, thus, the modeled hypoxia thickness is typically greater than the observations.

L336-337: the model showed more offshore extension of hypoxia than observation. Any possible causes?
Response: There is no model that can reproduce exactly the extension of the hypoxia area documented by ship-based observations. The model simulated more offshore extension of hypoxia may result from the overestimated offshore transport of water and materiel due to a relatively coarse spatial model resolution (~5 km) and parameterization of advection and diffusion processes. Compared to the existing modeling studies, we are confident that our model performed is pretty robust.

L346-347: the hypoxia area was separated around 92.5W instead of 91W shown in the model simulation? This may reveal a certain defect in the dynamics of model simulation in oxygen.
Response: Indeed, such spatial separation can also be found in the Shelf-wide observations (like in the summer of 2012, 2014, 2015, and 2018).

L349: why not include hypoxia area in the water depth<10m?
Response: The model hypoxic extents did not include regions shallower than 10 m, although the Shelf-wide cruises have some measurements around 6 m. We enlarge the analysis of model grids to 6 m as shown in Fig. 7 in this document and perform validation of hypoxic extent in Fig. 8 for the updated baseline simulations.

L346: the order of figure citation is a bit messy; the figure should be numbered according to the order of citation, not the other way around. For example, the order of Figure 10 is not optimal for reference.
Response: We will review the manuscript carefully and renumber the figure according to the order which the results are shown.

Figure 7: please adjust the x-axis as the other years for better comparison.
Response: We will adjust the x-axis follwoing the comment.

L351-352: this means no apparent bias of model simulation in the hypoxia area. How is this model performance compared to other model studies in this region?
Response: We will add relevant discussion revision.

Results

L432: use biogeochemical instead of biochemical throughout the manuscript
Response: We will correct it accordingly.

L433: denitrification process should not consume oxygen
Response: We will correct it accordingly. But here, we wanted to emphasize that the SOC scheme applied considered the aerobic mineralization, nitrification, and denitrification in the sediment as a 1-step process with the linear assumption applied as we stated in L170–173.

L453-454: what does it mean by saying contributions are limited? I suggest showing the contribution in percentage. What is DO balance and how it was calculated?
L450-457: the entire description and calculation is misleading and confusing. Generally, all DO budget terms including physical terms, photosynthesis, SOC and WCR should be calculated. The summary of budget terms should match the change of DO. I think the authors did not understand and explain the oxygen dynamics well. Please refer to the model studies with oxygen budget analysis and rewrite this part.

Response: Here the DO balance represents the local DO change due to biochemical processes only, i.e., local oxygen sources from phytoplankton photosynthesis (L152–153) and sinks through phytoplankton respiration (L154–155), zooplankton metabolism (L156), aerobic decomposition of PON and DON (L159), and nitrification (L158). We missed the description of aerobic decomposition of PON in Eq. (4) and will correct it in the revision. We plan to perform a more comprehensive comparison among different DO sink and source terms (SOC, water column respiration, local vertical mixing, horizontal advection, and diffusion) and the corresponding contributions to the total rate of change of bottom DO.

L455-456: does the biochemical process in this sentence represent water column respiration?
Response: The biochemical processes here represent the DO change in the bottom water layer due to local oxygen sources from phytoplankton photosynthesis and sinks through phytoplankton respiration, zooplankton metabolism, aerobic decomposition of PON and DON, and nitrification.

L475-476: please indicate the change of PEA quantitively (e.g. in percentage).
Response: We will provide a more quantitative analysis of the PEA (or vertical mixing) in revision.

L480-482: west-Mississippi nearshore did not show a change of current direction from westward to southward, rather it pointed to northward.
Response: The west-Mississippi offshore region was dominated by southward currents in July and August. We will correct this part accordingly.

L498-499: please justify the choice of GBMs method.
Response: We chose GBM method since it provides quantification of the importance of independent variables to the response. Likewise, random forest is another choice. However, the GBM is a better method when dealing with limited independent variables.

L498-511: Move detailed description of GBMs into method section.

Response: We will move it to the Methods section.

Figure 13(a) and Figure 10(a) conflicted in PEA contribution in nearshore West Mississippi?

Response: Figure 10(a) provided a qualitative analysis of the seasonal bottom DO pattern. It shows that the influences of PEA and SOC seemed to be comparable in the west-Mississippi nearshore region as we stated in L490–491. However, it was still hard to quantify their effects using Figure 10(a) alone. We thus conduct a statistical analysis using GBM to quantify the importance of SOC and PEA on the daily variability of bottom DO. The conclusion from Figure 13(a) and Figure 10(a) did not conflict since Figure 10(a) did not distinguish the importance of SOC and PEA while Figure 13(a) did.

L540: what does this statement mean? Please clarify it.

Response: A significant quadratic relationship was found between hypoxic volume and bottom hypoxic area, which showed that the hypoxic water in non-bottom layers can be result of vertical mixing and diffusion. We clarified it in L543–555 showing the influence of the bottom hypoxia condition on the evolution of hypoxia water above the bottom.

L543-544: how does it compare to other model studies? Is this parameterization better or not? Please add a more in-depth discussion here.

Response: We will add more comparisons against other numerical studies. And by far, we found that the fast-sinking velocity setup (15 m day^{-1}) is improper after the validation of SOC and SOC/overlaying water respiration. We present this in this document and have found that the sinking velocity of 5 m day^{-1} provides the most reasonable results. We thus update our baseline simulation with the sinking velocity set as 5 m day^{-1}.

L544-548: Figure 10 and Figure 15 looks very similar which is questionable to me. The previous studies suggested that sediment oxygen consumption dominated the hypoxia in the study area, while the water column respiration was still notable.

Response: According to observation and our updated simulations, SOC is indeed much greater than overlaying water respiration (Fig.2 and Fig. 4). The impacts of upper water column respiration could be important to bottom oxygen balance by advection and diffusion processes. We will add more analysis for all the oxygen budget terms and perform comparisons of their impacts on the bottom DO dynamics.