Responses to Comments by Referee #1

General responses:

We sincerely thank the reviewer who provided us with such detailed comments and suggestions and also his/her patience with our prolonged responses.

We reorganized our thoughts for this study and would like to focus on 1) the role of silicate cycling in the biogeochemical processes and bottom hypoxia development in the Louisiana-Texas shelf and 2) the impacts of a complex plankton community on the dissolved oxygen (DO) dynamics. Thus, we removed the original sections 4.1 and 4.2, which discussed the impacts of physics. Instead, we focused on the biogeochemical processes in DO dynamics.

We also separated the Results and Discussion sections. In the Results sections, we tried to answer the following questions. 1) What are the limited nutrients for PS and PL, respectively? 2) What is the dominant plankton group on the shelf? 3) How do different plankton groups contribute to the source and sink of DO in water columns and sediments? In the Discussion section, we reran all sensitivity tests and expanded the 3-year (2018–2020) simulations to a 9-year (2012–2020) one. We aimed to discuss the responses of biomass and DO to the reduced riverine nutrient supplies. Please find our detailed point-to-point as follows,

General comments:

Overall: This revision of Ou and Xue manuscript attempt to address some of the issues the original version that were pointed out by the reviewers. Indeed the authors provide more context in their introduction to justify their study as well as more validation that tend to demonstrate that the model agrees well with observations. The analysis is also better displayed. The use of Si limitation is also interesting.

That said, most of the analyses still repeat previous work (oxygen budget, effect of stratification, importance of SOC) and indeed the conclusions are similar. Apart from the model itself, Section 4.3. (nutrient load experiments including Si) is arguably the only novel part of the paper; I don't think Si has been included in such a way in nutrient management strategies/studies, although the potential for Si limitation has been discussed in various studies. That raises several questions that are critical to this investigation but that have not been or barely discussed.

Responses: Sections 4.1 and 4.2 have been removed. We now focus on the contribution of different nutrients (N, P, Si) and the effects of a complex plankton community to hypoxia development.

1) why was Si not included in previous models? the authors mention briefly in the Discussion section that previous work assume that Si is plenty and therefore not limiting. This is true but the authors do not provide strong evidence against these assumptions. Previous assumptions were based on observations. Also observations indicate that N (TN or NO3) is the main predictor for the mid-summer hypoxic area, which suggest that variations in N load control hypoxia.

Responses: Previous statistical models suggested that there is a strong correlation between the reduction in riverine nitrogen and reductions in hypoxic areas. However, there have been several observational studies indicating the importance of silicate in the study area. In this study we also found significant correlations among riverine nitrogen loads, phosphorus loads, and silicon loads (see figure below and table in Appendix C). Strong correlations can also be found between hypoxic area (or bottom DO) and phosphorus and silicon loads. In this sense, previous statistical studies might overestimate the effects of nitrogen. We added related discussions to the revised Discussion section.

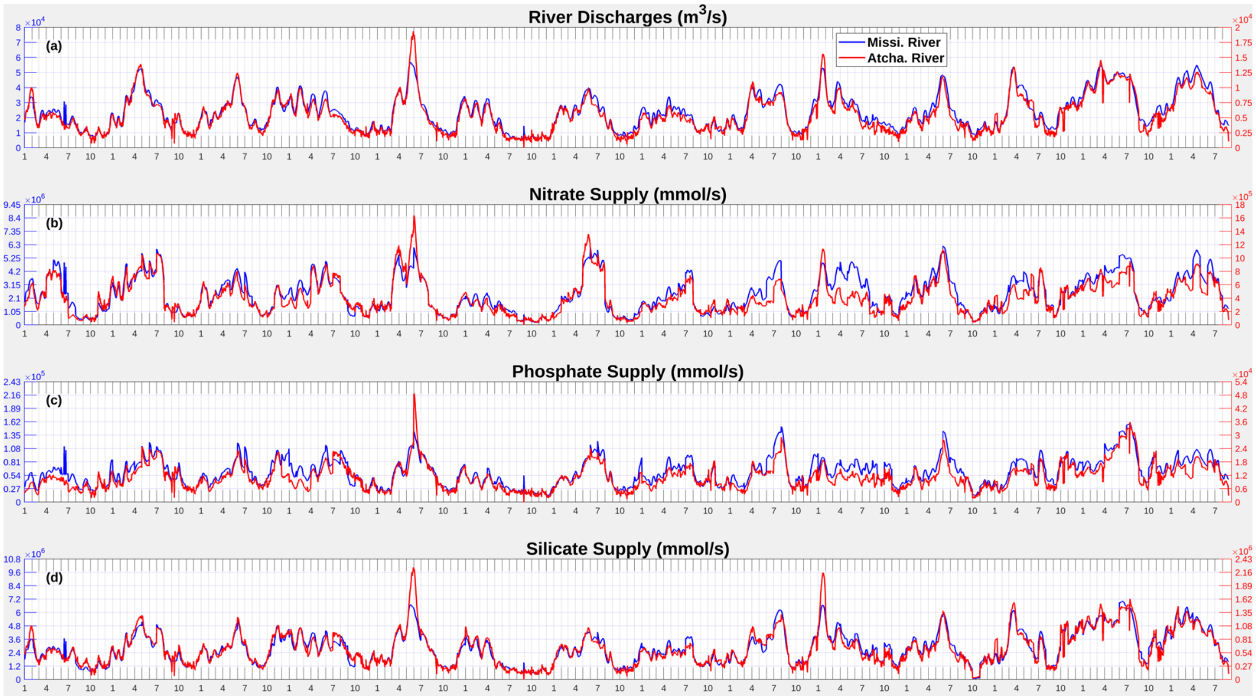

Figure C1. Daily time series (2007–2020) of river discharges of freshwater, nitrate, phosphate, and silicate from the Mississippi and Atchafalaya rivers.

Table C1. A correlation matrix of daily inorganic nutrient loads by the Mississippi River and the Atchafalaya River from 2007 to 2020. Correlation coefficients shown are all significant (*p*<0.001).

	Mississippi nitrate+nitrite	Atchafalaya nitrate+nitrite	Mississippi phosphate	Atchafalaya phosphate	Mississippi silicate	Atchafalaya silicate
Mississippi nitrate+nitrite	1					
Atchafalaya nitrate+nitrite	0.9123	1				
Mississippi phosphate	0.8328	0.7577	1			
Atchafalaya phosphate	0.7517	0.7913	0.9155	1		
Mississippi silicate	0.8583	0.7795	0.8759	0.7942	1	
Atchafalaya silicate	0.7938	0.7956	0.8131	0.8148	0.9520	1

2) Is this useful to include Si in an experiment to assess the effect of nutrient reduction strategies? can you provide examples of how this can be implemented? It might be more useful to focus on N and P reductions as in other studies (assuming constant Si, or extrapolating on future Si) and look in detail at the effects on the ecosystem. Mentioning that there are nonlinear effects to nutrient reductions is not enough.

Responses: In the Discussion section, we carried out and analyzed the results of six sensitivity tests to cover the effects of different nutrient reduction strategies on hypoxia reduction. Our model shows that N is not always the only limited nutrient. P and Si limitations can be dominant in waters that are deeper than 20m. Please see the Discussion section for more details.

3) the effect of nutrient management on phytoplankton community structure, trophic interactions, and ultimately organic matter deposition and hypoxia is very interesting. This should be the main focus of this manuscript instead of redoing previous work with a different model. For information, earlier models were validated against (bulk) phytoplankton biomass. Since section 4.3 this is an add on section at the end of the manuscript, these effects are only vaguely described. A thorough analysis of these effect would improve the manuscript significantly and make an important contribution.

Responses: We strongly agree with this comment, and we thank the reviewer for helping us to point this out. We have shifted the focus of this study to the potential silicate limitation and the effects of a complex plankton community on hypoxia and DO dynamics.

4) Are the simulated effects real? This is hard to believe that an 80% N load reduction will result in a 25% increase in hypoxia. By which mechanism? Si/P transport downstream? but then what is the source of NO3? Could you describe, through schematics the effects of the different load reductions? The model will always give results but the readers need to be convinced that those are realistic. Currently it makes me wonder if the biological model has been properly parameterized. How was the parameterization done after modifying the structure of the model? Using a predatory zooplankton without a proper parameterization may result in a top-down control of the system, which then lead to artificial nonlinearities in the response to decreasing nutrient loads. Zooplankton is often unconstrained due to the lack of observations, which may be why previous modelling studies used a more parsimonious approach to the model structure.

Responses: Most of the parameters of our model followed previous model studies of Laurent et al. (2012) for phosphate-related parameterization, Fennel et al. (2006) for parameterization of light inhabitation on nitrification, Shropshire et al. (2020) for the rest.

A linear function of mortality was applied for PS, PL, ZS, and ZL, while quadratic mortality was used for ZP, accounting for the predation pressure of unmodeled predators, like planktivorous fish. We can see both bottom-up and top-down effects in the biomass responses to the different nutrient reduction strategies. Please see the updated Discussion section for more details.

As for the counterintuitive responses of hypoxia to the reduced N supply, we updated our discussion in the manuscript. We reran our sensitivity tests and focused on the riverine nutrient reductions of 60% only. Results indicate that 60% of N load reduction would lead to an increase in the bottom hypoxic area. We attributed such a counterintuitive response to the dominated limited nutrient and the maintenance of positive DO contribution (net production) by the plankton community at layers within the bottom 2 m.

The N is usually limited for the growth of PS and PL, mostly in the shallow middle and west shelf (10 – 20 m) during summers, while P and Si limitations are more commonly simulated by our model. The reduction in N supply only would lead to a slight decrease in photosynthesis. DO contribution by the plankton community maintains positive at both the upper (surface to 2 m above the bottom) and bottom water columns (layers 2 m above the bottom) as in the control run. As total production in both upper and bottom layers would decrease with less N supported, less DO production in the water column could be found. Although SOC would decrease due to less PON_{sed}, the total effects of the changes in the three DO source/sink terms are likely to result in a decrease in the bottom DO and a slight increase in the bottom hypoxic area. A detailed discussion has been added in the Discussion section.

A less important but redundant issue is the use of "means" (or ratios) for validation and analysis. This is sometimes problematic because mean values, shelf wide for instance, are often not representative of the dynamics of the system.

Responses: In the Validation section, we added plots of 1) nutrient concentration bias against distance to the Mississippi River mouth, 2) DO concentration profiles averaged for different ranges of depth.

As for the SOC validation, in McCarthy et al.'s (2013) study, cruise periods were only listed by months, and thus, we averaged the model results over the corresponding month and performed a model-data comparison. In the comparison of bottom hypoxic waters, the modeled results are not averaged outputs; instead, they are a composite of different bottom DO snapshots corresponding to the cruise date. In the Results and Discussion parts, in addition to the mean value, we added model medians, minimum, maximum, and quantiles for a better demonstration.

Specific comments:

L134-135: you could mention Hetland and Dimarco parameterization as well.
Responses: parameters added.

L258-259: I don't understand this statement
Responses: This part has been removed from the revised manuscript

L268-270: you may want to reformulate this, you should say that complexity may be a factor instead of saying that all previous studies were wrong. So far there you did not provide evidence that this could be the case. Can you do that?

Responses: We have reformulated this sentence.

L349-351: this is not novel so you should provide evidence of why you think using more plankton groups would change previous findings. Otherwise you are only repeating previous work.

Responses: We have rewritten this part and restated the focus and novelty of this study.

L351-352: You should focus on your study here and mention that later in the discussion/conclusions
Responses: We have removed this sentence.

L414-418: was this model re-parameterized? if you change the structure of the model you will need to change your parameter set accordingly, e.g. manually or though optimization.

Responses: The parameterizations of our model largely followed previous existing studies, i.e., Laurent et al. (2012) for addon phosphate parameterization, Fennel et al. (2006) for parameterization of light inhabitation on nitrification, Shropshire et al. (2020) for the rest. We have validated our nutrient, SOC, and DO concentration in the validation section. Comparisons indicated that our model provided an improvement in nutrient and DO simulations than previous model studies. Previous studies barely validated nutrient profiles. Our model-data comparison for DO profiles is better than that of existing studies. Please see the updated validation section for more detailed descriptions.

L428-429: river P does not follow the Redfield ratio
Responses: The Redfield ratio was only applied to fill the missing riverine measurements. As Fig. 8 shows (also attached here), riverine N:P did not follow the Redfield ratio indicating that the missing P measurements are rare for the Mississippi and Atchafalaya rivers.

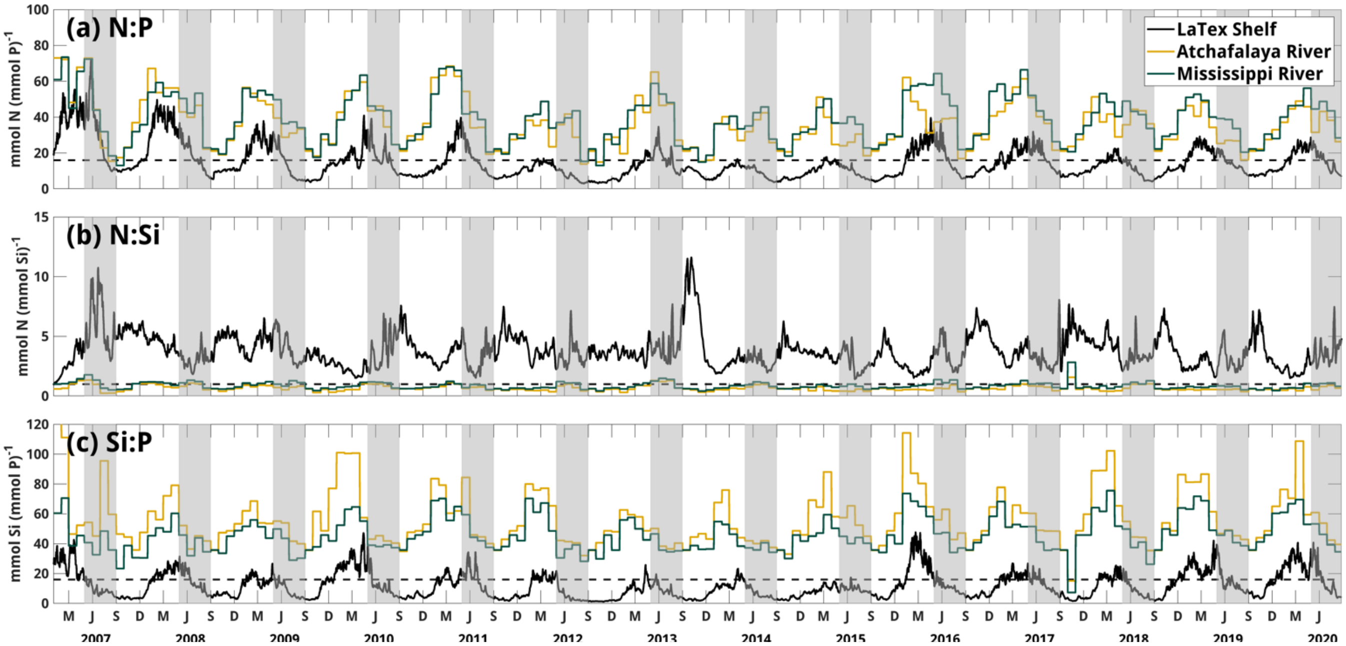

Figure 8. Daily time series of ratios of nutrient loads from the Mississippi and Atchafalaya Rivers and nutrient ratios averaged over the LaTex shelf (Fig. 2b) from the numerical results. Note that the latter ratios are derived based on the depth-integrated nutrient concentrations (in mmol m^{-2}). The black dashed lines denote the nutrient ratios of 16:1, 1:1, and 16:1 in (a), (b), and (c), respectively. The gray patches indicate the late spring and summer (May–August) period of each year. The capitalized letters of M, J, S, and D in the x-axes denote the first day of March, June, September, and December, respectively.

L522-523: the way it is presented is confusing, you should say that it is coupled nitrification-denitrification, as in the Fennel et al model, which implies that NO3 produced in the sediment is used for denitrification

Responses: The corresponding descriptions have been simplified.

L564: what is the horizontal resolution in km?

Responses: The horizontal resolution is about 5 km and has been addressed in the first sentence of this paragraph.

L683: you could show that with biomass data

Responses: The high diatom productivity could be found in Table 1. Instead of pasting the data here, we added a reference to Table 1.

L689: you are mixing vertical and horizontal locations, a 35m peak can either be located at the bottom in 35m waters or in the subsurface in deeper waters. Similarly, a 15m peak can indicate deep observations in nearshore waters or subsurface conditions in deep waters

Responses: We have removed this confusing statement as it would not provide evidence for our analysis. Here we aimed to provide validation for nutrient profiles rather than in-depth analysis.. Further, probability histograms were replaced by profile biases against distance to the Mississippi River mouth.

L691: can you provide the average bias?

Responses: Profile biases against distance to the Mississippi River mouth were added to Fig. 3.

L775-778: what is the purpose of these numbers? Are these values varying seasonally?
Responses: The probability histograms were replaced by profile biases against distance to the Mississippi River mouth. Thus, we removed this part but added a corresponding description for the average biases.

L779: It depends where and when these differences occur. If it is close to the river source in Spring then yes, but if it is farther downstream in summer then this is a significant difference.
Responses: Profile biases against distance to the Mississippi River mouth show observation biases are high only near the river mouths (<70 km for the Mississippi River mouth, ~250 km for the Atchafalaya River mouth). At other sites away for the mouths, biases are much lower. It supports that "The nutrient concentrations bias between simulations and observations is acceptable concerning the strong influences of high riverine nutrient loads on the shelf".

Figure 3: are these averages for the area in Figure 2b? this should be mentioned. Also why averaging over this very large region with very heterogeneous conditions?
Responses: The averages were not for the area in Figure 2b, instead, we extracted the modeled nutrient profiles at the locations shown as blue crossings (WOD-derived measurements) in Figure 2c on the date of the measurements and performed averages of these modeled profiles.

L787: see previous comment
Responses: The modeled averages were performed for the inner shelf and midshelf, respectively, according to the Fig. C2 in Appendix C. In Schaeffer et al. (2012) and Chakraborty and Lohrenz (2015), locations of sample sites were not provided (Figure 1 in both studies). We then restricted the lon/lat range of the inner shelf and mid-shelf according to the figures shown in their studies.

Table 1: can you add columns with biomass?
Responses: Schaeffer et al. (2012) provide biovolume, and Chakraborty and Lohrenz (2015) provided chlorophyll *a* of different plankton groups. Neither study provided plankton biomass data.

Section 3.4, Figure 4: SOC observations are often heterogeneous and it is not expected that your model exactly match observations at a particular location given the simple representation of the sediment. However, it would be useful to show how your sediment layer behave (since it is a new addition) with time series of PON accumulation/respiration and SOC at several representative locations of the shelf. In comparison you can compare with the measurements of McCarthy et al and others (Murrel and Lehrter?, e.g. 10.1007/s12237-010-9351-9).

Responses: We strongly agree with this comment. Model-data comparison suggested that our model can capture well the SOC magnitude. We have removed the comparison of SOC/overlaying water respiration. In McCarthy et al.'s (2013), the overlaying water was the layer ~20 cm above the bottom. In our model, as a sigma vertical coordinate system was used with 36 vertical layers designed, the thickness of the bottom layer is usually ~ 1 m. We have added a time series of integrated sedimentary PON (Fig. 11 in the updated manuscript; also attached here) over the LaTex shelf to illustrate the behavior of the PON accumulation.

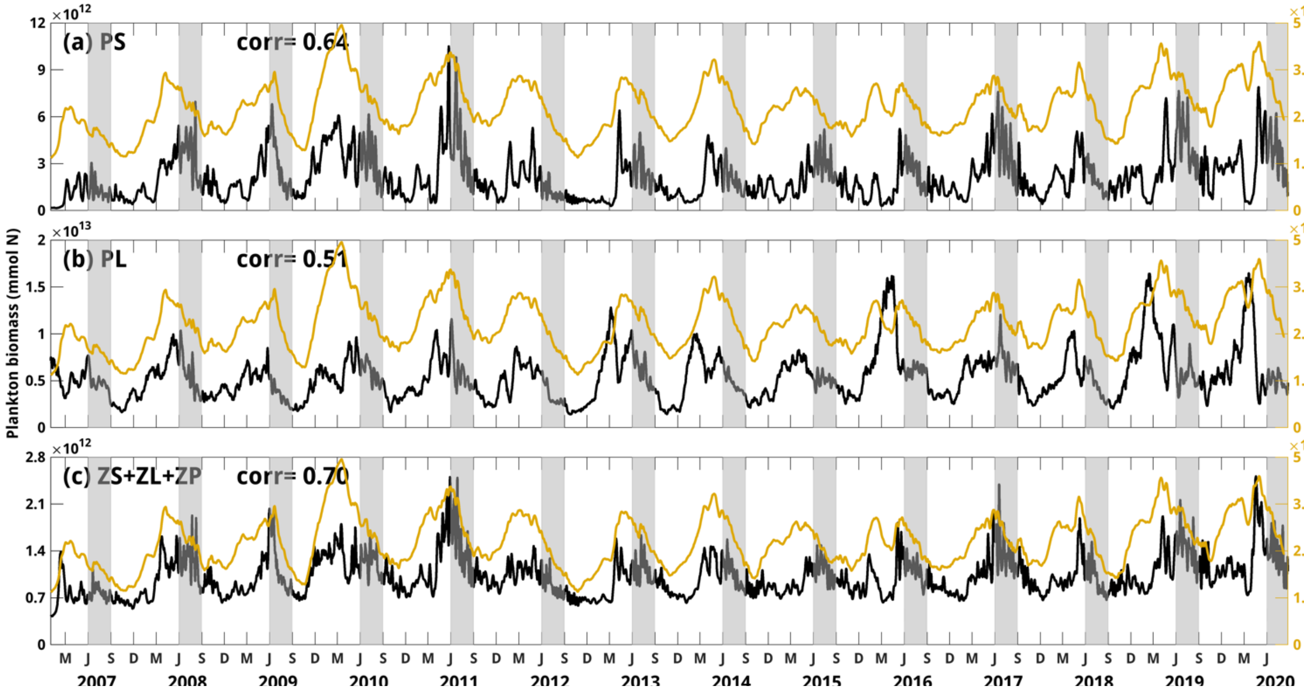

Figure 11. Comparisons between daily PON_{sed} and plankton biomass (i.e., (a) PS, (b) PL, and (c) secondary production). All biomass matrices were integrated over the entire water column and over the LaTex shelf.

Figure 5: it would be better to show a comparison of water column respiration instead
Responses: We did not provide a comparison of water column respiration here. On one hand, in the McCarthy et al.'s (2013), only samples at 4 layers of the water columns were collected (i.e., surface, middle, bottom, and overlaying). In contrast, there were 36 vertical layers designed in our model. Thus, there should be a great bias between the estimated depth-integrated respiration provided by the McCarthy et al. (2013) and our estimates. On the other hand, as the literature did not provide exact sample dates, biases are introduced between monthly averages of simulations and the cruise measurements. The model is not expected to capture well with the depth-integrated water respiration.

Figure 6. comparing average profiles of DO doesn't make much sense at the scale of the shelf. The envelope, which represents actual observations) indicates here that hypoxia tends to be found in shallower waters in the model with somewhat more severe conditions. The large differences in panel (d) is probably the result of this mismatch in space.
Responses: We updated this figure (also attached here), showing averaged profiles against the normalized depth for different depth ranges. Despite some overestimations (~ 1 mg L^{-1}) of DO profiles, our model results, in general, provided similar and even better performance than previous numerical studies. For example, DO concentration biases given by Yu et al. (2015) were within 2 mg L^{-1}. And to our best knowledge, no existing study has ever tried to provide one-to-one comparison between model-simulated DO profiles and observed ones.

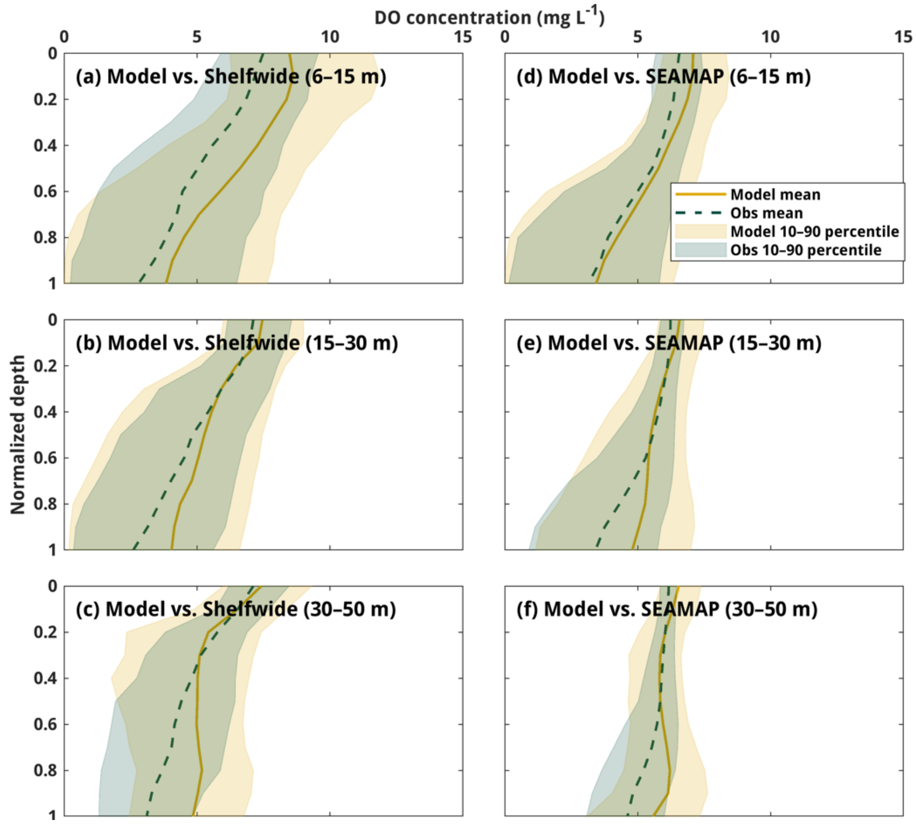

Figure 5. Comparisons of DO profiles between model hindcasts and measurements by (a–c) NOAA's shelf-wide cruises, and (d–f) SEAMAP. The normalized depths of 0 and 1 represent surface and bottom, respectively.

Section 3.6: can you also provide a mid cruise comparison in the appendix/supporting material? Can you also mention in the Methods (may be I missed it) how is the bottom sediment layer was initialized?

Responses: The model results shown in Figure 6 is not averaged results during the cruise periods. Instead, they were composites of different DO snapshots spanning over the cruise period. The cruises were always conducted from the east to the west, we thus "sampled" the modeled snapshots from the east to the west following the cruise dates. For example, on day 1, if the cruise reached 91W, model DO over the east of 91W was "sampled" as the 1st snapshot and was added to the composite first. On day 2, as the cruise reached 92W, model DO between 91W to 92W was "sampled" as the 2nd snapshot and was added to the composite, and so forth. We added relevant descriptions in the first paragraph of section 3.6.

The sedimentary PON, burial PON, sedimentary Opal, and burial Opal were initialized as 0.1 mmol m^{-3}. We have added it in the Method (Line 236).

Section 4.1. This belong in the Results section

Responses: We have removed this section.

L1052-1063: The way it is formulated it sounds novel but this type of regional budget has been done previously, you should mention that you are doing the same type of budget but with your new model.

Responses: We have removed this content.

L1065-1066: It make sense that water column biogeochemistry does not have a significant effect because you look at the bottom DO and therefore you only include respiration within this (thin) layer. However water column BG influence the entire water column and is relevant to bottom DO in the subsurface layer. At this scale you may find that water column BG is as important as SOC for bottom DO.
Responses: Our new results indeed showed that the biochemical processes in the water column (both bottom 2 m and layers above) are as important as SOC for bottom DO. Please see the updated discussion section.

L1122-1124: You point out the issue with this analysis (i.e. previous comment), it would have been more relevant to look at the bottom (lower 5 or 10m) or subsurface layer. Vertical diffusion might have shown a seasonal cycle then. Also, looking at the bottom layer only, you artificially increase the contribution of SOC on DO.

Responses: In the updated version, we focused on the depth-averaged DO concentration at layers within the bottom 2 m rather than the DO at bottom layer. We also removed the analysis of this part.

L1160-1161: isn't this expected?
Responses: We have removed this part.

1163-1459: not clear. It is difficult to interpret such pattern with shelf averages
Responses: We have removed this part.

Section 4.3: 3 years simulation including 2 years (2018, 2020) with mid summer wind events. Is this enough? What were the nutrient loads, river discharge during these years?
did you initialize your model from your long run?
you didn't show the dynamics of the sediment PON pool, that would be an interesting addition
Responses: We extended the length of simulations to 9 years (2012-2020). The river water discharges were kept the same as those in the control run. The riverine nutrient concentrations were reduced by 60% with different reduction strategies (i.e., -60%N, -60%P, -60%Si, -60%(N+P), -60%(N+Si), and -60%(N+P+Si)). Thus, the nutrient loads were decreased with the shelf hydrodynamic unchanged. The sensitivity runs were initialized from the long-term (15-year) control run. We have added relevant discussion about the PON_{sed} pool to the revised manuscript.

L1495: here you say that this is a mean for mid summer hypoxia and the below it becomes a multi year summer mean, which is very different. The good metric to show the effect of nutrient reduction is to take either a season mean or seasonally integrated hypoxia (the time integral of the hypoxic area)
Responses: We have updated this figure. Previous comparisons were all based on the mid-summer (during the shelf-wide cruise) mean matrices. In the updated version, we focused on the statistic matrices conducted for the May-August period of each year.

Figure 11: % change from what value? in Figure 8 you show that mid summer hypoxia is very small in 2018 and 2020(see general comments) These results are hard to believe. You need to be

more convincing. How do you explain that N80, P80 result in an increase of the diatoms, which are the dominant phytoplankton in the river plume. What is their source of nitrogen? River N load has been shown to be well correlated with mid-summer hypoxia. Your results suggest that Si load is the best predictor for hypoxia.

Responses: We have updated this figure. New results indicate that diatoms would not change much under the N60 scenario. But PS would experience a slight decrease. It was due to that the shallow parts of the mid and west shelf were usually limited by N in summer. However, in other parts of the shelf, P and Si limitations were more common. Previous studies showed that riverine N loads are well correlated with mid-summer hypoxia. However, riverine P and Si loads are both highly correlated to riverine N loads (see below table), which indicates that riverine P and Si loads are both well correlated with the mid-summer hypoxia as well. Here we would like to investigate, by using a mechanistic model rather than a statistical one, whether the introduction of the Si cycle and a more complex plankton community would provide different results from previous model studies. We agree that the Si limitation in the LaTex shelf was rarely shown in published studies. Yet our new model results indicate that it is worth conducting more data collection, including the possible contribution from Si to hypoxia development.

Table C1. A correlation matrix of daily inorganic nutrient loads by the Mississippi River and the Atchafalaya River from 2007 to 2020. Correlation coefficients shown are all significant (*p*<0.001).

	Mississippi nitrate+nitrite	Atchafalaya nitrate+nitrite	Mississippi phosphate	Atchafalaya phosphate	Mississippi silicate	Atchafalaya silicate
Mississippi nitrate+nitrite	1					
Atchafalaya nitrate+nitrite	0.9123	1				
Mississippi phosphate	0.8328	0.7577	1			
Atchafalaya phosphate	0.7517	0.7913	0.9155	1		
Mississippi silicate	0.8583	0.7795	0.8759	0.7942	1	
Atchafalaya silicate	0.7938	0.7956	0.8131	0.8148	0.9520	1

L1556-1558: you talk about nonlinear response but you don't provide the mechanisms. please explain the mechanisms, diagrams would be useful to support these explanations
Responses: In the updated manuscript, we focus on the explanations of the nonlinear responses, starting from nutrients to biomass and then to DO dynamics.

L1572: what about phytoplankton biomass?
Responses: Please see our response to comments for Table 1.

L1640-1648: This is the novel part of the study but I have a hard time to believe these results, they do not make sense ecologically. The authors should show the mechanisms, nonlinearities that explain their results.

Responses: Please see our response to general comments 4). We have provided related discussion in the revised Discussion section.

L1650: does that follow the BG policy or should the results be available from a repository?
Responses: We will deposit our model data in a public repository.

Minor comments/typos:

L258: Fennel et al is cited twice (later in the Discussion as well)
Responses: Have corrected.

L347: typo, 3->2 and 2->3
Responses: Have corrected.

L352-354: this is not necessary
Responses: We have removed this sentence.

Figure 7: can you show 2020 as well?
Responses: We could not find the source data of 2020 shelf-wide cruise.

L1506: for mid summer or for whole summer?
Responses: Previous results were based on the mid summer statistics. In the updated version, we focused on the statistic matrices conducted for May-August period of each year.

L1144: sediment biogeochemistry
Responses: Have corrected.

Figure 12: your color scale is a bit counter intuitive
Responses: We have removed this figure.

Responses to Comments by Referee #2

General responses (we also included this in our response to Referee#1):

We sincerely thank the reviewer for providing us with such detailed comments and suggestions and also his/her patients for our prolonged responses as we carried out a series of new experiments and reshaped the article.

We reorganized our thoughts for this study and would like to focus on 1) the role of silicate cycling in the biogeochemical processes and bottom hypoxia development in the Louisiana-Texas shelf and 2) the impacts of a complex plankton community on the dissolved oxygen (DO) dynamics. Thus, we removed the original sections 4.1 and 4.2, which discussed the impacts of physics. Instead, we focused on the biogeochemical processes in DO dynamics.

We also separated the Results and Discussion sections. In the Results sections, we tried to answer the following questions. 1) What are the limited nutrients for PS and PL, respectively? 2) What is the dominant plankton group on the shelf? 3) How do different plankton groups contribute to the source and sink of DO in water columns and sediments? In the Discussion section, we reran all sensitivity tests and expanded the 3-year (2018–2020) simulations to a 9-year (2012–2020) one. We aimed to discuss the responses of biomass and DO to the reduced riverine nutrient supplies. Please find our detailed point-to-point reponse to your comments as follows,

This version of manuscript by Ou et al. improved the previous version of manuscript somewhat. The reviewer appreciated the modification on the bottom DO budget analysis and additional sensitivity runs on the sensitivity runs to evaluate the model's robustness regarding different parameterizations. The contributions of different biogeochemical and hydrodynamic processes on bottom DO, the bottom DO's response to the reduction of riverine nutrient loads, and the impact of diatom on hypoxia dynamics were better analyzed. Although I see improvement of the manuscript, there are still some issues that needs to be addressed, which I think is significant and necessary. In particular, this paper missed real discussion section including the comparison with previous work, and advanced understanding on the topic. In other words, although much work has been done for this research, this paper is still organized like a technical report of result, lack of the understanding of both physical and biogeochemical mechanisms. My major comments and concerns are listed below.

Responses: We have separated the Results and Discussion sections. We added comparisons with previous studies in the Discussion section (please refer to the revised manuscript).

(1) "Compared with existing modeling efforts, our model, for the first time, included a silicate cycle as well as multiple plankton functional groups in the modified biogeochemical model, the importance of which has already been addressed in previous studies yet not included in hypoxia modeling efforts." As the authors indicated in the reply, the highlight of this paper should be adding the silicate cycle as well as multiple plankton functional groups contribute to the hypoxia simulation and understanding in this area. However, although the authors validated the new model well with nutrient profile, SOC and diatom percentage, the bottom DO as well as subsurface DO was significantly overestimated (Figure 6), and the hypoxia area was underestimated correspondingly. Moreover, the interannual variability of SOC was actually not well captured by the model (shown by Site F5 and C6). Therefore, there might be systematic defect with the model, either hydrodynamics or biogeochemical processes, that need to be at least thoroughly discussed.

Responses: We change the focus of this study to the potential contribution of Si cycling in hypoxia development and the effects of a complex plankton community on DO dynamics.

SOC observations are often heterogeneous. Model-data comparisons suggested that our model can, in general, capture the SOC magnitude. As the model results were monthly averages, it would be hard for the model to fit perfectly with the measurements.

We updated Figure 6 and reordered it to Figure 5 (see below), showing averaged profiles against the normalized depth for different depth ranges. Despite the overestimations (~ 1 mg L^{-1}) of DO profiles, our model results provided similar and even better performance than previous numerical studies. The DO concentration biases in Yu et al. (2015) were within 2 mg L^{-1}. And the bias of our model outputs is less than 1 mg/ L^{-1}, which gives us confidence about the model performance.

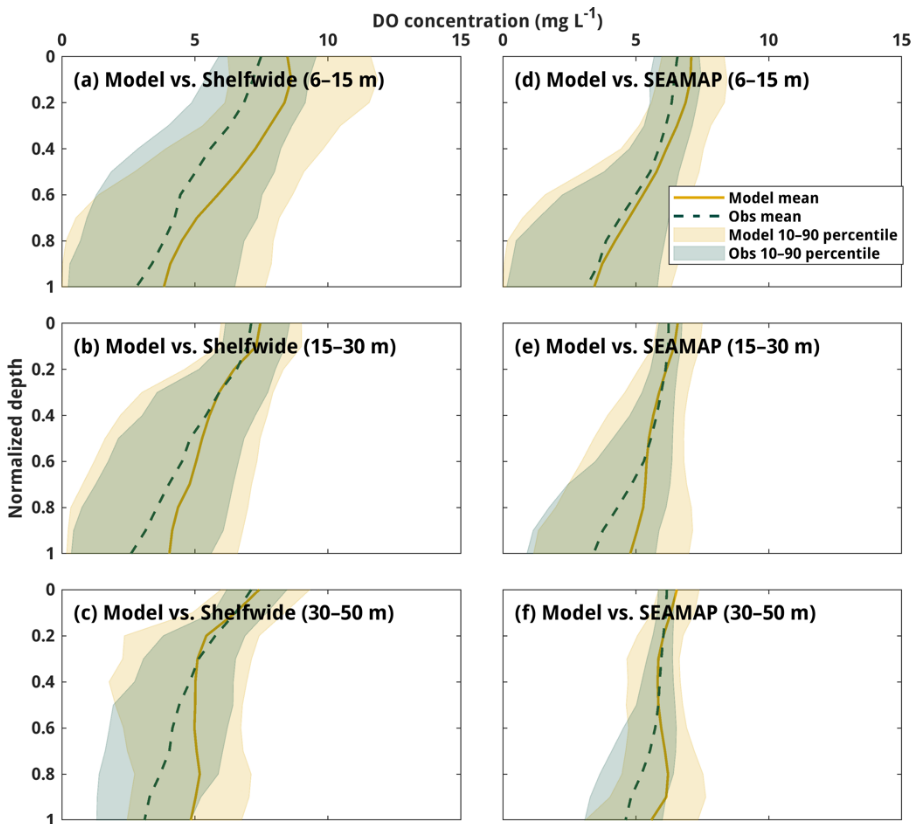

Figure 5. Comparisons of DO profiles between model hindcasts and measurements by (a–c) NOAA's shelf-wide cruises, and (d–f) SEAMAP. The normalized depths of 0 and 1 represent surface and bottom, respectively.

(2) The authors applied oxygen budget analysis on the bottom DO, tried to discern the major contributors for bottom DO. This part of manuscript did not make much contribution to the study and what have been shown were mostly covered by previous studies, i.e. SOC was the dominant term especially in the nearshore area. Also, the physical mechanisms between stratification and advection terms were not explained well. Therefore, the reviewer suggests removing or shorten this part and focus on the nutrient and lower tropic community impact on bottom DO.

Responses: The sections 4.1 and 4.2 have been removed following the reviewer's suggestion.

(3) Following my previous comments, this paper still lacked comprehensive discussion and

comparison with previous modeling study on the model performance, simulation results and conclusions, specify and address what were the agreement, what were the disagreement, and what was new. The authors claimed that the bias of overestimation on DO was acceptable, however, the quantitative comparison should be provided with other model works in this area, like RMSE, bias, correlation coefficient, etc., to prove the improvement of this new model. Even if the new model underperformed compared to previous models, discussions on the potential causes were needed.
Responses: We separated the Results and Discussion sections and added the comparison between our model results and previous simulations in the Validation section (also see responses above).

Detailed comments:
Abstract
Modify "biochemical" to "biogeochemical" in the title and throughout the entire manuscript. The biogeochemical cycle links the living biomass in the water column to the sediment. Biochemical is only part of the processes discussed and less used for the topic.
Responses: Have corrected. But we would like to keep "biochemical" in our title as this is the PartI of our duo-paper, the PartII paper about applying Machine Learning in hypoxia prediction is published in 2022.
Ou, Y., Li, B., & Xue, Z. G. (2022). Hydrodynamic and biochemical impacts on the development of hypoxia in the Louisiana--Texas shelf -- Part 2: statistical modeling and hypoxia prediction. *Biogeosciences*, *19*(15), 3575–3593. https://doi.org/10.5194/bg-19-3575-2022

L13-15. 16-17: that's basically all the terms… which are more important in the physical terms and why in the aspects of hydrodynamics?
Responses: The sections 4.1 and 4.2 have been removed. The Abstract has been updated accordingly.

L20: add period after the bracket
Responses: corrected.

L20: how about water column and sediment nutrient recycling associated with change in oxygen condition? (Kemp and Testa et al, 2012)
Responses: In this study, we focus on how the complexity of the plankton community will lead to different DO responses. Please see the updated Discussion section.

Introduction

L38-40: clarify the terms: overlying water, surface water

Responses: The overlying water mentioned in McCarthy et al. (2013) was the water layer about 20 cm above sediments. We did not mention surface water here.

L52, L80: modify "biochemical" to "biogeochemical"

Responses: please see our earlier response.

L57-65: this paragraph is suggested to move after paragraph L67-76; sentences should be added in this paragraph about how phytoplankton species affect hypoxia size

Responses: Have modified.

L80: these is no such term as sedimentary biochemical… consider biogeochemical processes or other technical terms

Responses: The term "sedimentary biochemical" has been replaced by "biogeochemical processes at sediment layers".

L80-81: remove (Fennel et al., 2016)

Responses: Corrected.

L85-91, 102-103: The necessaries of using higher level representation of plankton community should be highlighted and further specified in the introduction. Reviews of previous work on the defects (for example, poor model performance on reproducing observational Chl-a) of using simplified representation on trophic level should be expanded in addition to L57-65. It was noted by the authors that the influence of the community complexity can be reflected in the SOC and eventually in the bottom DO variability. The goal of this study should prove model performance improvement by adding additional plankton groups by providing better model validation compared to previous works.

Responses: We have specified these in Lines 84–91 where we also posted our focuses of this study. We also compared our model performance with previous in the updated Validation section. For example, DO concentration biases against profile measurements were found within 2 mg L^{-1} in Yu et al. (2015) but were found within 1 mg L^{-1} in our study.

L97, 102: please check how many groups of phytoplankton and zooplankton carefully.

Responses: Sorry for the typos. The numbers of phytoplankton and zooplankton groups should be two and three, respectively.

Biogeochemical model validations
L303-304, Figure 3: which are the profiles shown? If it is statistics of multiple profiles, please normalize the vertical depth.
Responses: It is statistics of multiple profiles. We aimed to provide nutrient profile validation in this section rather than to provide any in-depth analysis. In this revision, we replaced probability histograms with profile biases against distance to the Mississippi River mouth. Please see the updated section 3.2.

L323, 324: plot nutrient concentration bias according to the distance of observation stations to the river mouth would help to discern the influence of river load and model itself
Responses: The plot has been updated according to the suggestion (updated Fig. 3).

L354: how thick is the bottom water column of overlying water? Overlying water is a very confusing term, please describe it with depth range or other more accurate way
Responses: In the incubation study by McCarthy et al. (2013), the overlying water layer above sediments (depth 20 cm above sediments) was isolated for a separate incubation. Respiration rates at the overlying water were then measured. However, in our model, there was no overlying water layer added. In our model, as a sigma vertical coordinate system was used with 36 vertical layers designed, the thickness of the bottom layer is usually ~ 1 m. It is hard to expect the model can reproduce well enough the measured overlaying water respiration with the simulated bottom water respiration. So, we removed the comparison of SOC/overlaying water respiration.

The interannual variability of SOC was actually not well captured by the model (shown by Site F5 and C6). The authors tried to explain the bias with river point sources diverting at the computational grids. However, if that was the case, the interannual variability should have been captured although there might be a discrepancy on the magnitude. Therefore, there might be a systematic defect with the model, either hydrodynamics or biogeochemical processes, that need to be discussed. This can be found in Figure 8 for the overall model underestimation of HV, especially for the year 2007, 2010, 2012
Responses: It is hard for monthly mean simulations to capture exactly the measured SOC. We acknowledge the overestimation of DO profiles, especially at lower layers, which may be ascribed to the coarser vertical resolution near the bottom and the vertical mixing parameterization applied. We have posted relevant discussion in the updated section 3.2 and 3.4.

L379-381, Figure 6: why (a) showed model significantly overestimated DO compared to the measurement from subsurface to bottom layer, while histogram (d) did not show any bias? The model showed consistent overestimation of surface DO, suggesting there might be issues with model air-sea oxygen flux calculation or phytoplankton production dynamics.
Suggest validating DO profile data by season (spring, summer, winter) to discern the probable causes.
Responses: We have posted relevant discussion in the revised section 3.5.

Results and Discussion

L434-439: please specify the oxygen budget equation applied, the integration area, depth, and method, etc. in the method section or result section. Use equation to represent all the terms, including the individual terms in the water column respiration. When you mention bottom DO concentration balance, what is the depth of water column?

Responses: The section 4.1 and 4.2 have been removed. The revised manuscript focuses on the biogeochemical term only (DO source by photosynthesis, DO sinks due to phytoplankton respiration, zooplankton metabolism, nitrification, microbial decomposition of dissolved and particulate organic matters, and sedimentary oxygen consumption). In Eq.4, we provided expressions of these terms. Also, we updated Eq. 4 as the decomposition of particulate organic matter was missed in the previous submission.

In the updated manuscript, DO source/sink terms were averaged over the layers within the bottom 2 m and over layers above, respectively. The bottom DO concentration was represented by the average DO concentration over layers within the bottom 2 m.

L441: why only selected the bottom water column, instead of using the water column beneath pycnocline as previous study?

Responses: See the above responses.

Figure 9: the DO budget terms are generally shown with signs rather than absolute value. Positive sign indicates replenishment of DO while negative sign represents consumption. Horizontal and vertical advection mostly cancel out each other. Therefore, using absolute value can be misleading. Please modify Figure 9 accordingly.

Responses: The section 4.1 and 4.2 have been removed.

Replace "water column biochemistry" with "water column process" throughout this section and figures

Responses: Done.

L450-451: where did you get this conclusion? Please explain a bit more.

Responses: The section 4.1 and 4.2 have been removed.

L455: again, what was the depth range of bottom layer of this study?

Responses: See the above responses.

L494-496: Lack of explanation on the physical mechanisms of stratification on advection terms. Please add the physical understanding here, including the circulation and mixing dynamics in LaTex area. Why the enhanced stratification would suppress DO advection, which is not the case in other hypoxic estuaries like the Chesapeake Bay? What did it mean that "the enhanced water stratification in summer usually leads to less DO exchanges due to advection at the bottom layer"?

Responses: Sections 4.1 and 4.2 have been removed.

L501-505: with this explanation, the negative correlation between PEA and vertical diffusion of DO makes little sense since the bottom DO concentration was the controller. The stratification just covaried with seasonal freshwater discharge and/or wind strength in this area.

Responses: Sections 4.1 and 4.2 have been removed.