**Responses to Anonymous Referee #2**
Below the review is reproduced in black font and our responses interspersed in blue.

**Comments:**
This publication presents a numerical model study of the hypoxia events off the Changjiang Estuary.

The combination of the different modeling components is a priori convincing and appropriate: 3D oceanic circulation model, biogeochemical model, sedimentary oxygen consumption module, river discharge (nutrient and freshwater load), and atmospheric forcing from reanalysis.

My main concern is about the model validation or skill evaluation before any use.

**Reply:** We appreciate this overall positive assessment and believe we can address the Reviewer's concerns as described in more detail below.

The model-data comparison is presented in section 3.1, Figure 2 only, and some other in the Supplement. The display of Figure 2 is problematic: color points (data) having the same color (same value) as the background (the model) do not appear. It is really difficult to see the observational structure and to evaluate the agreement with the model. (same for Figures S2, S3, S4). It could be separated figures (data distribution and model). Figure S6, including the bottom line, is much more speaking.

**Reply:** We would like to add an additional model-data comparison for nitrate in our revision to illustrate that the model reproduces nutrient distributions well. Also, we agree with the Reviewer about the data points blending in the background and will reproduce the comparison figures for the in-situ observation comparisons of simulated salinity and temperature in better quality. We are happy that the Reviewer finds the 2D histograms in Figure S6 (and also Figure S1) informative. However, such graphs only make sense when a large number of data points is available (usually only for satellite data), which is not the case for the in-situ observations in Figures S2 to S4. Those will be improved by choosing a more appropriate, colourblind-friendly colourscale and better demarcation of the data symbols. In the revision we will add a subsection dedicated to model validation.

The authors aim at reproducing the observations from 9 cruises from march 2011 to september 2013. Therefore, the simulation starts in 2006, uses climatological observations from this period to force the model, 2006-2007 are used as spin-up, and the model is run in 2008-2013 for analysis. Regional models may be very dependant on the boundary conditions. Nothing is demonstrated about the robustness of the inner region : is there any drift in the total budgets (nutrients, oxygen, intensity of the Primary Productivity) ? The model is set and used. I would be more confident with the results if any sensitivity test would be performed. By example, it would be possible to run the model for the same duration (8 years) but with repeating the same annual forcing (e.g. 2006), in order to control that the inner structure of the PP and hypoxia are repeated or if any trend exist. It would also evaluate the model-internal-variability, not to be confused with the variability induced by the varying forcing (winds, river discharge).

**Reply:** We are confident that there is no significant drift in the domain and will provide suitable plots to show this. We would be happy to add this to the Supplement during the revision. Also, we appreciate the Reviewer's suggestion to contrast the original simulation with realistic forcing and a climatological simulation that repeats the same forcing every year. We are currently conducting this simulation.

The model is used in its "optimal" configuration, but the evaluation to reach to this configuration is not presented. The model here is not used to make any sensitivity experiment. Part of this is explaned late in the paper (line 384, just before the conclusion): there is a companion paper (Grosse et al.) that presents modeling experiments to quantify the relative importance of the processes responsible for hypoxia. This is important since the authors just infer the importance of processes (lines 325-334), without proceeding to the sensitivity test to their hypotheses. In this case, I would indeed recommand to proceed to a simulation while removing the nutrient load (which seems to be done in the companion paper). It should be presented from the beginning that part of the modeling analysis is done somewhere else.

**Reply:** We will make clearer early on in the revised manuscript that the sensitivity to nutrient load is investigated in the companion paper by Grosse et al.

Concerning the main conclusions of the publication, the analysis of the main contributors to hypoxia, in the whole water column and in the bottom layers, is relevant. It is important to be able to evaluate the relative importance of Water Respiration versus Sedimentary oxygen consumption. But once again, data are missing, or at least a more rigourous model-data evaluation. As an example: Figure 3 focuses on the patterns of the hypoxia events from 2008 to 2013, and different behaviors or chronology could be distinguished (that is very interesting in itself, and the modeling tool is really appropriate for this kind of studies). Unfortunately, it is unsufficiently documented, how does this relate to observations ? Same for the discussion about the influence of wind events (4 typhoons) on the hypoxia extent.

**Reply:** We appreciate the Reviewer's assessment that our main conclusion about the contributions of water column versus sediment respiration is relevant and important. We also have to acknowledge that a rigorous model-data comparison is desirable but hampered, to some degree, by the relatively limited availability of observational data. We will add information about observed rates of SOC to the revisions. Also, we recently became aware of a nutrient data set for the region and will add comparisons to this in our revised manuscript. Futhermore, we present model-data comparisons of satellite-derived SST and Chlorophyll, and model comparisons against in-situ data of temperature, salinity and oxygen. We believe that these comparisons provide the best attainable and sufficient level of confidence in the model's ability for us to present model results. However, we fully agree that more would be much better. If the Reviewer is aware of any additional in-situ data that are available, we'd appreciate hearing about these and would happily include them.

With regard to the interannual variations shown in Figure 3: Unfortunately, we do not have sufficient observations to validate the interannual differences in hypoxia development in these years. However, we would argue that the model-data comparisons we have (with the additional comparisons for nitrate) do provide sufficient confidence for using the model to analyze the oxygen dynamics in the region and to present simulated interannual differences. In fact, the Reviewer seems to agree with us that this is an appropriate use of the model. Of course, future models will be better, but this is the best that is available now and we believe it provides useful insights.

I would recommand to improve the model-data evaluation in order to convince that the modeling of hypoxia events are (1) not biased by model-depending behaviors (2) close to observations.

**Reply:** We believe we can satisfactorily address both of these points. See responses to the comments about model drift and validation above.

Specific comments:

The model includes a light-attenuation term dependent on water depth and salinity (lines 177-181). Could you confirm that places where the light attenuation is applied ( f(z,S) ) are indeed places where particles (RDOM, Detritus , phytoplankton, . . .) are present and induce this shadowing effect ? Some other parameterisations exist that compute the shading directly in situ from the biogeochemical species. Using depth and salinity has the convinience to put this effect where it has been observed, but has the inconvenience to decouple the modeled biogeochemistry from its shading effect.

**Reply:** Thank you for raising this point. We have to clarify that light is attenuated everywhere in the model domain by seawater constituents (specifically chlorophyll and detritus) and seawater itself. In addition to this, light is also attenuated by suspended sediment according to the parametrization referred to above. Observations show relatively higher suspended sediment concentrations, and thus light attenuation, in shallow areas (Bian et al., 2013; Chen et al., 2014). To account for this additional contribution to light attenuation by suspended sediment, which are not explicitly modeled, a simple parametrization depending on bathymetric depth and salinity is implemented. We will clarify this in the revised manuscript.

Minor comments

line 60. ref. Fennel & Testa : missing comma

line 187. "based on"

Figure 3 : labels a, b and c are missing on the figure itself.

**Reply:** We will address all minor comments in the revision.