Biogeosciences Discussions

Biogeosciences Discuss., referee comment RC2
https://doi.org/10.5194/bg-2022-4-RC2, 2022

Comment on bg-2022-4

Anonymous Referee #2

Referee comment on "Hydrodynamic and biochemical impacts on the development of hypoxia in the Louisiana–Texas shelf – Part 2: statistical modeling and hypoxia prediction" by Yanda Ou et al., Biogeosciences Discuss., https://doi.org/10.5194/bg-2022-4-RC2, 2022

In their manuscript, Ou et al develop a novel statistical model to forecast/hindcast the size of the hypoxic area in the northern Gulf of Mexico. They use the model to test the feasibility of using HYCOM output and atmospheric data (reanalysis and forecast) to forecast the size of the hypoxic zone. The manuscript is well written and the statistical model seems to be able to retrieve the hypoxic area simulated with the ROMS model (part I paper). I am not familiar with the GLM/GAM statistical techniques and hopefully another reviewer can verify this part of the methodology. My overall assessment is that some improvements are required before the manuscript should be considered for publication. There are a few points that I think are important and would like raise below. Other, more specific comments are listed afterwards.

1) In its current form, the manuscript is mostly methodological and therefore I don't know if BG is the best fit for it. This could be solved with some improvements. For instance, the Discussion section present an example of how to use the forecast model. This is a really interesting approach but it feels like a quick addition to justify the model development, that will be "further improved" in the future. A proper set of "forecasts" that are tested against observations would make a much more compelling case for the models ability to forecast hypoxia. 1985-2021 mid-summer observations are available for this test; I believe that HYCOM and atmospheric forcing data are available in the recent years to carry out this analysis. The forecast input data come with (high?) uncertainty and it would be interesting to know the effect on the hypoxia forecast (compared with the reanalysis input).

2) My second point is a follow up from above. The manuscript relies exclusively on models. This is fine as a methodological paper but not if the authors aim at improving the current (seasonal) hypoxia forecasts and providing a tool for managers. For instance, it is assumed that the ROMS hindcast is a true representation of LaTex hypoxia. This is obviously not the case (as with any models) and it seems important to include observations in the manuscript to see how/if the forecasts drift away from the observations as we go from ROMS to GLM/GAM to HYCOM. Also note some of the reviewers comments on the Part I paper referenced here. Furthermore, the model provides a highly temporally resolved forecast, but it is not clear to me if, as a forecast, it does better than the seasonal forecast models (cited in the Introduction) that are, for some of them, spatially and temporally resolved. Some comparison with those (available annually through NOAA, e.g. https://www.noaa.gov/news-release/noaa-forecasts-average-

sized-dead-zone-for-gulf-of-mexico) would strengthen the manuscript.
3) The part that needs significant improvement is the Discussion, which is not really available in the current version of the manuscript. Rather, the Discussion section presents an attempt at a "real" forecast using HYCOM. This could be moved to the Results section and a real Discussion section should be provided. What does this new technique brings to the knowledge of LaTex hypoxia? How does it compare with earlier models? How is this useful to managers? What are the caveats and limitations? what are the future developments? How is this technique portable to other systems? All of those are legitimate points that should be discussed.

Specific comments

L36/53: Those are seasonal forecast and cannot include the wind since it is not predictable at this time scale.

L56: Stratification is included indirectly in the statistical models

L58: They are not pseudo forecast, they forecast the mid summer hypoxic area (well in advance). Therefore, they are seasonal forecasts, which is different from the short-term forecasts provided by HYCOM.

L58-59: "fail whenever winds are strong in summers": Note that some of these models provide information on the effect of the wind on the forecast

L76: FYI (related to the main comment above), looking at the comparison between ROMS and observed mid-summer hypoxic area in Part I manuscript, the r-square is 0.58.

L79: could you define the geographical limits that you use for the LaTex shelf? That would be helpful to have a sense of your comparisons as it is not clear if you use the same area as the mid-summer sampling cruises to calculate the hypoxic zone.

L91: what do you mean by up to?

L148: It might be helpful to include these equations here.

L158: Can you discuss the biological meaning of this time lag? It seems to indicate that

mid-summer hypoxia is fuelled by early summer loads and therefore that there is no relationship between May load and summer hypoxia.

L176 (Table 1): "Hypoxic area" would be better than "Area of extremely low dissolved oxygen concentration"

L194: Figures are not presented in order, please reorder

L198 (Figure 1a): The lack of relationship between SOCalt and botT is a bit concerning, can you comment?

L198 (Figure 1g): What is the time range of these data, all year, spring-summer, spring-fall?

L223-228: I didn't get how this added term solves the high level of correlation between predictors

L258: "impaired"

L264-274: Not sure if that is a good test of model skill. Excluding randomly half of the years (or 30-40%) would have provided a good dataset for testing. Can you discuss why you did not split the hypoxia data into years, since hypoxia is a seasonal process?

L281 (Figure 4): You should add observations.

L289: the correlation doesn't seem to be significant

L293 (Table 2): What is Pr? does it make any sense to provide a Pr of <1e-16?

L299: "procedure"

L316-317: Early summer or spring? It looks like hypoxia develops in Spring in the time

series

L316-323: Do you see all that in Figure 5?

L343: This is not a discussion, see main comment above.

L373: Why not doing that for the entire time series?

L376: It is an interesting technique but lacks observations, why didn't you do a real forecast, i.e. a week ahead of the mid-summer cruise, for each year where the input data are available?

L377: "slight": ~20+% difference

L386: Your model forecast doesn't seem to do better than the seasonal forecast in 2019 and misses the pre-sampling mixing event, can you comment? The 2020 mid-summer hypoxic area is also largely overestimated (~20,000 vs 5,000) and seem to be doing worst than seasonal forecasts despite the model ability to take into account the effect of wind (there was a tropical storm before the mid summer sampling that year)