Comment on bg-2022-4

Anonymous Referee #1

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-RC1, 2022

General Comments:

This manuscript applies an ensemble regression approach to produce daily predictions of hypoxic area for the Louisiana-Texas shelf. The manuscript is well written and provides more than adequate descriptions of the methods used to develop, train, and apply the multiple regression models considered for application. Although the ensemble model's application to global HYCOM seems an important aspect of this work, it's sole focus in the discussion feels somewhat like an afterthought. HYCOM presentation in the discussion also presents material that seems better suited for the methods section. Although not essential for publication (in my opinion), I would ask the authors to consider expanding the discussion of HYCOM application in the manuscript by addressing relevant technical details in the methods section and focusing on model outcomes for both the ROMS and HYCOM model in the discussion.

I have no major issues with the manuscript as presented (beyond my recommendation for the discussion of HYCOM), and recommend publication with mostly minor revisions as described below.

Specific Comments:

- Line 55: "The effects of water column stratification are not included or only partially considered" In the previous paragraph you describe several models as incorporating water reaeration and wind velocity in the regression model. Are these not at minimum proxies for water column stratification? Suggest this be re-phrased to address the need to include stratification explicitly.
- Line 155: where was the "temperature-dependent decomposition rate of organic matter" derived?
- Line 179: Is a shelfwide average appropriate for all predictors? Was any attempt made to derive predictors for different longitudinal zones of the shelf? For example, the ROMS grid extends far to the western limits of the hypoxic zone along Texas, and thus stratification predictors averaged across the entire domain may be less dynamic than otherwise expected.
- Figure 4: The 95% CI are not visible. Is this because the confidence intervals are extremely tight and not visible at this multi-annual scale?
- Figure 7: Please clarify in the caption where the hypoxic area from the ROMS hindcast is coming from. Is it from the ensemble model, and if so, should 95% CI's be applied here?

Technical Corrections:

- Abstract line 20: suggest changing "is by far the first one providing" to "is the first"
- Line 24: superscript "L^{-1}"
- Line 56/57: "The information of future conditions is limited although some models are built upon multiple predictors, thus these forecast models are indeed "pseudo-forecast" ones." This sentence is awkward. Suggest rewording to: "Information on future conditions is often limited to few predictors, thus limiting these forecast models to "pseudo-forecasts""
- Line145/146: "However, by far, global forecast model systems like HYCOM does not include biochemical fields" This sentence is a little confusing. When you say "fields", do you mean "parameters"? HyCOM is strictly a hydrodynamic model, so it is sufficient to say "However, global forecast models such as HYCOM do not simulate biochemical parameters. Therefore, the biochemical-related term SOC needs to be replaced by an alternative term (denoted as SOCalt)."
- Line 150: It may be more appropriate to describe nitrogen as available for plankton growth, not bloom.
- Line 164/165: "For simplification, we denoted this variable as (Qh), W3, and ð□□□ as PEAheat, PEAwind, and DCPTemp, respectively." It took me a few reads to figure out what this sentence was trying to say. Suggest removing the word "this" and modifying to "denoted the variables"
- Line 316: Change "It implies" to "This implies"