We sincerely thank the reviewers for providing comments on our manuscript. We provie a point-by-point response to these comments below. In general, most of the revisions address concerns raised by all three reviewers, including 1) improvement in writing style for clarity and brevity, and 2) additional information on sampling methods. We have also addressed concerns of Reviewer 2 about potential nutrient cycling by adding a section in the discussion and concerns of Reviewer 4 regarding additional detail on the simulation modelling.

## Reviewer 1

The authors (a very long list of authors) have described the sequence of events that followed a phosphate mining spill. The paper is long and tedious- and in the end cannot conclude any strong relationships with red tide occurrence. I wish I could be more positive about this paper, but it has the feeling that authors were rushing to get something in press, rather than taking the time to write a well crafted manuscript. There are some very good writers on the list of authors- so I am a bit surprised that they all agreed to this final version.

Thank you for your comments. We have made revisions to improve the writing, by reducing length in many locations and making additional edits where noted. Please see the details below.

graphic abstract- much too complicated; fonts too tiny, too much text

* **Response**: The graphical abstract was simplified and the font size increased.

highlights- the term Piney Point is meaningless except for locals

abstract- poorly written

Introduction- needs a complete rewrite- it takes more than 3 pages to get to the point that there was a spill. The general background is interesting but you need to grab the readers attention first.

p. 5- inadequate explanation for why the focus on N. Why not both?

Methods- the first paragraph on monitoring is all about modeling, not monitoring

p. 7- how were the Karenia data quantified…they didn’t just appear in a database.

p. 7 - eyes on seagrass is meaningless (except for locals)

stats section- poorly written, confusing, wordy

Results- rambling, disorganized and non-quantitative. Paragraphs don’t make sense (try using topic sentences and structuring concepts accordingly)

Discussion- I was left wondering what I learned. No conclusions drawn

In all, it seemed like an interesting exercise to describe all the patterns, but this does not make a quality manuscript that will stand the test of time.

## Reviewer 2

Beck and colleagues report on the short- and mid-term effect of a very large nutrient-rich release from Piney Point. The authors focus on a few water quality indicators, but also include data for several additional responses (e.g., seagrass coverage). The results are largely what you would expect, initial rapid response of phytoplankton with cascading secondary responses.

First, a few of the things that make this a worthwhile contribution to MPB. The authors should be commended for their quick response to data collection and the large number of data streams that were brought to the study. It really is quite impressive. Further, compared to citizen monitoring, I expect the data is high quality and consistent. Second, the writing was quite good; there were some sections where flow or organization could be improved (beginning of results; last section of discussion); however, these considerations are very minor.

Second, a few things that could be improved or warrant some critique. What is the rationale for the lack of continuous line numbers? Journal suggestion? Reviewing is so much easier when line numbers line up with the text and continue throughout the document. I realize this is not something that should be fixed now, but it is bothersome. That said, my one major critique of the manuscript, one that is difficult to reconcile with the available data, is the speculative nature of last section of the Results, titled Potential nutrient cycling. This is clearly a very important section and, in my opinion, warrants additional consideration. The title of this section is correct, yes; thus, I would like to see 1) this section moved to the discussion, 2) the MAJOR lines of evidence supporting this thesis outlined and discussed consecutively in one section, 3) major alternatives entertained and addressed. The authors might also think about what it would take to make the connections more explicit, if possible (stable isotopes, modeling, etc.).

MINOR COMMENTS/SPECIFIC SUGGESTIONS

P3/L56: Recommend removing “in other countries”. When I read that passage it seems to suggest a broader critique, one that is more political. That may not have been your intention, but that was how I read it.

P4/L21-24: I am not sure if “unanticipated releases” captures the spirt in which these releases occurred. In fact, the title of the FWC report was “Response of estuarine nekton to the regulated discharge of treated phosphate-production process water”. This regulated discharge seemed to occur after extensive consultation with the EPA.

P4/L51: worth indicating that these concentrations, especially for TP, are several orders of magnitude greater than typical surface waters?

P4/L56: deliberately “released”

P5/L4: I see the use of the term lower Tampa Bay; how does that compare to Figure 1, the areas of interest? Ultimately, lower Tampa Bay is used extensively, but I don’t see it defined. Also, is this region ecologically based or defined by currents?

* **Response**: The boundary between middle and lower Tampa Bay was added to Figure 1. These boundaries are used by the Tampa Bay Estuary Program to track annual attainment of water quality targets under a Reasonable Assurance plan for TMDL reporting to the Florida Department of Environmental Protection. We do not define the rationale for the boundaries in the main text because the regulatory delineations of the bay are not of interest to a broader audience, but we use the delineation as it is useful for our discussion of trends in the paper. Our area delineations (areas 1 - 3) were meant to provide a more relevant grouping for the paper.

P5/L22-27: I think a citation may be warranted here; while predictable, numerous studies ultimately supported the statement that nitrogen is limiting in Tampa Bay.

P5/L29: interesting that discharge was to Port Manatee and not Bishop Harbor like the release discussed previously.

P6/L24-39: that is an impressive and extensive list of agencies and partners; any potential drawbacks to so many data streams?

P6/L54-65: is it routine to take water samples directly from the surface? Ultimately, were the authors worried about any potential stratification, particularly of Karenia brevis.

P7/L32-37: why get data from Tampa Airport and the Airfield in St. Petersburg? There is an NWS location not far from Piney Point.

* **Response**: We have updated our weather data to include wind and precipitation from Albert Whitted Field in St. Petersburg and have excluded the precipitation data from Tampa International Airport. Our initial rationale for including Tampa was a longer period of record. However, we have revised the historical baseline in the plot to include only years 2006 to 2020, consistent with our comparison of the water quality data with the historical observations. Also, the NWS location is in Ruskin, Florida and is farther from the bay proper than the location in St. Petersburg.

P9/L27-29: not sure what to make of the aggregating data to weeks or months; isn’t there a way to maintain the actual sampling regime?

P9/L49: This is a bit confusing; is this comparing “between” pre and post?

P11/L12-14: as personal preference, I see no reason to lead off the results with a general statement about the location of a figure summary. Instead, the authors might reference the figure when they describe, in general, that a bloom was observed ~2 weeks after the release began, red tide was observed within x weeks, the bloom peaks, and conditions become stable on xx.

P11/L20-: this first part of the results has a choppy structure

P11/L31-34: the authors indicate the Tampa Bay influx was related to an ongoing coastal bloom; is this indicating the red tide bloom was ongoing prior to the Piney Point discharge? Also, not clear what “related” means in this context. This important statement may require additional context.

P16/L26: not sure I see the rationale for analyzing the data using Braun Blanquet estimates, which, in my impression, was developed as a rapid assessment

P17/Potential nutrient cycling: Unfortunately, much of the latter portions of this section are speculative and better presented in the discussion. Further, it would have been nice to see something a bit more concrete linking the nutrient shift from photoplankton to macroalgae and release of macroalgae nutrients to K. brevis; stable isotopes perhaps; the nitrogen sources could have been highly enriched

P18/L18: or a shift in the relative availability, N:P

## Reviewer 4

This is an interesting study in that it utilizes a quasi-Lagrangian approach to quantify the influence of a nutrient-rich waste plume on an estuarine ecosystem and is the first that I am aware of that documents the impacts of the Piney Point episode. Nonetheless, there are some points that require clarification to give the reader confidence in the findings. In general, more details are needed on several aspects of the methodology, and I provide specific details below on this as well as other issues.

Page 1, Line 38 - “Elevated levels of phytoplankton” - needs to be clearer… is this referring to biomass or abundance?

Pages 6, Lines 4-22 - The sampling regime was largely dictated by output of a model of plume evolution. It would be useful to see more details on how the model is calibrated, and if/how its output is cross-verified. In other words, are field observations of T/S compared with model output to evaluate its performance in terms of accuracy of plume location? This seems like critical information for assuring that the sampling regime actually followed the plume.

Page 6, Lines 46-54 - More information is needed on laboratory procedures. In particular, with many different entities running samples and presumably utilizing different analytical approaches for each variable, it is important to know how comparable the data is between programs.

Page 7, Line 22 - Were the breakpoints for K. brevis, or total phytoplankton abundance?

Page 7, Line 37 - What about from ungauged flows?

* **Response**: Our methods for estimating inflows into Tampa Bay are based on methods that are similarly presented in a technical report published by the Tampa Bay Estuary Program in 2012. This report provides the basis for compliance assessment of our TMDL for reporting to the Florida Department of Environmental Protection and represents our best estimate at hydrologic load inputs in to the bay. We have provided a citation to this report in the paper.
* Janicki Environmental, Inc., 2012. Development of a screening level tool for estimating annual hydrologic loadings to Tampa Bay (No. 05-12). Tampa Bay Estuary Program, St. Petersburg, Florida.

Page 11, Line 30 - Need to put Anna Maria Sound on the map. Likewise for Port Manatee.

* **Response**: Locations were added to the map.

Page 12-14, Water quality trends section - Very little attention is given to the role of the hurricane that hit the area in early July. Would be interesting to hear more about its effects on water quality in the study area.

Page 14, Macroalgae and seagrass trends section - This appears to be one of the weaker components of the project, although it is outside of my expertise. My biggest concern is that it is not clear to me how useful “% occurrence” is. As far as I can tell, this is not equivalent to biomass or areal coverage. Perhaps I am mistaken? Regardless, it would be useful to see a better justification for this metric. Also, how did the 2021 data compare to historic spatial-temporal trends?

Page 17, Lines 17-20 - How did the 2021 fish kills compare to historical kill spatial-temporal trends?

Page 17, Lines 37-40 - Could there have been a role for low D.O. in the post-storm fish kills?

Discussion - In general, it would be useful if the team would spend some time establishing the prevalence of N-limitation in the system, either based on their own data or previous studies. The interpretation of plume impacts is predicated on the system being N-limited, so that needs to be established up front.

Page 20, Line 22 - Ammonium, not ammonia. Or was pH such that ammonia would be the dominant form?

Page 20, Line 39 - Might be worth having a separate paragraph that talks about the timing/duration of the K. brevis bloom in 2021 compared to historic conditions.

Page 20 (last paragraph)-21 (first paragraph) - Without showing methods and/or data, I feel that it is inappropriate to introduce the nekton abundance/composition here.

Page 21, Lines 32-47 - Much of this text on seagrass seems speculative, and again, unclear about use of % occurrence as a metric.

Page 24, Lines 17-52 - Speculative without showing methods of data collection or the actual data.