*Response to reviewer comments “Water quality trends following anomalous phosphorus inputs to Grand Bay, Mississippi, USA”, by M. W. Beck, K. Cressman, C. Griffin, and J. Caffrey.*

*We thank the reviewers for providing helpful comments on our manuscript. Our responses to each of the comments are in italics. Reviewer comments are edited for brevity where appropriate.*

**Associate editor comments:**

Overall, the reviews are quite good, and request only minor revisions.

One reviewer did note some potential statistical issues dealing with autocorrelation. Essentially, the authors divided their time series into 6 time periods, and then used Kruskal-Wallis one-way ANOVAs to test for differences between these groups. In their revision, I'd like to see the authors justify their use of these tests, and discuss any potential autocorrelation in their data and how they dealt with that (as requested by the reviewer).

*Closer evaluation of the model residuals showed some autocorrelation for orthophosphate at BL and BN and chlorophyll at BN and PC. We filtered the time series using an additive seasonal decomposition to remove components of the time series that were strongly correlated. Autocorrelation plots of each station and nutrient parameter suggested that the filtering was effective, but some residuals still exhibited minor correlation. We include these plots at the end of this document. We emphasize that the important result is the clear impact of the spill events, such that our conclusions remain the same, independent of any statistical artifacts from autocorrelation in the data.*

*The methods were revised to explain the filter.*

*‘Further, nutrient time series were filtered using a seasonal decomposition to reduce autocorrelation among model residuals. All nutrient time series were evaluated after filtering the seasonal component using the decomp\_cj function of the SWMPr package for R (Beck 2016).’*

*Results were updated where applicable (see the section ‘Differences between sites and time periods’).*

Also, I'd like a bit more information on the change point analyses that they mention very briefly in lines 166-169 ("These divisions [the six divisions within their time series] were generally supported by a change point analysis that identified continuous periods of time within which the statistical properties of the phosphorus time series were similar (Killick and Eckley, 2014))." Were all of the divisions supported by the change point analyses, or just some? Why not use the change point analyses as the basis of the divisions, rather than simply a way to check whether or not the divisions were "generally supported"?

*We explored the use of changepoint analysis in an earlier version of this manuscript to identify natural breaks in the time series. This exploratory work demonstrated that it was possible to define the time periods this way and that the results generally confirmed our a priori designations. However, it was difficult to identify divisions that were applicable across all stations and parameters. For example, breaks in the data that were identified for the BL station were very dissimilar from those identified at the PC station. Rather than following a formal statistical approach to define the time periods, we chose an alternative approach where we identified the time periods following our visual interpretation of the time series plots and knowledge of deviations from background conditions. We felt this approach was more applicable to our analysis that evaluated differences in time and space, and not necessarily discreet points in time specific to each station and parameter from changepoint analysis. The latter approach was simply too sensitive to noise in the data. Our mention of changepoint analysis was meant to provide a justification for our time periods, although in hindsight it confuses rather than clarifies. In our revision, we’ve removed any mention of changepoint analysis to identify time periods and have provided a better justification for the time periods:*

*‘The a priori identification of time periods was primarily based on visual interpretation of nutrient parameters at each station and the authors’ knowledge of deviations from background conditions in Grand Bay. This approach was preferred over more formal statistical approaches to identify breaks in the data (e.g., changepoint analysis) to prevent identification of spurious changes that likely differed between parameters and sites. The analysis below provided a thorough evaluation of between-group comparisons using additional statistical approaches.’*

**Review 1:**

One major concern is whether the statistical analyses are appropriate for autocorrelated time series data. The analyses may be fine as is; however, the manuscript does not mention whether or how potential autocorrelation issues were addressed (see # 4 below). Please explain further how the analyses are appropriate, given the non-independent nature of the data.

*Please see our comments to the associate editor above.*

I also have a few minor suggestions, as outlined below.

1. Lines 39-41: Can you elaborate a bit more in the abstract on how the results improve understanding and have management implications?

*This study provides a rare example of using long-term monitoring data to evaluate phosphogypsum inputs in a nitrogen-limited system. Describing these trends in the primary literature will have value because they are not reported extensively in other systems. We have added the following to the abstract, but have not expanded on these ideas due to word limit constraints: ‘The results provide new information on the effects…’ The potential value of this work is reported in the introduction (e.g., line 50-52, 92-94)*

2. Line 128: Please explain further why water column stratification warrants exclusion of this station from the analyses.

*This site was excluded from analysis for several reasons, but this was initially not explained for brevity. The following was added:*

*‘For this analysis, we did not include BH data* *because site-specific characteristics prevented meaningful comparisons with the remaining sites (i.e., frequent stratification, regular freshwater dominance, sustained summer hypoxia, and groundwater influence).’*

3. Lines 168 & 230: Can you elaborate on which changepoint method you used (change in mean, slope, variance, etc.) and what, specifically, the changepoint analyses results were?

*Please see our response to the associate editor above.*

4. Lines 182-196: Were the data autocorrelated? Can you still perform these tests if the data are not independent? Did you conduct any tests to check that the data meet the assumptions of the statistical tests?

*Please see our response to the associate editor above.*

5. Line 385: Can you elaborate on how Fig. 5 illustrates tidal advection of nutrients?

*This figure shows diel sampling of nutrients (orthophosphate) and salinity before and after the second event. The point was to demonstrate that nutrients were transported 7km from the spill site to Bayou Cumbest. The diel patterns in nutrients after the spill follow the tidal pattern, suggesting physical transport by tidal advection. The previous sentence was added to the figure caption.*

6. Figures overall: color blind people generally cannot tell the difference between red and green – to maximize readability, you might consider changing the color palette in figures that contain both red and green.

*All palettes were changed to color-blind friendly (BrBG) or linetypes/shapes were used to discriminate plot objects.*

7. Fig. 3 E2A Salinity, Turbidity, and pH: Does the straight line right after Sep 15 indicate missing values? If so, it’s best to remove the line to make it clear that data are missing during that period.

*Lines were removed.*

8. Fig. 5: It’s sometimes considered poor practice to plot multiple axes on one panel, although there are exceptions to the rule. In this case, I think the plot would be more easily readable with three stacked panels, rather than one panel with three axes. It is disorienting to match the symbols of each line with the symbols in the legend with the axis labels.

*We generally agree that overplotting can be abused, but we feel it’s justified in this example. We don’t care so much about the relative values here, rather the correlation of trends after the spill and this graph does a good job showing the relationship. However, we’ve modified it slightly to reduce some of the clutter (no more dashed lines, moved/resized legend, removed legend border). We haven’t changed the red/green colors because the lines can be identified by shape.*

9. Figs 6, S1, and S2: What do the colors of the boxes represent? Also, it’s difficult to differentiate between the shades of green.

*The colors represent relative magnitudes of the median for each box. This information was added to the figure captions.*

**Review 2:**

This manuscript describes a time-series with analysis of several key events which triggered changes in water quality. A combination of change point analysis and comparisons of events to baseline is provided. I think the paper is balanced, well written concise and I would like to see more time series data like this published throughout the coastal and estuarine science literature. This paper should be accepted with copy edit revisions as explained below.

INTRODUCTION

Ln. 45 “our” Is this the authors or is this the field of coastal scientists?

*This meant the collective ‘we’, but it was changed for clarity: ‘Anthropogenic nutrient inputs into natural water bodies have been well-described and the ecological understanding of eutrophication continues to evolve with new information (e.g., Cloern 2001).’*

RESULTS

TABLE 2 Number of decimel places varies in DOsat and PO4 (and NOX). I’m not sure what your precision for Chl a is but I would think 1 decimal place is sufficient. Instead of listing the units in the table legend, suggest that you puth thm next to the parameter in parentheses e.g. Chl-a (ug L-1)

*Significant digits for chlorophyll, DO saturation, pH and salinity set to one, all others as two. Units were placed next to the appropriate row variable and removed from caption.*

Ln. 210; Sentence starts with 2012, suggest that you change this to start with a word.

*Changed to: ‘Years 2012 and 2013 were in the upper 80th percentile, …’*

Figure 3; It would help to put the year on the graphs rather than in the legend.

*Year labels were added to the left panel of each column.*

Ln. 235 Sentence starts with PO4, should start with Phosphate (PO4)

*Changed to orthophosphate.*

**Review 3:**

General Comments

Several places in the *ms* was difficult for me (the reader) to follow, primarily because results

would be reported for one parameter,then another, then back to the pt parameter (e.g., paragraphs beginning on lines 216, 249 and 286). Of course, I may have misinterpreted the results reported, but none the less, clarification should be considered.

Specific comments:

Line 27: Example/Suggested Wording (SW)- We examined spatial and temporal patterns...

*Changed.*

Line 29: If incorporate Line 27 SW,then delete ‘and time periods’

*Changed.*

Line 30: Suggest changing Phosphorus to Orthophosphate

*Changed.*

Line 42: Other keywords to consider- phosphate, fertilizer, spills. MS

*The key words were changed to ‘estuary, monitoring, phosphate, phosphogypsum, spills’ to keep within the word limit.*

Lines 45-52: Move paragraph to Line 74, with suggested changes

*Moved and suggested changes incorporated.*

Line 53: Make this the beginning (1st paragraph) of Introduction

*See above comment.*

Line 57: SW- this process, which was stored in large stacks as waste, and is rich...

*Changed.*

Line 62: Considers starting a new paragraph at the beginning of this line.

*New paragraph added.*

Line 63: sw- cm fell in...

*Done.*

Line 69: How much wastewater? If you know, then provide, as did for the April 2005 event

*This information was received through a FOIA request to Mississippi DEQ and the following information was added: ‘**MPC released 90 million gallons of wastewater over three days into Bayou Casotte to the west (public records request to Mississippi Department of Environmental Quality, December 2017), where a fish kill was subsequently observed (MDEQ 2015).’*

Line 69: SW- west, and concomitant elevated levels...

*See above comment.*

Line 74: Insert the previous mentioned paragraph here (Line 45-52).

*Moved.*

What was Line 46: SW - Like Grand Bay, most estuaries are nitrogen...

*Added.*

At the end of this newly placed paragraph, SW for last sentence of paragraph -The potential effects.. .less studied, and there are few...nitrogen-limited systems, especially regarding impacts of concentrated fertilizer waste loads.

*Added.*

Line 75: Therefore, will omit ‘…to nitrogen-limited estuaries on ecosystem conditions.’

*Omitted.*

Line 75: Start a new paragraph that begins with SW-The few studies ...

*Added.*

Line 114: Where is Escatawpa River? Its not on the map. How far is it from Grand Bay? 8-16km? How is it connected to Grand Bay?

*The Escatawpa River is north of the reserve and cannot be viewed on the map. We have revised the text as follows:*

*‘The reserve is a retrograding deltaic system and does not have significant freshwater inflows. The Escatawpa River is located approximately five km north of the reserve and minor inputs into Grand Bay may occur during flood events through small channels. ‘*

Line 123: SW- The current network includes four instrumented sites (Bayou Heron, BH: Bayou ....) that continuously measure physiochemical water quality, and one weather station (Figure 1). Monthly nutrient sampling at the four continuous stations

began in March 2005 and...

*The first part was changed:*

*‘The current network includes five continuous stations: four that measure physiochemical water quality (Bayou Heron, BH; Bayou Cumbest, BC; Bangs Lake, BL; and Point aux Chenes, PC), and one that records meteorological data (Figure 1).’*

*We left the “monthly nutrient sampling” as-is (“at the water quality stations”) because the weather station is also continuous.*

Line 132: SW-At the continuous water quality collection sites, YSI dataloggers measure and store temperature, specific...every 15 minutes. The datalogger sensors were deployed 0.5m above the bottom before August 2005 and 0.25 thereafter.

*Changed to ‘**Water quality parameters logged every 15 minutes by YSI dataloggers were water temperature, specific conductance, salinity, dissolved oxygen, pH, turbidity, and depth. Dataloggers were deployed 0.5 m above the bottom before August 2005 and 0.25 m thereafter. Meteorological data were logged…’*

Line 150: SW- (NH/), nitrite+ nitrate (N02-/N03- ),

*Changed*

Line 150: orthophosphate (P043-)... Now that this is established as what was measured, from hereafter, be consistent using either one of these terms, but mostly use the P043-. Change the occurrences that use other terms that are intended to mean the same thing (e.g., phosphorus and phosphate).

*We have edited all occurrences for consistency, PO43- was used in all cases except at the start of a sentence or in the introduction.*

Line 158: Define "left-censored observations"

*Added (below detection limit) after “left-censored”*

Line 160: SW ~~phosphate~~ P043

Done

Line 164: SW- where orthophosphate was higher...

*Changed to PO43-.*

Line 165: SW-were also identified as the baseline...

*Done*

Line 168: SW- orthophosphate

*Line removed with revision.*

Line 172: The document reports that the current stations were established in 2005, and implies there are earlier samples being considered that predates the current stations. Doesn't there need to be some description of them?

*Water quality stations were established in 2004 and this was reflected in analyses; nutrients were not added until 2005. To clarify, we revise:*

*‘Nutrient samples (2005-2015) and water quality parameters (2004-2015) were evaluated using descriptive statistics to interpret changes over time in relation to each event…’*

Line 182: SW- orthophosphate

*Changed to PO43-.*

line 219: Figure 3 seems to indicate the spike was over 125 NTU, more than the 110 reported on this line

*This was in error; the spike was measured as 129 NTU. Sentence was revised.*

Line 216: It seem the reporting starts jumping around here between parameters and timing related to the rainfall (how long it lasted) and when the breach occurred. The following is some SW to provided order and be consistent with Figure 3 –

Precipitation began on March 27, 2005 and was heaviest during March 31-April 1st, producing a total of 43.2 cm of rainfall. Prior to the storm, salinity at BL was 20-21 psu, then fell during the event and reached its minimum of2.8 cm on April 2nd, one day after the storm ended (Figure 3). Salinity thereafter started trending up, which included some spikes during the first week of recovery. The levee at MPC was breached on April 14th, two weeks post-storm. Salinity ranged from about 7 to 12 psu during, and the days immediately following the breach, as it continued its upward trend mostly unaffected. Turbidity at BL was normally < 10 NTU before the storm, but surged to -100 just prior to the storm, fell to normal concentrations during the rain event, then spiked to 125 NTU immediately following the storm. However, despite some reoccurring spikes, turbidity quickly declined over the next several days, after which it usually remained< 20 NTU. Like salinity, it seemed to be unaffected by the breach (Figure 3).

Unlike salinity and turbidity, pH changed significantly following the breach. The BL water quality station is sufficiently shallow...(rest of paragraph)...returned to pre-spill levels on April 19•11, five days after the spill.

*Paragraph was revised to include the above.*

Line 232: SW – Like pH, orthophosphate was significantly impacted by the breach. At BL on April 25th, 11 days after the breach, PO43- averaged 4.29 mg P/L (Figure 4). It decreased gradually to.. (rest of paragraph).

*Changed, ‘it’ replaced with ‘orthophosphate’.*

Are there any numbers to indicate what P043-concentrations were prior to the breach; before, during, and/or after the rain event. In the current draft, nothing is included (like for other parameters) regarding what effect the rain had, which leaves the reader to wonder if the high P043 concentrations were just part the result of lingering freshwater influence (low salinity). If no data is available, maybe it is enough to more specifically point out somewhere in this paragraph that, since salinity was impacted (lowered) by the storm, but pH was unaffected, AND pH was impacted by the breach, but salinity was not- therefore, it must have been P043-that caused the pH changes related to the breach. [Uh, I wouldn't really use those words, but I hope you know what I mean and the author can create the appropriate inference]

-

*Unfortunately, we do not have enough nutrient sampling prior to the breach to fully address this question but we agree with the reviewer’s logic. One sample exists prior to the rainstorm which measured PO4 below detection and an appropriate sentence was added.*

*‘Only one nutrient observation was available prior to the rain event, which measured PO43- below detection.’*

*We have also added the following to the first paragraph of the discussion:*

*‘These water quality changes are likely a direct cause of the levee breach rather than freshwater inputs from the rain event. Salinity decreased with precipitation, whereas dramatic pH changes did not occur until April 14th when high concentrations of PO43- were also observed.’*

Line 237: As per Table 1, the beginning of E1C is July 2006, not April as indicated on previous line

*The reference to E1C was removed as this was incorrect.*

Line 240: Should it be ‘period through June 2006’? Concentrations decreased for 2 months at the beginning of ECl in July, then spiked...

*Yes, this was changed.*

Line 243: SW- As per the seasonal Kendall tests, the reduction in P043 from the beginning to

-

end of.... (Table 3).

*Changed to:*

*‘For the trend tests, the reductions per year in PO43- during the first event were a decrease of approximately 236% at BL and 80% at BN (Table 3, percent change based on slope per year divided by median).’*

Line 244: Can it be assumed the pre-event concentrations were at or below the detection limit at all four stations,then the reduction from the maximum value to that at the end of E1C was nearly 100%. If so, then probably worth stating.

*As noted above, we only have one data point before the breach. Based on several years of data between events that were below detection, we believe the baseline was probably at or below the detection limit, but we can’t back it up with data from the time period. We have added a sentence in the discussion to clarify this point.*

*‘The reduction of PO43- concentrations to levels near or below detection limit during the non-impact years suggests background concentrations prior to the first event were at similarly low levels.’*

Line 246: Figure 3 seems to indicate the average sea level is about O.5m, so the surge would be about 1.5m?

*The figure indicates ‘sensor depth’, not ‘sea level’. This was changed in the figure legend for clarity.*

Line 246: SW- rain and approximately 1.5m of storm surge at BL was produced...

*Changed.*

Line 249: SW- affected. A steady increase in turbidity from ~ 23 to over 100 NTU was observed with the storm surge. As the surge receded on August 31st, the observed pH and salinity decreased to a minimum of 6.5 and 5.6 at 18:30, respectively. After the storm passed, pH remained low, varying between 6.5 and 7.0 from August 31st to September 3rd, with occasional increases to 7.8. After September 13th, pH observations stayed above 7.0 and by September 31st, one month after the initial storm surge, pH had nearly attained its pre-storm level (Figure 3). Salinity after reaching its post-surge low began to steadily increase, with daily tidal variations that could exceed 5 psu. Like pH, one month after the initial storm surge, salinity also had nearly fully recovered. Also like pH and salinity, turbidity rapidly decreased after the storm (< 10 NTU), but unlike them, it quickly returned to pre-storm conditions a day after the surge.

*The paragraph was revised.*

Line 350: SW – ~~inorganic~~

*Removed.*

Line 260: sw-At BL during this same time period, average daily pH also was increasing, reaching 9.6 on March 2nd, concomitant with an elevated, albeit declining salinity condition typical of late winter (Figure 2).

*This was changed to: ‘**At BL during this time period, pH was also high, with peaks of 9.5 on February 16th and 9.6 on March 2nd.’*

*We did not use the phrase “average daily” because this is referring to 15-minute values. Salinity was not mentioned either because it was actually lower at that March 2nd peak (8 psu).*

Line 261: New Paragraph? SW- In contrast to the other 3 stations, changes in monthly P043 at BC were small and short-lived (Figure 4). However, a more delineated temporal analysis of iSCO diel samples collected at BC in 2012 depicts a stronger c4ange in P043-pattems before and after the storm (Figure 5). Orthophosphate...storm, when concentrations were near detection at low tide, but increased to 0.2.....storm. (Figure 5)

-

*New paragraph was added and the text was modified.*

Line 267: New Paragraph? SW- Even though the duration of the two events were the same, they were different in the amount of P043 recovery toward pre-event conditions. The reduction in P043 from the beginning to the end of the second event, based on the Kendall tests (Table 3), was a decrease.....135% at BN. In addition, the P043 concentration at the end of the second event was greater at three sites compared to the first event, despite having a lower maximum value at two of them (BN and BL) when E2A began. These differences and slower recovery were related to rainfall, with the post eyent years of 2013 and 2014 being very wet years compared to 2006 and 2007 after the first event (Figure 2).

-

-

*We did not include the last two sentences of this suggestion for two reasons. The initial orthophosphate value was sampled more days after the second event relative to the initial value for the first event and it is not valid to compare the initial observations between events. Second, we don’t have sufficient evidence of rainfall being a key factor related to recovery for the second event.*

Line 271: This 1st and 2nd sentences of the paragraph needs clarification. Not sure what stats test was performed to claim BN concentrations were greater for the entire period of this evaluation, and also enabled hierarchical ordering of others. Just looking at Figure 6A is not enough.

*The relative averages/medians can also be verified from Table 1. A citation to the table was added to the first sentence.*

Line 274: SW- orthophosphate

*Changed to PO43-.*

Line 275: If omit 1st two sentences of this paragraph (above comment), then include reference to

Figure 6A at end of next sentence...statistical different (Figure 6A).

*See response to above comment.*

Line 276: SW- orthophosphate

*Sentence removed with revision.*

Line 278: SW- periods for BN, BL, and the first event at PC. Comparisons within time periods showed substantial...

*Changed.*

Line 280: SW- ….orthophosphate concentrations (Figure 6B). Interestingly, orthophosphate concentrations at BL and BN were.....

*Changed to PO43-.*

Line 287: Reclassify Figure S2 as Figure 7

*Done.*

Line 288: Move the sentences in this paragraph to the end of the previous paragraph (Line 285), such that SW- remaining sites during the non-impact periods. The highest median P043

-

concentration for BC occurred in ElA. However, none of the time periods were significantly different....(Figure 6A). At PC, there were....E2C time period (Figure

6B).

*Changed.*

Line 292: Bring down the results of the other parameters that was started on line 286 to merge with this paragraph, with some rearranging of sentences to improve flow. For example,this new paragraph SW -Within time periods, there were no significant differences between sites in N02-IN03- (Figure7), but concentrations were significantly greater during ElA than E2A at BLand BC (Figure St-8). For NH/, there was no difference between sites during time periods, except BN and BC were greater than the other two sites during Nil (Figure7). Seasonal Kendall trend tests for NH4+ during sampling within time periods indicated reductions within sites, although none of these changes were significant (Table4). For chlorophyll-a, there were no significant differences between sites for any time period (Figure7). An increasing trend was generally indicated at most sites during advancing time periods (Figure Sl 8), although none were significant. However, the seasonal Kendall tests of samples within time periods showed a significant positive change at all sites during NlA, ranging from 32% to 56% (Table st-5). Increases were also observed during the second event (E2A and E2C), but these were not significant.

*The paragraph was similarly modified and includes updates after removing autocorrelation. Supplementary tables were also added to the main text.*

Line 300: SW- orthophosphate

*Changed to PO43-.*

Line 301: SW- characterized by extremes in P04 3- concentrations.. .

*Sentence was changed: ‘The first event was characterized by extremes in PO43- concentrations and both low salinity (< 3) and pH (<5).’*

Line 302: Is it similar trends or values in salinity and pH?

*See response to previous comment.*

Line 325: Define DIN. sw-DIN concentrations (dissolved inorganic nitrogen = N02-/N03- +

NH4+) have been...

*Added.*

Line 333: What ratios and where were they reported in this ms that were comparable? Also, no results were provided for DIN, just NO2-/NO3- and NH4+ separately.

*This sentence was revised for clarity. The ‘comparable’ qualifier applied to the monitoring stations used in the cited study and ratios for fertilizer used in the study region.*

*‘A similar study showed that water quality stations near a fertilizer plant in the Kavala Gulf, Greece had N:P ratios (2.5) that were comparable to fertilizer used in the region (3.2, Sylaios et al. 2005).’*

*We explain later in the paragraph why we did not evaluate N:P ratios.*

Line 334: SW- These studies are similar to the results of Grand Bay, such that elevated nitrogen (NH4 and N02-/NO3-) also was observed with elevated P043- concentrations (Figures 6-8).

*Changed.*

Line 337: DIN results not reported in this ms.

*Changed to ‘nitrogen’.*

Note: In the discussion, prior to Line 365: Include more development of spill consequences on ecosystem and biota. In the Introduction, reference is made to Viskup (lines 65-67) and the damages of $2 million. Perhaps pull that sentence from the Introduction and save it for the Discussion where it can be further developed (types of damages and correlation to events), along with potential thresholds. Compared to Rekik 2012? Same for the MDEQ 2015 report on fish kills. Do these reports/results have management implications which can be provided in the following section?

*As mentioned in the ‘management implications’ section, the acute effects of these spills are clear. An obvious but not always preventable solution is to keep these spills from ever occurring as there will be immediate, negative impacts. We realize that prevention is not an absolute solution and spills are likely to occur in the future. Understanding the chronic impacts remains a research priority. However, the chronic, long-term effects are not clear and difficult to understand using only water quality monitoring data. We hope that our final paragraph has made this point clear, and more generally, that this manuscript adds to the limited research on phosphogypsum impacts in coastal waters.*

Line 578 Table 2: Tables and Figures should stand alone, so include (PO43-) since PO43- is used in table

*Added.*

Line 578: sw -ammonium and nitrate/nitrite (NH4+ and NO2-/NO3-)...

*Added.*

Lines 588 and 589: Use ortho, since that is what is used on the actual graph label and even later in this figure description (Line 592). Therefore, SW- pH, and orthophosphate for Bangs Lake, within Grand Bay National Estuarine Research Reserve. All...excluding orthophosphate, which was sampled monthly.

*Changed.*

Line 592: SW- categories (Table 1) in relation to...

*Added.*

Line 598: Earlier reported the rains began March 27th on Line 216

*The precipitation began on March 27th, but the heaviest rainfall occurred form March 31st to April 1st). The following was added: ‘…centered on April 1st, precipitation began on March 27th…’*

Line 601: SW- orthophosphate...since that is what is used on the actual graph label

*Changed.*

Line 602: SW - Distance of stations from spill increases from top to bottom panels.

*Added.*

Line 606: Since PO43- used as a graph label, SW- orthophosphate (PO43-)…

-

*Added.*

Line 607: SW- spill event on August 28, 2012 at Bayou Cumbest station, approximately 7 km

from the spill site. Elevated P043is observed

-

*Added date.*

Line 610: SW- time period. Boxes depict median and middle quartiles of the values, and whiskers depict the 5 and 95 percentiles. Outliers are represented as points. Black boxes indicate??? and green boxes???. Boxplots within each...

*Added, but whiskers depict 1.5 times the interquartile range. Also added ‘Boxes are shaded by relative magnitudes of the median.’*

Lines 615 and 621: Switch Figures S1(becomes Figure 8) and S2 (Figure 7).

*Done.*

Line 615: Then would be- FIGURE 7. Boxplot summaries of nitrogen and chlorophyll-a data for sites grouped by time periods. Boxes depict median and middle quartiles of the values, and whiskers depict the 5 and 95 percentiles. Outliers are represented as points. Color shading of boxes represent ??? . Boxplots within each...

*Added ‘nitrogen and chlorophyll a’, added boxplot descriptions as for figure 6.*

Line 621: Then would be- FIGURE 8. Boxplot summaries of nitrogen and chlorophyll-a data

for time periods grouped by site. Boxes depict median and middle quartiles of the values, and whiskers depict the 5 and 95 percentiles. Outliers are represented as points. Color shading of boxes represent???. Boxplots within each...

*Added boxplot descriptions as for figure 6.*

Figure 2.

• Date ranges (width) of green bars do not seem to be correct, compared to supporting text. For example, during 1st event's precipitation, the bar depicts a low rainfall period.

*The first bar indicates the date of the levee breech (April 11th). This was changed in the caption for clarity.*

• For Orthophosphate, indicate this a log scale, and on they-axis, show the values that correspond to the guideline between .01and 0.1, between 0.1and 1.0, and especially for the guideline above 1.0.

*Caption was changed to indicate ‘log-space’. The minor grid lines were removed.*

• On bottom x-axis, show more than every 5 years; preferably every year, with additional hash marks for months. It currently is hard to follow the changes reported in the text related to time periods.

*X-axis was changed to show grid lines every year.*

Figure 3.

• Same issue with the green bar...please check if accurately depicts date range of events.

*As noted above, the green bar is centered on the period of highest precipitation.*

• Add units for Salinity

*Journal formatting specifies no units for salinity.*

Figure 4.

• Same issue with the green bar...please check if accurately depicts date range of events.

*Caption was changed to match info about the green bar from Figure 2.*

• For these graphs, indicate they are on a log scale, and on the y-axis,show the values that correspond to the guideline between .01and 0.1, between 0.1and 1.0, and especially for the guideline above 1.0 for BN and BL.

*Caption changed to indicate log-scale. Y-axis breaks changed to show maximum.*

• As above, on bottom x-axis, show more than every 5 years; preferably every year, with additional hash marks for months.

*X-axis now shows every year.*

Figure 5.

• Add units for Salinity.

Journal format specifies no units for salinity.

• On X-axis indicate this is a time line,provide labels:

2012 Dates

(Time: EST)

*Information was added to figure caption.*

Figure 6.

• For these graphs, indicate they are on a log scale,and on the y-axis, show the values that correspond to the guideline between .01and 0.1, between 0.1and 1.0,and above 1.0 where applicable.

*Added ‘log-scale’ to the figure caption. The minor grid lines are not labelled because the formatting does not work in log-scale.*

• Change "time frames" to "time periods," since that is what is used in text.

*Changed.*

• 6B's lettering to denote significance is opposite of all other graphs. Reverse it, so that "a" represents the greatest values, "b" the next, and so forth.

*Changed.*

• As mentioned above in Figure Legends, what is the meaning of the black and green boxes?

*Description added to figure caption.*

Figure 7 (Was S2)

• Indicate the graphs that are on a log scale. For those graphs, on they-axis show the values that correspond to the guideline between those that are numbered.

• What is the meaning of the black and various shades of green boxes?

*See above comments.*

Figure 8 (Was Sl)

• Indicate the graphs that are on a log scale. For those graphs,on they-axis show the values that correspond to the guideline between those that are numbered.

• What is the meaning of the black and various shades of green boxes?

*See above comments.*

*Autocorrelation plots for each station and nutrient parameter, up to twenty lags.*

**