State University of New York
COLLEGE OF ENVIRONMENTAL SCIENCE AND FORESTRY
One Forestry Drive, Syracuse, NY 13210 USA

Department of Environmental Biology

Professor Sean Crowe
Department of Microbiology and Immunology
University of British Columbia

30 August, 2023

Dear Sean,

Thank you for the opportunity to revise our manuscript titled "Marked recent declines in boron in Baltic Sea cod otoliths – a bellwether of incipient acidification in a vast hypoxic system?". We appreciate the thoroughness of the reviews and thank the reviewers for the comments. Below, I have pasted in the review points (blue), our responses (red), and **revised text** in black Times New Roman font.

-- *** --

The manuscript entitled 'Marked recent declines in boron in Baltic Sea cod otoliths – a bellwether of incipient acidification in a vast hypoxic system?' is an interesting study focused on temporal trends in otolith boron concentrations and how they relate to water chemistry, salinity, and pH.

Overall, the manuscript has clear objectives and the subject matter is quite relevant. However, I have several issues with the manuscript that I feel need to be addressed and therefore I recommend this manuscript be revised and resubmitted for review. Specifically, there several instances in which statements are made that are beyond the scope of the study. There are also several instances in which an analysis is reported in the results section yet no description of it was provided in the methods section.

My detailed suggestions are listed below.

Abstract:

Line 14: remove the following sentence 'B:Ca is positively proportional to pH in carbonates, as B in the form of borate is taken up in the CaCO3 matrix' Done

Line 19: remove 'healthy'. Done

Line 21: What does anti-correlated mean? Does it mean negatively correlated or no correlation? It means negatively correlated. We changed the wording to "negatively"

Line 23: Making a statement about the 'decomposition of large algal blooms' is beyond the scope of this study and should be removed. We feel this is a valid linkage, and would prefer to retain the statement.

Introduction:

Line 42: Sentence is too long. I suggest revising it into 2 sentences. For example, 'Despite buffering from alkalinity sources, pH is highly variable in space and time with its recent tendency to decline at greater depths in many parts of the Baltic Sea (citation). This decline in pH at greater depths is associated with the ongoing eutrophication and higher vertical export of organic matter that leads also to worsening deoxygenation (Kuliński et al., 2022).' Done

Line 46: Remove 'past' from this sentence. I would prefer to retain the word "past," because the fishery is closed and I don't want to imply that this is ongoing. (Now on line 47)

Line 49: Can you provide more detail about 'worsened [body] condition'? For example, was it length to weight ratio? Indeed, that's correct. We revised the sentence as "Cod exposed to hypoxia were shown to have lower growth rates and worsened **body** condition, **quantified by the Fulton K index of weight:length ratio** (Casini et al., 2016; Limburg & Casini, 2018)." (lines 49-51)

The 2nd paragraph could use some background information before launching into otolith chemistry and hypoxia. Below I provide a suggestion. 'The population of the Eastern Baltic cod (Gadus morhua, hereafter referred to as EBC) has been severely impacted by a number of factors that include past overfishing, hypoxia, parasite infections, and seal predation (Eero et al., 2020), leading to reduced growth and age at maturity (Eero et al., 2016; ICES, 2019). For example, EBC exposed to hypoxia were shown to have lower somatic growth rates and poorer body condition (Limburg & Casini, 2018, 2019). Maybe a sentence or two citing other studies that have shown negative links between hypoxia and fish health (e.g., somatic growth/body condition), then the third paragraph could focus on hypoxia and its measurement in otoliths. Thank you for the suggestions

3rd paragraph: Recent studies that reported links between hypoxia and fish body condition have relied on the measurement of manganese concentrations in otoliths (ear-stones) (Limburg et al., 2015). Briefly, otoliths, the calcified structures that form part of the hearing/balance system in teleost fishes, are sectioned and then analyzed by ablating micro-transects along the major growth axis and analyzing by mass spectrometry (see Methods). grow continually by the deposition of layers of calcium carbonate and protein, and as these layers form, ions are deposited onto the otoliths' growing surface often at concentrations that reflect those in the environment. The resulting elemental data are lifetime concentration histories, being incorporated at the time of exposure. The use of otolith manganese to track hypoxia exposure is one of the emerging biomarkers in fisheries ecology (Reis-Santos et al., 2022).' We thank the reviewer for the suggestions. We have re-structured the 2nd and 3rd paragraphs as suggested. They now read as follows (lines 46-62):

The population of the Eastern Baltic cod (*Gadus morhua*, hereafter referred to as EBC) has been severely impacted by a number of factors that include past overfishing, hypoxia, parasite infections, and seal predation (Eero et al., 2020), leading to reduced growth and age at maturity (Eero et al., 2016; ICES,

2019). Cod exposed to hypoxia were shown to have lower growth rates and worsened body condition, quantified by the Fulton K index of weight:length ratio (Casini et al., 2016; Limburg & Casini, 2018).

Recent studies that reported links between hypoxia exposure and body condition have relied on the measurement of trace elements in fish otoliths (ear-stones). Briefly, otoliths, the calcified structures that form part of the hearing/balance system in teleost fishes, are sectioned and then analyzed by ablating micro-transects along the major growth axis and analyzing by mass spectrometry (see Methods). Hypoxia exposure was quantified by means of a proxy developed from analyzing manganese concentrations in otoliths (Limburg et al., 2015), in ratio to magnesium which corrects for growth influences on Mn uptake, since both are affected by growth but only manganese is affected directly by hypoxic conditions (Limburg et al., 2018; Limburg & Casini, 2018). The resulting elemental data are lifetime concentration histories, being incorporated at the time of exposure. The use of otolith manganese to track hypoxia exposure is one of the emerging biomarkers in fisheries ecology (Reis-Santos et al., 2022).

Line 50: This sentence has a lot of information and is somewhat confusing, especially 'ratio to magnesium which corrects to some extent for growth influences on Mn uptake'. How does it correct for growth? We tried to clarify in the third paragraph (text above, and repeated here):

Hypoxia exposure was quantified by means of a proxy developed from analyzing manganese concentrations in otoliths (Limburg et al., 2015), in ratio to magnesium which corrects for growth influences on Mn uptake, since both are affected by growth but only manganese is affected directly by hypoxic conditions (Limburg et al., 2018; Limburg & Casini, 2018).

Line 60: I suggest adding your preliminary findings after this sentence. 'Recently, with acquisition of more sensitive instrumentation, we began to experiment with quantifying elements having sub-ppm concentrations in cod otoliths, including the trace element boron. **Interestingly, preliminary results from this work showed elevated levels of boron in 2000 but not in 2019. Boron is noteworthy because it is an indicator of salinity. Specifically,** Boron in seawater generally correlates with salinity (Kuliński et al., 2017) predominantly in the form of weak boric acid (H3BO3) at standard seawater salinity (35 PSU) and pH of 8…. '. We added the suggested wording (bolded above) to the paragraph (lines 66-68).

Following the above text I would then put a new paragraph that describes your study. 'In this study, we explored the extent to which otolith B:Ca varied through time and whether its values were correlated with pH, salinity, or other otolith derived values that are proxies for environmental factors.' We kept the next paragraph and figure (see next comment) intact, and added a new paragraph (lines 85-91) before the questions that we pose at the end of the Introduction.

Line 74: 'It was therefore surprising when we began to analyze otoliths of EBC captured in 2019 and discovered greatly reduced boron concentrations (Figure 1(b)). The measurements were repeated for verification.' Figures are not typically referenced in the introduction. If the journal permits it, great, otherwise I suggest removing the figure and just describe that you observed elevated levels of B:Ca in 2000 but not in 2019. See note above.

Line 84 & 86: No justification is given as to why Mn:Mg and P:Ca are used. I suggest moving lines 146-153 from the methods to the introduction at line 80. Thank you for pointing this out. We've re-worked text to produce the final paragraph near the end of the Introduction to include both P:Ca and Mn:Mg (which was explained in lines 56-59). The new text (lines 85-91) reads as follows:

In this study, we explored the extent to which otolith B:Ca varied through time and whether its values were correlated with pH, salinity, or other otolith derived values that are proxies for **environmental or physiological factors. In particular, otolith P:Ca is known to be under physiological influence (Thomas et al., 2017; Heimbrand et al., 2020). In cod, higher values of P:Ca occur during the growing season and are thus hypothesized to be associated with growth and other activities. We selected this as well as our proxy for hypoxia exposure, otolith Mn:Mg,** and posed the following questions:

(followed by the questions)

Methods:

Line 91: When were the otoliths collected? Were fish of a certain size/age targeted? Answered in the next response

Line 91: Break into 2 sentences: 'Otoliths of Baltic cod (N = 156) were obtained from both fishery-independent and fishery-dependent surveys conducted by the Swedish Fisheries Board and its successor, the Department of Aquatic Resources, Swedish University of Agricultural Sciences. Fish were collected in the summers of 2000 to 2019 from ICES sub-divisions (SD) 24, 25, 27, and 28 (Figure 2). Otoliths were extracted, measured, weighed, and embedded in epoxy…' We did as suggested and broke this up. We also added a sentence to clarify about our selections (starting on line 104):

Fish were collected from ICES sub-divisions (SD) 24, 25, 27, and 28 (Figure 2), spanning the period 1988-2021. To construct the time series, we selected previously analyzed (but not for B:Ca) otoliths that were comparable in age range to those collected in 2019-2021.

Line 127: Additional details are needed about fish age and ageing techniques used; this text could be placed after the mention of otolith polishing. Also, was age and the location of annuli (ie otolith radii) determined by multiple readers? We added a short paragraph (lines 128-132) to explain about aging:

Baltic cod were aged by examining seasonal patterns in otolith Mg:Ca and P:Ca, following Heimbrand et al. (2020). Icelandic cod were also aged by otolith chemistry, although several of them had been previously aged visually by age readers in Iceland. For the latter otoliths, the seasonal chemical patterns produced age estimations that matched exactly with the visual estimates.

It sounds like that the ablation derived trace element data was pooled within each annulus and then an average for each trace element was generated. To do this you would first need to know the otolith radius that coincides with beginning and end of each annulus for every fish, and then bin the ablation data based on those otolith radii. Yes, this is correct.

Ablations were run across the entire otolith, so was the ablation data from both halves of the otolith (ventral and dorsal) averaged? In other words, was the trace element data corresponding to the first annulus pooled from the transect running through both the dorsal and ventral portion of the first annulus? We only used the dorsal axis for quantitative analysis. We often also lasered the ventral axis to check for symmetry. We added a new sentence to explain this (lines 145-147):

Although often edge-to-edge laser transects were made (to check for symmetry), only data collected on the dorsal (longer) axis were used in our analysis due to greater spatial resolution.

Line 131-135: Sentence too long. I believe that the reviewer didn't see that it is actually two sentences.

Line 138: More otolith growth occurs in the summer rather than the winter, and therefore trace element ratios averaged within an annulus correspond to summer more so than winter. If that is correct, shouldn't the water data also focus more so on summer time than winter? 'Annual mean water values were computed and matched to corresponding otolith chemistry data (annual means of Element:Ca data parsed to calendar years as described above).' The reviewer makes a good point; we cannot parse the data finer than per year. As it happens, there are also more water data collected during the growing season months.

Line 142: what R packages were used? We used base R and ggplot2; this is now mentioned in lines 163-164.

Line 143: 'Statistical analyses were separated into examination of relationships of B:Ca values to potential environmental drivers (primarily pH and S, but also AT, DO, and T) and to trace element ratios the internal variables P:Ca and Mn:Mg.' (no response needed)

Line 145: This sentence describes a result and should be moved to the Results section. We moved the errant sentence and accompanying figure to the start of Section 3.4, lines 242-247. The figure is now Figure 5.

Line 146-153: Move these sentences to the introduction to justify your reasoning for using P:Ca and Mn:Mg. Done. Line 149-153 is too long of a sentence and it should be broken up into at least two sentences. This is now two sentences and as stated above, is at the start of Section 3.4 (lines 243-247).

Line 154: define what you mean by 'graphical analysis'. This was removed. Also, specify that PCA was performed on water data and why was it used. Will the resulting PC factors be used in subsequent analyses? We explain as follows now (lines 167-169):

For the environmental drivers, a PCA analysis was first performed to examine relationships among them. Thereafter, B:Ca was regressed on the first principal component eigenvector.

Describe why you used repeated measures regressions and on what data you used. Hopefully, these next two sentences suffice to clarify (Lines 169-172):

This was followed by mixed linear models of environmental drivers on B:Ca, with each fish's age (nested within fish ID) being treated as a random effect. This analysis took account of repeated measures within the same individual, and was intended to examine more closely the effect of individual drivers.

Line 154: clarify what is the dependent and independent variables used in the mixed models. Please see the first line in the re-written sentences directly above. B:Ca is the dependent variable and the drivers are the independent variables. Hopefully that's clear now.

Line 158: Is there a reason why trends among decades were examined and not at an annual basis? Maybe it has to do with sample sizes, but please clarify. Although the data are annual, it is easier to resolve trends decadally, what with the noisy data in the Baltic Sea.

Results:

Line 165: specify that the first two components explained 91% of the variation in the water dataset (71% and 20% on component 1 and 2, respectively). This was an older analysis and should have stated 72.4% and 20.5% respectively. We have added a sentence (lines 184-185):

A principal components analysis of the water variables found that the first two components explained 92.9% of the variance (72.4% by PC1 and 20.5% by PC2).

Figure S2 should be placed on a separate page in landscape orientation to increase its size. As per a later comment on the Discussion section, we moved Figure S2 out of the Supplemental Information and into the main article. It is now Figure 3, and we have now put it on a separate page in landscape orientation, as suggested.

Figure S3 (now S2): I suggest overlaying the datapoints in each plot. Also, there isn't any mention of how otoliths were aged and by whom. We prefer not to show the noisy data, in order to focus on the general directions of the trends. The aging information is in the Methods now.

Line 174: remove words 'box plots of'. Done

Figure 4: remove 'within year' from y-axis label. Done

1. Define what the thick line, box, and whisker correspond to. The following sentence was added to the figure legend: Boxes represent the upper and lower quartiles, vertical lines the maximum and minimum values, and the horizontal lines are the median values.

2. For any given year of B:Ca in this plot, does the box and whisker correspond to the B:Ca from several individuals and potentially different annuli? Yes

3. B:Ca is plotted twice. I think B:Ca should be plotted once and both salinity and pH included as a line plot using the second y-axis. Time trends of salinity and pH, together with the other environmental drivers, are presented in the preceding section in the new Figure 3 (former Figure S2). We prefer the visual effect of showing the box plots colored by the salinity and pH, to bring together the data for greater understanding.

Table 1: Why was the PCA performed if the resulting PC factors, which summarize the water data into two varaibles, weren't used in subsequent analysis? I suggest removing the PCA from this paper. We appreciated this comment; many times, PCAs are used to understand relationships, and this was the case here too. We were somewhat surprised to see the near-orthogonal nature of temperature to the other four environmental drivers; this tipped us off that it would have less of an effect. That said, we also took the reviewer's advice and also regressed B:Ca on the PC1 scores. This produced a negative slope, suggesting that it follows the PC1 axis: higher dissolved oxygen and pH associate with higher B:Ca, and higher salinity and alkalinity associate with lower B:Ca. The re-written paragraph (lines 184-189) is as follows:

A principal components analysis of the water variables found that the first two components explained 92.9% of the variance (72.4% by PC1 and 20.5% by PC2). The analysis showed three groupings: (1) DO and pH that were opposite to (2) salinity and A_T, respectively; and (3) temperature was on a separate axis (Figure S1). Linear regression of B:Ca on the first principal component yielded a negative relation (B:Ca = 0.000485(PC1) + 0.00246, $R^2 = 0.21$, $p < 0.0001$).

Line 193: I'm confused. What kind of two factor tests were performed. These two factor tests were multiple regression models with B:Ca as the dependent variable and paired combinations of pH, salinity, and alkalinity to see which had a stronger effect when only two independent variables were used. This was to further assess the relative strength of their influence on B:Ca. We have tried to explain (lines 222-225):

Because of hypothesized positive relationships of salinity and pH on B:Ca, as well as the large effect size of A_T, we examined pairwise (two-factor) models of these independent variables on otolith B:Ca (Table 2). This allowed us to examine the relative strength and direction of trends (slope parameters, Table 2) on B:Ca when testing two independent variables together.

Line 195: clarify what you mean by the 'overall model'. I don't see this analysis described in the methods section. Please clarify. The regression option in JMP allows one to subset (filter) by specific portions of the data, but also to test the effect of all the data. We re-worded the sentence on lines 226-227 to read:

Aside from an overall model **using all the data**, these were filtered by decade to examine finer scale trends.

Line 211: What are the r-squared of figure 5 plots? 'Annual average P:Ca was highly positively correlated with B:Ca in Icelandic cod otoliths, but negatively correlated in Baltic cod with a great deal of scatter' We have added R^2 and p-values to the two plots in the figure. The figure is now Figure 6.

Line 217: Remove 'Box plots' from this sentence. We chose to retain this for clarity.

Line 218: I don't see an ANOVA described in methods. Please add a description of how this analysis was used to the methods section. A sentence was added on lines 179-180 in the Methods:

These were tested using ANOVA with slopes as the dependent variable and decades (for Baltic Sea) and Iceland as grouping variables.

Line 225: 'We used a similar analysis of individual fish regressions of Mn:Mg, our hypoxia exposure proxy, on B:Ca, as was performed with P:Ca.' This should be described in the methods section. This is already in the Methods section. Hopefully our re-writing improves the clarity.

Discussion:

Line 247-249: remove this paragraph because it is irrelevant to the study. Done

Line 257: I suggest more details: 'In this study, univariate regressions of otolith B:Ca on five abiotic water variables all were highly significant, with dissolved oxygen and pH showing positive relationships while negative relationships with alkinity, temperature and salinity (Table 1).' We made the addition as suggested (lines 295-297).

Figure S2 is referenced quite a bit in the discussion and thus I think it should be moved from supplementary section to the main body of the manuscript. As mentioned above, it was moved into the manuscript as Figure 3.

Line 294: '….we can observe a regime shift and a clear decoupling…'. These are substantial statements about the environment changing. Conclusions about regimes shifts are not within the scope of the study and thus I don't feel they are needed. I would just rephrase these sentences to talk about trends and/or variability. We have re-worded and replaced "regime shift" with "change in trends" which we hope is satisfactory (line 333).

Reviewer 2 comments, and our replies

Summary:
The authors examine B:Ca changes in cod otoliths across a 30+ year period in the Baltic Sea where there is a good record of environmental data that are hypothesized to drive observed ratios. The article was generally well written with appropriate references. With a few minor to moderate clarifications I suggest below in the Concerns section, would make an excellent addition to the wider otolith chemistry literature where researchers are diving much deeper into the meaning of many of the markers used for stock discrimination and uncovering useful relationships with environmental variables. The Comments section are mostly editorial changes that help readability or alternative words that create clarity.

Dear Reviewer 2, thank you for your vote of confidence, and for your comments below. I will address them in blue.

Concerns:
Line 126: The spot size (110 μm) is large and travel speed (7 $\mu m \cdot sec^{-1}$) is slow relatively speaking. How many data points is the ICPMS collecting across the full transect? Many of these points are going to overlap, though the repeated measures analysis used does account for this by nesting age within fish. Figures 1 and 3 both indicate much more variation than I would have expected using this combination of spot size and travel speed. My expectation would have been a much more muted record with smaller peaks and valleys.
Reply: we run a continuous transect rather than point analyses. The number of replicates depends on the travel speed and the number of analytes we run. I would say we typically collect 100-300 replicates per transect. Some researchers collect a lot more points, but we set the mass spectrometer software to do a little smoothing before reporting data; and as pointed, out, there is some smoothing from some overlap. Many researchers smooth their data with running averages; that would also result in smoothing. That was not done in this analysis.
I have run comparative analyses in the past, testing whether I obtain better results with discrete point analyses vs. continuous transects. In terms of the quantitative results, there is little difference; but setting up the laser to run points results in a noisier transect. I've also tested running at different speeds and spot sizes. While obtaining more detail at slower speeds, there is the trade-off of analysis time (and money). Thus, we have settled on the spot size to collect as many detections as possible, and a reasonable speed.

Line 142: The analysis section could be clarified somewhat. While I understand the mechanics of the comparisons made, I did need to read this section a second time when I got into the results. Line 155-158 are the bigger picture analyses. It would be helpful to highlight this analysis more by reordering the paragraph or adding to these lines so that it leaves a larger impression later

in the manuscript.

Thank you. We have re-organized the Introduction and Methods for greater clarity, following suggestions of Reviewer 1.

Line 135: How much does the selection of different water depths to average (pre and post 1995) impact the results? I presume cod were found deeper post 1995. If just cod from overlapping depths (40-60 m in this case) were chosen, would the differences have been as dramatic? While 10-15 m is not much in the open ocean, in the nearshore environment, small changes can dramatically affect chemistries due to terrestrial inputs. Figure S2 illustrates this concern well. While the 40-75 m and 30-60 m annual averages follow similar trends, there are times where the differences are large particularly across the recent 2000-2020 period.

This is a good question. We selected data from the Gotland Deep monitoring station in the Central Baltic as a representative site. We did not perform tests on the relationships of otolith B:Ca to water variables at a single depth across the time series. However, I will note that in recent years, as reported both by Casini et al. (2021) and by Almroth-Rosell et al. (2021, "A regime shift toward a more anoxic environment in a eutrophic Sea in northern Europe," Frontiers in Marine Science 8: 799936), the depths of the hypoxic and anoxic layers have been rising, particularly in the last ca. twenty years.

Line 185: Why do the degrees of freedom differ in table 1? I would have expected df to be the same for each comparison. Are some of the annual environmental parameters missing? While an exact accounting of missing data isn't necessary, the existence should be stated in the methods.

Yes, some of the water variables have missing data relative to other water variables. We can note this in the table legend.

Line 247: This paragraph seems to be an orphan in the discussion. It is useful information and would fit better in the paragraph beginning on line 60.

We have removed the "orphan paragraph" at the suggestion of Reviewer 1.

I did not read the Journal guidelines but I do not like a heavy reliance on supplemental information. It's certainly not a deal-breaker and I do understand page limitations for print. My preference would be for greater inclusion in the manuscript or simply writing with less reliance on supplemental figures.

At Reviewer 1's suggestion, we moved Figure S2 into the main text.

Comments:

Line 21: anti-correlated should be inversely correlated

Done

Line 39: built could be composed

Done

Line 151: reducing instead of reduced

Done; the sentence was moved to address a comment by Reviewer 1

Line 234-237 should be moved to the discussion

Now moved to the end of the Discussion.

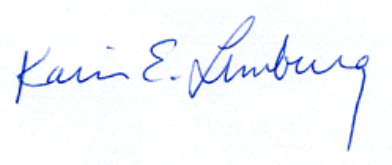
-- *** --

As you can see, we did a substantial reorganization of the manuscript and supplemental material. For clarity, I'm including here a table to show the reorganization of figures:

Original Figure Number	New Figure Number
1	1
2	2
3	4
4	5
5	6
S1	S1
S2	3
S3	S2
S4	S3
S5	S4
S6	S5

We hope that the revision will now be sufficiently clear that it warrants acceptance, but of course we can do more if needed. Thank you for serving as editor, and we thank the anonymous reviewers for the careful reading and commenting on the original draft.

Sincerely,



Karin E. Limburg,
Distinguished Professor
Email klimburg@esf.edu