

## Reviewer 2 comments, and KL 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  $\mu\text{m}$ ) is large and travel speed (7  $\mu\text{m}\cdot\text{sec}^{-1}$ ) 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.