Author Note:
We thank all referees for their insightful and constructive comments on our manuscript “A high-resolution synthesis dataset for multistressor analyses along the U.S. West Coast.” We appreciate the opportunity to incorporate and respond to these thoughtful comments and improve our manuscript. Below, we discuss the comments from Reviewer 2. We have included all original comments, with our response to each point raised bulleted below.

Rev 2
This paper documents the development of a large dataset of observations of dissolved oxygen, pH (and other carbonate chemistry parameters), and temperature in addition to a few other low priority ad hoc variables (e.g., nutrients). The dataset will be very useful to the broader scientific community and the paper is generally well-written. I inspected the data posted the public repository and it is in excellent shape. I think the paper is ready to be accepted after some minor comments listed below.

Major comments

The only thing approaching a major comment is that I got confused about the number of observations in the data. At one point in the methods, it sounds like the aggregation of the data into daily averages reduced the dataset from 12.7 million rows to 1.2 million rows but then in the results it sounds like there are 12.7 million rows and the aggregation wasn’t done. Please be very careful about this and report accurately how many observations are in the final (data available to user) dataset and propagate throughout.

- Thank you for bringing our attention to our confusing wording. We have now explicitly defined an “observation” in our Methods section to be a row in our dataset, in which one or more individual parameter measurements will be associated with a dataset, date, time, and location (lines 191-196). We have further replaced “measurement” with the more specific “individual parameter measurement” to make it clear when we’re discussing the amount of data associated with a single parameter (e.g., 13.7 individual temperature measurements, but only 3.3 million individual dissolved oxygen measurements). Additionally, we have clarified the Results section discussions of the aggregated daily dataset by breaking the previous section “3.1 Overall Data Totals” section into two smaller sections, “3.1 Overall Measurement Quantity” and “3.2 Aggregated Daily Data Totals.”

The metadata table (MOCHA_dataset_metadata_table.csv) is easy to understand and seems largely complete though dataset 2 does not have a name. The other fields missing data make sense.

- We have fixed the error with dataset 2. We are pleased to hear that the dataset metadata table is easy to interpret!

I examined a subset of the data ("47_to_49_pre_2015.csv") and it is in great shape. The column names are all super intuitive and the values in the columns are all correctly formatted. All of the data that you would expect to be complete is complete. Wonderful.

Minor note for future submissions: please use continuous line numbering (not 5 line intervals). Do everything you can to make the reviewers job easy – this will keep them happy!

- We appreciate this feedback and have passed it on to ESSD for potential incorporation into their manuscript templates.

Minor comments

- Unless otherwise indicated or discussed, all of the following minor comments and suggestions have been fully incorporated. We appreciate the reviewer's close attention to detail.

Abstract

24 – Stressful or favorable, plus what's stressful for one organism might not be stressful for another

31 – could you work the focus on hypoxia and ocean acidification risk a little earlier in abstract?

32 – stats on the time span of observations should get mentioned

Introduction

43 – could shorten “effluent from coastal settlements and agriculture” to “coastal runoff”

43 – it’s not clear to me the mechanism for “diverse and highly productive ecological communities” to drive local deviations from global patterns

- Here, we were referencing how local biomes like seagrass meadows and kelp forests can significantly alter the local chemical environment (e.g., Ricart et al., 2021). We have changed this phrase to “high local productivity”.

52 – “e.g.” is missing a comma after it (like in line 40); ensure comma is added throughout

61 – Free et al. (2023) (https://doi.org/10.1111/faf.12753) provides an update to Cavole et al. 2016 paper and is explicitly about this region

71 – Can you make it clear that conditions have gotten shallower without using the word “shoaled”? It might not be familiar to everyone.

89 – What regions do they apply to?

- The specific references to CenCOOS and SCCOOS have been removed..

97 - “…for the CCS and is newly archived and available at…”

102 – Hoping to see that the unincorporated sources of info get mentioned later

107 – no need to capitalize MPAs

108 – The stats on number of observations, sources, and time span should get mentioned in last paragraph of info

Methods

119 – how was the literature search conducted?

- We did not do a formal literature search to find datasets and meant here to refer to sourcing some datasets from published literature. More accurately, we accessed public data portals and federal government datasets, contacted colleagues to request their assistance in locating datasets, presented the project

at conferences for three years requesting community participation in the project, and completed a scan of published literature that likely included published datasets that could be incorporated into the project. While this is not considered exhaustive, one benefit of the NCEI data platform is that we can continue to update the available data as we become aware of and process new sources. We have updated our description of sourcing datasets to this more detailed description (lines 150-153).

138 – suggest adding (1), (2), (3), and (4) here to orient reader

153 - suggest adding (1), (2), (3), and (4) here to orient reader

158 – this should at least be a supplemental table in this paper; its annoying to have to go look elsewhere for info on the dataset documented in this paper

- The “dataset metadata table” has now been included as a supplement in addition to being available at NCEI.

191 – What does “as normal” mean here?

- This phrase has been altered to “quality controlling them further following the practices described for all other incorporated datasets.”

199- suggest adding (1), (2), (3) here to orient reader

205 – Can you give examples in the supplement? Reads as vague now

- We have added an example of our formatting and flagging practices as supplementary information and have made the code and data associated with it available on our project Github repository (github.com/egkennedy/DSP_public_code).

210 – Examples drawing from this would be useful

- We believe this request has been addressed through our supplemental flagging example.

224 – “i.e.” should be followed by comma – correct throughout

Results and Discussion

244 – Again, time range would be helpful.

244 – Isn't 12.7 million incorrect? Didn't you reduce down to 1.2 million by aggregating to daily level as stated in Line 226. I'm skeptical of all the sample sizes reported here b/c of this.

- Following our response to the reviewer's major comment above, we have revised this paragraph to make it clearer that we are discussing the full, disaggregated data set here, rather than the aggregated dataset we used for oceanographic interpretations. The aggregated dataset is now discussed in its own subsection just below.

273 – "malfunction, 2)"

273 – I think either means between two options

Conclusion

474 – No need to capitalize MPAs

Tables and Figures

Figure 1. The figure would be more useful if it showed the density of points along a raster grid (potentially hexagonal) so that the reader understands data density spatially. The panels should all be the same size, 1 row, 3 columns would be an improvement. The density could be the number of points within a cell or the number of unique year-months in a cell. I leave it to the authors.

- We appreciate this suggestion. There is inherent tension between showing every available data location and the data density over space or time. We believe both are valuable, but have kept this figure showing all of the individual data locations since we show the spatiotemporal data density more clearly in Figure 3. We have taken the suggestion to make all maps in this figure the same size.

Figure 2. Y-axis is a proportion, not a percentage. Align the word choice with what is shown. Spell out acronyms in caption.

Figure 3. It would be nice if the panels were labeled with the parameter so the reader doesn't even have to read the caption. The width of the latitude should be stated.

Eyeballing the figure. Data looks to be most common between 2015-2020 and not 2010-2015 at the authors state, Bar plots of annual totals would be a good way to examine the temporal bias alone.

- We have added parameter titles to the plots and updated the caption to clarify that the spatial coverage of observations is most complete between 2010-2015, whereas the total number of observations is highest between 2015-2020.

Figure 6. The caption is confusing about what the points are. Are these all observations with 50 km of shore in the top 50 m? State what it shows. Currently, it's written like a results section. Define the lines but exclude all of this results interpretation.

Figure 7 caption also includes lots of results interpretation.

- Both the captions for Figures 6 and 7 have been rewritten to exclude results interpretation. Figure 7 has also been substantially revised and now replaces Table 3.

Figure 4. Y-axis should read "Percent of observations."

Table 1. Define acronyms in parameters column in caption. Consider making this a supplemental figure given its size.

- While Table 1 is long, it was important to us to give credit to the constituent datasets of this synthesis compilation so we hesitate to bury it in a Supplement. We have added a DOI/Citation column to the table and full references to all datasets to our References section. This has allowed us to simplify the titles in Table 1 and improve the readability of the table, though it is admittedly still quite long. We have defined the parameter acronyms as suggested in the caption.

Table 2. Add comma to 3rd column. Eliminate 2nd decimal spot in fourth column. Spell our Parameter acronyms in caption.

Table 3. This would be more compelling as a multi-panel figure of scatter plots with regression fits. Caption is mostly results interpretation.

- Thank you, we have replaced Table 3 with this suggested figure and removed the previous Figure 7.