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 3. We have included all original comments, with our response to each point raised bulleted below.

Rev 3
Kennedy et al. collate, quality control, and synthesize temperature, salinity, and biogeochemical data in the nearshore region of the U.S. portion of the California Current Ecosystem. This data product does show promise for addressing temporal and spatial variability and multistressor dynamics within this region, however, the associated manuscript does not provide enough information for a potential data user to fully understand the appropriate applications for the data product or how the data are manipulated. It also does not provide fair credit for the contributions of the original data providers and funders.

Major comments:

The authors claim the science applications of this data product are broad, including characterizing seasonal variability and spatial variability along the U.S. West Coast. However, the results illustrating variability only focus on the portion of the data sets within 50 km of the coast and < 25 m depth. Either the results need to be expanded to include analysis of the entire data product, or the data product should be restricted to the shallow, nearshore environment and the title and introduction should reflect that the product is focused on the nearshore.

- It is our intention to provide a data product that lends itself to a wide range of spatial and temporal scientific questions, rather than limiting our data compilation to the use of the specific case studies we discuss in this paper. We imagine future research using the MOCHA synthesis that defines its areas of interest via socioeconomic boundaries such as the U.S. Exclusive Economic Zone, bathymetric boundaries such as the region shoreward of the continental shelf break, ecological boundaries such as viable kelp habitat, or other bespoke regions that extend from the shoreline. We have added language to paragraph 5 of our introduction to frame our shallow, nearshore examples as examples, rather than exhaustive analyses (lines 120-126)

Given the data set itself is not interoperable or compatible with other products that include offshore biogeochemical data (e.g. gridded NetCDF files of the Surface Ocean CO2 Atlas or Biogeochemical Argo), it would be difficult for a user to combine and utilize them for assessment of biogeochemistry spanning offshore to nearshore.

- We appreciate the reviewer noting the lack of interoperability between many coastal and oceanic datasets. Even without considering interpolated gridded products, the difficulties of working with data from cruises, moorings, and shore samples was a large part of the motivation for developing this synthesis. Existing coastal syntheses such as SOCAT or CODAP-NA are highly specific – only surface CO_2 measurements in the case of SOCAT and only oceanographic cruise discrete samples in the case of CODAP-NA. These are excellent products, but very difficult to augment with the wealth of high resolution sensor data also available in the region. Sensor and hand-collected datasets, on the other hand, are currently primarily available through a bewildering variety of formats and databases with no standardized metadata, quality control, or data organization. Our work to bring these sensor datasets into a common format with oceanographic cruise observations was a significant endeavor that responded to a real need and will improve our ability to map and understand the coastal ocean. Our priority for this synthesis was to bring nearshore and coastal datasets together, but we have included data extending well beyond the continental shelf to allow investigators some ability to examine biogeochemistry from offshore to nearshore environments even if this synthesis is not fully compatible with existing offshore gridded products.

Given upwelling- and respiration-driven low pH and low oxygen conditions manifest first in bottom waters, the way these conditions are explained in section 3.5 as within 50 m of the surface is confusing. It would be more intuitive to assess these conditions in the entire nearshore water column based on a bathymetric definition of nearshore, rather than defining nearshore as 50 km from the coast.

- Thank you, we really appreciate the improved data visualization suggestion. We have remade Figure 5 with data from within the 100m depth contour, which strikes a good balance of data availability, distance from shore (99% of the data is within 30 km of the mainland, with all “outlying” data associated with the Channel Islands in California), and ecological considerations of “nearshore” environments (e.g., environments where appreciable light still reaches the benthos). Interestingly, the conclusions from this figure were very similar to those from the original figure since most data is coming from shallow, coastal moorings, but we agree that the bathymetric cutoff provides more intuitive support for these conclusions. Using a bathymetric cutoff for this figure also serves to show an additional way of interacting with the MOCHA dataset.

"Surface" and "near-surface" are used interchangeably, both defined in different parts of the manuscript as < 25 m. "Nearshore", "surface", and "near-surface" should all be defined early on in the results and used consistently throughout.

- We have replaced all “surface”, “nearshore”, and “near-surface” shorthand with explicit descriptions of the depth and spatial range.

The description (and potentially the application) of the secondary data quality control is inadequate. First, the original non-QC’d OOI data sets (section 2.3) need to be QC’d using recommended best practices specifically developed for OOI biogeochemical data sets (doi.org/10.25607/OBP-1865).

- Thank you, OOI published these recommended QA/QC practices just after our group worked with the data, so we were unaware of this publication. The OOI data was handled in close collaboration with OOI staff using code, best practices, and data cleaning techniques provided directly by them and now published in Palevsky et al., 2022. We have added references to these published OOI protocols to section 2.3 to clarify our actions around this data.

Second, the description of the QC for the remaining data sets (section 2.4) sounds qualitative as written, as if the QC’er simply looked at property-property plots and flagged data points that looked bad. What the authors consider an “outlier” needs to be defined. Were outliers identified as a certain number of standard deviations of the linear (or some non-linear) relationship between parameters? What were the criteria for identifying “suspicious observations” (line 214)? Data QC routines need to be well documented and applied consistently throughout the data product using statistical analyses and thresholds to characterize quality.

- Thank you for this comment. In response to the first point, that QC practices should be objective and statistical, we generally agree but add some additional context. Most data pulled into this synthesis had been published and at least subject to automated QC processing. For these published datasets, our further quality control role was akin to both the "human in the loop" and "comparisons among co-located data" steps of the OOI data's recommended best practices (Palevsky et al., 2022). Our secondary QC relied more on human judgment since automated QC practices are liable to miss clear instances of biofouling or significant sensor drift in automated sensors, as well as data that is unreasonable for the location, time, and depth while remaining within "normal" limits for the whole dataset. This is, by its nature, somewhat qualitative, but that is a response to the diversity of sampling schemes, observation frequencies, and habitats we were sourcing data from. The reliability of data associated with a tight time series dataset with samples every 20 minutes for 5 years is very different from a sporadic time series that includes data from three different seasons spaced across five different years, though both datasets can be effectively interpreted and quality-checked by a team with oceanographic expertise. In all cases, we opted towards data inclusion - flagging only data that was "unreasonable" or "unreliable" by the standards of that data set rather than a more aggressive stance. To clarify our QA/QC practices, we have added a flagging example to the Supplementary Information. This flagging example shows the raw data and previously published quality flags, the standard property-property and time series plots we used to double-check published data, and our changes and additions to the quality flags. All code and data associated with this example is fully available on our project Github repository (github.com/egkennedy/DSP_public_code).

It is also a best practice to state the constants used in carbonate chemistry calculations. In addition to Dickson et al. 2007, the authors should refer to more recent best practices for the use of constants in a broader range of temperature and salinity: doi.org/10.1016/j.marchem.2018.10.006; doi.org/10.1016/j.gca.2021.02.008; doi.org/10.1016/j.marchem.2014.07.004.

- Thank you for these reference suggestions. We realized that including calculated pH data in our manuscript provided little additional information and unnecessarily expanded our methods section, so we have removed all calculated pH data from our figures and discussion.

Lastly, the data products I am most familiar with all have a substantial acknowledgements section including funders of the observations, a long list of citations, and many coauthors because they include the major data providers in the data product development. At minimum, Kennedy et al. should include all the data citations in the list of references. That requires referring to the metadata for each of the original data sets and including a data citation in the references if the data provider requests one be cited. I see citations provided in a table within NCEI Accession 0277984, but that is not trackable by the data providers. Those data citations are critical metrics that funders use to make decisions about what observational programs to support.

- We completely agree, and apologize that our misunderstanding of what could be included in the References section meant that we did not include dataset citations or DOIs in Table 1. We have since added those following an excellent example from Sutton et al., 2019 (https://doi.org/10.5194/essd-11-421-2019). Table 1 and the References section now both include full citations for each dataset. Now that the datasets can be identified by their citations, we have also shortened their titles and improved the overall readability of Table 1. The MOCHA dataset metadata table currently available on NCEI is also now included as a supplement.

Minor comments:

Line 44: Given this data product excludes seawater pH values derived using glass electrodes (for good reasons) they should consider referencing here the many other papers discussing coastal biogeochemical variability and change and not papers that utilize glass electrode data for estuarine and coastal pH monitoring. Many of the providers of the original data sets have published papers on this topic that could be cited instead.

- Thank you for pointing out this inconsistency. We have updated our references here to point towards higher quality pH monitoring efforts (lines 47-48).

Line 133: From my review of the original data sets, the product likely includes sensors using a membrane-based spectrophotometric method (the SAMI-CO2 as cited) and an equilibration-based method paired with an infrared gas analyzer (the MAPCO2). The phrase “autonomous equilibrium-based spectrophotometric pCO2 sensors” is a mix of the two.

- This has been corrected (lines 170-171).

Line 138: What is meant by “devices”? Are these sensors integrated into a CTD-rosette equipment package?

- Here, meant to reference all the sensors attached to a CTD-rosette, which might also include dissolved oxygen, pH, and chlorophyll sensors. This has been updated to “‘CTD for observations from ship-side profiles with autonomous sensor arrays,” (lines 178-179).

Table 1: Entry 52 title references the Cha Ba buoy, but the product only includes the cruise data for validating the buoy, but not the buoy data?

- That is correct. The Cha Ba buoy data will be included in future updates of this synthesis.

Lines 198-200: If the data product is going to propose and use a new set of flags, this section should explain why the authors chose to deviate from community-developed and widely-used standardized flagging schemes.

- We appreciate the suggestion to use a widely-known flagging scheme like QARTOD rather than developing a new scheme. Unfortunately, while some datasets we incorporated into this synthesis came with QARTOD flags, many others came with different quality information that was not easily mapped to the QARTOD scale. Additionally, what separates a QARTOD flag of 3 (“questionable/suspect”) from a flag of 4 (“bad”) depends on the project and investigator. Given the diversity of initial quality information available to us, we chose to use a simpler scheme of 1 = “plausible and reliable”, 2 = “unevaluated”, and 3 = “unreliable” to have an easily interpretable scale that we could map all datasets to. This reasoning has been added to section 2.4 (lines 257-259).

Section 2.5: Since nearshore biogeochemistry is heavily influenced by sub-daily processes, how do the authors account for potential bias in daily means when data are missing or flagged bad for a portion of the day?

- This was not a significant concern for this dataset for two reasons. For a given high resolution sensor, “unreliable” flags were applied either to individual outlying points or to multiday sections of data that showed extreme sensor drift and evidence of biofouling. In the former case, the removed data represents at most 1/24 of the day’s information, making the loss of one measurement minimally impactful. In the latter case, the period of sensor drift was identified by a date range within which all data was flagged, but there should be no days impacted significantly by flags applied only to a substantial portion of the day. There was also no evidence in our datasets for flags from the original data providers producing noticeable bias in this way. For close examinations of a given time series, we highly recommend data providers screen more closely for daily bias if necessary for their projects. In the context of data that is being compared to lower resolution sensors and daily, weekly, or seasonal discrete observations, the potential for bias in daily averages of a high resolution dataset is no more concerning than the practice of presenting discrete data observations as “the” oceanic conditions.

Top of page 21: First continued table entry looks incomplete.

- This table entry has been removed entirely since we no longer include calculated pH data.

Line 329: Figure 5 illustrates the ability of the data product to capture a monthly climatology. Seasonal variability could be interpreted as capturing all seasons over the entire time range of the data sets.

- Thank you for the clearer terminology. We have revised our discussion of Figure 5 to reference “monthly climatology” instead (lines 425-427).

Lines 335-337: Could differences in data density between those time periods be impacting this result?

- The data density is very comparable between the April to June and July to September periods in all regions, so the differences in variability and mean conditions between these two time periods are likely real.

Table 3 and associated discussion: I was surprised to not see a comparison to, or at minimum a mention of, previously-published TA-S relationships for these regions.

- We have substantially revised this section following the discovery of a years-old quality-control issue with some autotitrator data supplied by the author team (new text in lines 485-516). The data impacted have been removed from the MOCHA compilation and this manuscript. As such, we have replaced the old Table 3 and

Figure 7 with an updated Figure 7 that shows the regional and nearshore (< 2 km distance) and offshore TA-S relationships. We now discuss these results in context with Fassbender et al. (2017), Cullison Gray et al., (2011) and others.

Line 419: "Saildrone" is a company. These types of oceanographic platforms are commonly called Uncrewed Surface Vehicles (USVs).

- This has been fixed.

Line 418: This product should be named and cited.

- We now explicitly name and cite SOCAT here (lines 565-566).

Line 423: The mention of considering deeper water here is confusing because there is no mention of a desire to assess bottom water earlier in the manuscript. The analysis is focused on varying definitions of surface water.

- We have made it clearer throughout the manuscript that our case examples represent just a few of the potential uses we envision for the MOCHA synthesis. While we focus primarily on surface water for our case examples in order to highlight the coastal autonomous sensors and moorings this compilation includes, we have taken care to ensure that the MOCHA synthesis is also supportive of investigations focused on deeper waters.