

Reviewer One:

General Comments

The review 'Synthesis of Data Products for Ocean Carbonate Chemistry' presents an extensive and thorough summary of multiple oceanic carbonate system data products. This work includes an overview of cruise data compilations, time series data synthesis products, gridded and derived data products, multi-product analyses, and model based data synthesis products for ocean carbonate chemistry. Key information on data availability and access links, with associated references, is reported concisely and clearly. In addition, it is explained that overlaps, such as the use of SOCAT and/or GLODAPv2 in the majority of data products, allows for additional quality control and the test and intercomparison of the different approaches used to generate the respective products. This work is a very useful tool and valuable contribution and will considerably benefit the scientific community in several disciplines. The preprint manuscript is well written with key information regarding each data product reported in a series of tables. I recommend publication following minor revisions, as detailed in the following.

Response: Many thanks for the supportive and kind words about this manuscript. We are adopting all the suggestions in full. Details of our responses are provided below.

Specific Comments

Check for repeat definitions of certain acronyms, e.g. DIC, TA, as the subsequently repeated definitions can be removed to make the text more concise. Make sure to use the acronyms in the remainder of the text, for consistency and instead of writing out in full each time.

Response: All dissolved inorganic carbon (DIC) has been changed to DIC except for the first one. Similarly, all total alkalinity (TA) occurrences have been changed to TA except for the first one. We also went through the manuscript and did the same thing for other variables, e.g., dissolved oxygen, aragonite saturation state, as well.

Check consistency of certain variables, specifically the saturation states as Ω_{arg} , Ω_{arag} and Ω_{Ar} styles are used, for example. Or if the different notations that are used in the manuscript text are due to the specific notation of that variable in the data product, then perhaps retain use of consistent acronyms, if they are the same, e.g. dissolved inorganic carbon (DIC), and re-define variables using a different acronym/notation per dataset, e.g. aragonite saturation state (Ω_{arag}).

Response: All of them have been changed to Ω_{arag} , or Ω_{calc} .

Check consistency with defining the pH scale used, e.g. pH on total scale or pH on the total hydrogen ion scale.

Response: All occurrences have been changed to total scale.

Technical Corrections

Line 111 is ‘... 1690 to 1730 Gt of Carbon ...’ a global average? what is defined as surface ocean (depth)?’

Response: “1690 to 1730 Gt of Carbon” refers to the total amount of dissolved inorganic carbon in the surface ocean, the sentence was rephrased as below for clarity. “Since the beginning of the Industrial Revolution, the total amount of dissolved inorganic carbon (DIC) in the layer from 0 to 200m has risen from 1690 to 1730 Gt of Carbon, and from 35,400 to 35,560 Gt C below 200m (Sabine et al., 2004a; Müller et al, 2023a).” See Lines 122 - 124 of the revised document for more details.

Line 113 regarding ‘...the oceans’ buffer capacity...’ add details to further explain this concept on first usage; buffer against?

Response: A new sentence, “Buffer capacity refers to the ocean’s ability to resist changes in pH, and thus also the partial pressure of CO₂ (pCO₂), when CO₂ or any other acid or base is added or removed.” has been added. See Lines 125-126 of the revised document for more details.

Line 120 replace ‘... parameters... ’ with ‘variables’ for correctness and consistency as used on Lines 347, 348, 385 for example; this is the case for the use of this word in other places in the text

Response: All occurrences of “parameters” have been replaced with “variables”.

Line 129 is there a word missing at the end of the statement ‘... weakened seawater buffer capacity by biologically induced CO₂...’?

Response: The sentence has been replaced by the below: “Furthermore, eutrophication and hypoxia in coastal regions may exacerbate OA in oxygen-deficient bottom waters, as biologically produced CO₂ weakens the natural buffering capacity of seawater.” See Lines 142 - 144 of the revised document.

Line 184 replace ‘... parameters... ’ with ‘variables’; this is the case for the use of this word in other places in the text

Response: Done, as mentioned above.

Line 221 acronyms ‘... dissolved inorganic carbon (DIC), total alkalinity (TA)...’ are already defined earlier in the text

Response: Done, as mentioned above.

Line 243 replace '... parameters...' with 'variables'; this is the case for the use of this word in other places in the text

Response: Done, as mentioned above.

Line 266 replace '... parameters...' with 'variables'; this is the case for the use of this word in other places in the text

Response: Done, as mentioned above.

Line 292 replace '... parameters...' with 'variables'; this is the case for the use of this word in other places in the text, which won't be indicated for each further occurrence beyond page 10 to limit the repetition

Response: Done, as mentioned above.

Line 349 acronyms '... dissolved inorganic carbon (DIC), total alkalinity (TA)...' are already defined earlier in the text

Response: Done, as mentioned above.

Line 414 replace '...alkalinity...' with '... TA...', assuming it is total alkalinity or otherwise please specify

Response: Done. See Line 563 of the revised document.

Line 488 is the statement '... surface-ocean carbonate conditions ...' referring to carbonate ion concentrations or carbonate system variables, please clarify

Response: The latter. The word "carbonate" has been replaced with 'carbonate system'. See Line 652 of the revised document.

Line 690-691 check font and type setting

Response: Fixed.

Line 717-718 check all that acronyms previously defined could be used for all variables listen in full

Response: All three terms have been replaced with their corresponding abbreviations. See Line 915 of the revised document.

Line 735 has acronym '... OA ...' been defined?

Response: We just double checked. It was defined on Line 130 of the revised document.

Line 841 has '... [H+] ...' been defined/explained in full?

Response: We just double checked, and it was defined on Line 133 of the revised document.

Reviewer Two:

This paper encapsulates a wide arrange of carbonate system chemistry data including bottle samples, sensor data, gridded products, interpolations, and modeling outputs of varying complexities. This work will act as a useful synthesis, and currently reaches the goal stated in the abstract to “help researchers identify and access data products that best align with their research objectives”. The authors effectively divide a wide variety of data products into cohesive groups and make clear to the reader how at its core, all of these products stem from direct measurements. On the whole, I believe this paper will shortly be ready for publication and prove to be a useful scientific resource, pending minor revision to improve clarity and utility for the reader.

Response: Thank you very much for your supportive and insightful comments. We are adopting all the suggestions in full. Details of our responses are provided below.

General document comments

1. There is inconsistency across the document with respect to acronyms. A uniform system should be applied throughout the document. Understandably, with 35 pages of text it may make sense to redefine acronyms more than once as a reader may be looking at select sections. Especially in a research community where acronyms are used very often, acronyms could be included in this document to educate the reader, but consideration should be given to the purpose of abbreviating throughout the document.

A selection of examples highlighting inconsistent acronym use:

- The full name and acronym are used (line 443 “European Station for Time Series in the Ocean (ESTOC”).

Response: We provided both the full name and the acronym for this case, because it’s the first time this acronym appears in the manuscript.

- Only full names are used, even when the acronym was defined earlier or is commonly used (line 956 “Global Carbon Budget”, line 952 “Sixth Coupled Model Intercomparison Project”).

Response: We fixed this throughout this manuscript.

- The full name is used and acronym is not defined, then acronym used later in the text for No. 55 “Global Ocean Biogeochemical Models”...“GOBMs”.

Response: The following cases of “Global Ocean Biogeochemical Model” has been replaced with “GOBM”.

- Acronyms are used as titles for most, but not all, datasets.

Response: We went through all the data products and tried our best to use acronyms wherever possible. Below are some examples:

- “Global interior ocean mapped climatology from GLODAPv2” has been replaced with “GLODAPv2 Climatology”.
 - “Salish Cruise Data Package and Multi-stressor Data Product” has been replaced with “Salish Sea Multi-stressor Data Product”.
- The acronym is defined far earlier in the text (“LDEO” defined on line 173, used without definition on line 573.)

Response: All following cases of “Lamont-Doherty Earth Observatory” have been replaced with “LDEO”.

- In some cases, the acronym is used first and full name included in parentheses (line 1060 “CARINA (CARbon dioxide IN the Atlantic Ocean)”) but in most cases full name is used and acronym is defined in parentheses (line 1065-1066 “Global Ocean Ship-based Hydrographic Investigations Program (GO-SHIP”).

Response: Fixed. See Line 277 of the revised document. Now all of them have the full names in front of the acronyms wrapped in parentheses.

- Acronyms are defined but then not used later in the text (line 817 “World Ocean Atlas 2013 (WOA13”).

Response: It has been changed it to “WOA 2013”. See Line 1083 of the revised text.

2. Descriptions with repetitive information

This project requires significant synthesis of many datasets, and as a result is a long document. One way you might consider abbreviating the text would be to eliminate sentences in the description of each data product that restate the utility of the measurements, which you’ve already explained to the reader (very convincingly!) in the introduction. For example, data set 16 states that “These synergistic measurements have contributed to global ocean carbon observation networks (e.g., the newly released SOCATv22), which have improved our ability to characterize natural and anthropogenic drivers of ocean carbon uptake and acidification.” One might revise this to simply say “the measurements are included in global ocean carbon observation networks (e.g., the newly released SOCATv22)”.

Response: This sentence has been removed.

Another example is Dataset 20 which states “[a]ltogether, SPOTS’ pilot increased the readiness of biogeochemical time-series (Lange et al., 2023) and facilitates a variety of applications that benefit from the collective value of biogeochemical time-series observations”. This is likely true, but similar statements could be made for each dataset. Uniformity across the descriptions should be assessed. If I had to guess, a scientist from each project wrote a summary of their dataset and understandably is accustomed to explaining the utility of the data along with the description of the data itself.

Response: This sentence has been removed.

3. Interpretation within description

Similarly, a few of the datasets include a sentence on interpretation. For example, dataset 13 states “The findings underscore significant variability in the seasonality and interannual trends of surface carbonate chemistry across different regions and reef zones.” While this is likely true, it stands out when reading the descriptions because most do not include this type of interpretation or description of data patterns, and it is outside the aims and scope of ESSD to include.

Response: The identified sentence has been removed. We’ve also gone through the entire manuscript to remove any sentences that are not essential to the products they describe.

3. Connecting datasets

There are some datasets which have different versions or very closely related products but have been listed as subsets (26a and 26b) while others have separate entries (44 and 47, 58 and 59). Some clarification could be made about why this choice was made in the methods, or a uniform decision about closely related products could be made.

Response: These products have now been organized together. We also did the same thing to additional products.

4. Submission of future datasets

It’s a great service to the research community that you intend to update this collection of data in the future. In order to get submission from scientists who weren’t involved in the first iteration, it may be worth mentioning this fact in the abstract, so that those who may have data to offer know they can add theirs even if they don’t read further into the paper.

Response: A new sentence on this topic has been added to the Abstract.

Line edits

Line 112 - Reconsider word choice for “belies”, sentence should be simplified for clarity

Response: The sentence has been changed to “The seemingly small increase of 0.5% results in a significant depletion of the oceans' buffer capacity (DeVries, 2022).”

Line 158 - are cruise datasets ordered in any particular way?

Response: Not really. Generally, open ocean data products are listed first, followed by the coastal ocean. In Lines 172 - 173 of the revised text we inform the reader that “The products are organized based on end-used needs and listed within each class with no particular order...”

Line 160 - same as above for time series data.

Response: See above.

Line 260 -Reconsider word choice for “aegis”, sentence should be simplified for clarity

Response: The phrase - “the aegis of” has been removed. Below is the revised sentence: Metzl et al. (2024) aggregated over 44,400 measurements of DIC and TA from a series of research cruises and SOOP across various oceanic regions from 1993-2022, under several French research programs, to create a product called “Service National d’Analyse des Paramètres Océaniques du CO₂ (SNAPO-CO₂)”.

Line 468 - Dataset 19: Were DIC and TA samples truly collected hourly for 6 years at this site? I am skeptical (but highly impressed if so!). Clarification about if these are calculations from pH/pCO₂ sensor data should be described more fully in the text.

Response: DIC and TA were calculated from high frequency measurements of temperature, salinity, and pCO₂. The sentence has been revised to the below: “It is the first high-frequency (1 hour), multi-year (6 years) dataset of salinity, temperature, pCO₂, pH, as well as calculated DIC and TA in the High-Arctic Ocean (Gattuso et al., 2023a).”

Line 524 reads “Self-Organizing Map-Feed-Forward-Network (SOM-FFN)” but has a misplaced hyphen and should read “Self-Organizing-Map Feed-Forward-Network (SOM-FFN)”.

Response: Fixed. Thanks for catching that typo.

Line 719 - It seems some descriptions of the reconstructed or estimated pCO₂ data include uncertainty and misfit values (such as #38), but most do not. I would either leave this out here, or consider whether uncertainties should be added to all datasets (probably not).

Response: Sentences about the uncertainties of this product have been removed.

Line 752- 41) Refers to resolution in 3 nautical miles in the text but uses degrees in the data table, probably worth aligning in the text.

Response: The sentence has been changed to “Bittig et al. (2024) used biogeochemical model output to inform the mapping of sea surface pCO₂ observations in the Baltic Sea and to build a mean monthly climatology for the period 2003 to 2021, with spatial resolutions of 0.10° × 0.05° (approximately 3 nautical miles in both directions).”

Line 887-888 It doesn't seem needed to list individual depth levels, as levels aren't listed for other products.

Response: The listed depth levels have been removed.

Line 902- Table 4.

Clarity and consistency between in text descriptions and tables would be helpful for the reader, as well as consistency in where specific information is listed. I have identified a number of examples, but each of the tables and datasets should be checked over. Below is an example.

No. 42 and 43 list temporal resolution as “adjusted to 2002” and “adjusted to 2000” respectively, then state that they are climatologies in the highlights column of the table. Alternatively, No. 44 includes both pieces of information in the temporal resolution (“monthly climatology, centered around 2010/2011”). Please clarify if being centered around and referenced to a specific year are the same and if so align language, as well as aligning where in the text and tables the information is included.

Response: “centered around 2010/2011” has been replaced with “referenced to January 2011”.

Line 856- No. 49 Appears to be a typo in text or title. Clarify years covered, description says 1994, 2004, and 2014, title says 1994 to 2007, and table says 1994 and 2007.

Response: Here is the new text, “Gruber et al. (2019a) estimated the decadal time-scale changes in the oceanic content of anthropogenic CO₂ (ΔC_{ant}) between 1994 to 2007.”