

RESPONSE TO REVIEWER #2

We would like to thank Reviewer #2 for taking the time to review this manuscript and provide constructive feedback. We respond to each comment below, and where needed, outline changes made to the manuscript based on reviewer feedback or provide explanations as requested. We greatly appreciate the feedback and believe that it has substantially improved the revised manuscript.

Hussein Sayani and Rachel Walter,
On behalf of all authors

RC2: ['Comment on essd-2022-172'](#), Anonymous Referee #2, 21 Sep 2022

Summary:

Walter et al. present the PAGES CoralHydro2k database of coral $\delta^{18}\text{O}$ and Sr/Ca records for the Common Era. This work represents a large collaborative effort to collate and standardize coral proxy data into a machine-readable format. The publication of this data product will be of great benefit to the paleoclimate community, and I anticipate that the database will be used in many future studies. Overall, the manuscript is very well-written, and I recommend minor revisions.

The introduction clearly discusses the utility of coral geochemical proxies and their climate applications. The subsequent Methods and Key Characteristics sections are well-organized and easy to follow. I found it helpful that the authors provide the six tables with the metadata fields in the main text.

I also appreciate that the database includes shorter coral records that may have been excluded from other PAGES 2k data compilations, as these shorter records are still useful for reconstructing snapshots of tropical climate variability. As another beneficial outcome of this work, it is worth noting that an additional 27 previously unarchived records were submitted to the NOAA NCEI database. I am also glad to see that the authors provide a plan for updating the database annually.

My major suggestion is to expand the Usage Notes in Section 4 to provide a more comprehensive overview of how to access and query the Coral Hydro2k database. If there are space constraints, then additional details and specific examples could be provided in the Appendix. The authors mention that a MATLAB script will be provided, but I highly recommend also including example scripts written in Python and R to benefit a larger number of users. For example, a sample Jupyter Notebook that queries the database and performs some simple time series analysis would be very helpful.

Response: We have updated the usage notes section to clarify how the database can be searched. More importantly, we have also developed a new python demo and R example script, that will be archived along the MATLAB example script, to increase accessibility of the database.

The peer-review process could be a valuable opportunity for individuals not involved in CoralHydro2k to test the database and assess whether additional documentation and/or step-by-step guides would be beneficial. Similarly, to the other reviewer's comments, my major comments focus on the current availability of the database and its documentation. I suggest providing more comprehensive documentation in Section 4.2 of the main text that will help the reader/user successfully query the database. I was not able to access any of the files or the MATLAB example script using the links provided in the manuscript to test this out myself.

I recognize the authors are likely hesitant to make the database publicly available before the manuscript is accepted but given that this database is intended to be widely used in the paleoclimate community, it would be helpful to provide access to the reviewers. This would allow me to test the database, better connect the Methods section in the manuscript to the actual database, and importantly, provide more meaningful, constructive suggestions to improve the documentation and sample code.

Response: In addition to providing a python demo and R scripts with step by step instructions on accessing the database, we have now made the database and all accompanying documents fully accessible on both the NOAA archive and the LiPDverse page. The links on the NOAA page were previously offline, while reviewers were provided with an active link to the database on the LiPDverse website. However, we feel that this was not clearly communicated to reviewers. We hope that reviewers can now fully access all resources and test out the database.

Once the manuscript is published, I also suggest the authors utilize the benefits of open source and publish their code and documentation on GitHub. A community-based approach will allow the code to evolve and improve with time. For example, database users could post questions and submit issues if they experience any bugs in the code. As mentioned in the Summary, I also recommend the authors include example scripts for Python and R to help support a wider group of users.

Response: We have created GitHub repository (<https://github.com/CoralHydro2k/ch2k-notebooks>) with example code available in Python, R, and MATLAB.

Minor Comments and Questions:

It may be helpful to have a more obvious 'Submit New Data' button on the main NOAA NCEI CoralHydro2k landing page that directs people to the Google Form for submitting new records: <https://www.ncei.noaa.gov/metadata/geoportal/rest/metadata/item/noaa-coral-35453/html>.

Response: We have requested that the link be renamed as suggested, but there are limitations on the customizability of the NCEI landing page. We recommend visiting the NOAA Study Page, linked on the NCEI landing page.

The 'NCEI Direct Download' for the dataset files on the NOAA/WDS Paleoclimatology website currently goes to the Google Form for submitting new records. I am unsure if this is intentional or if this will need to be updated.

Response: We have requested an update to this link, however, as there are limitations on the customizability of the NCEI landing page. We are working to resolve this and recommend using the NOAA study page, linked on the NCEI landing page.

I recommend providing additional details about modern versus fossil corals in the introduction. Fossil corals are briefly mentioned in L240, but it is important to highlight that they are essential for reconstructing tropical climate variability during the earlier parts of the Common Era. The authors could briefly discuss and reference a few key studies that use fossil corals records to reconstruct tropical climate variability prior to the 1800s.

Response: We have added a few references as requested.

For the fossil coral records, does the database have a way to point to the modern coral dataset that was used to develop the SST calibration?

Response: This is unfortunately not included in this version of the database, however, it is something we will strongly consider for a future release. For now, the geographic metadata can

be used to identify modern corals and fossil corals from the same locations, and in most cases, the modern coral was likely used to generate an SST calibration for the fossil coral. Where available, the calibration information used by the original publication is also included in the calibration metadata fields (see Table 6).

How many of the original studies publish reconstructed $\delta^{18}\text{O}_{\text{sw}}$ values from paired Sr/Ca and $\delta^{18}\text{O}$ measurements? In these instances, does the database include the original Sr/Ca and $\delta^{18}\text{O}$ time series in addition to the $\delta^{18}\text{O}_{\text{sw}}$ time series?

Response: There are 19 studies in the database that published reconstructed $\delta^{18}\text{O}_{\text{sw}}$ values, and each $\delta^{18}\text{O}_{\text{sw}}$ record is accompanied by their original Sr/Ca and $\delta^{18}\text{O}$ time series under the same unique identifier.

None of the metadata field names are italicized in Table 2. I recommend double checking the fields and adjusting the italicized text accordingly.

Response: Thank you for flagging this. The italicization was lost when transferring the manuscript to MS word prior to submission. We have corrected this and updated the definitions in Table 2.

I recommend noting in L160-162 that the unstructured metadata fields are those not italicized in Tables 2-6.

Response: We have added a note as requested.

I am unsure what the 'paleoData_TSid' metadata field means in practice. Additional details in the description would be helpful.

Response: We have updated the descriptions for this and other related metadata fields in table 2 to better explain their purpose.

If any subset of the CoralHydro2k database is used (not just a small subset as discussed in L430), I think it would be beneficial to cite all the relevant original publications especially if this does not cause the new study to exceed the reference limit for the target journal.

Including the original DOIs for each database entry in Appendix A would facilitate this process.

Response: We agree and have updated the text in this section as requested. Appendix A includes all relevant citations for each record and their bibliographies, including DOIs, are included on the reference list.

Citation: <https://doi.org/10.5194/essd-2022-172-RC2>