RC2: 'Comment on essd-2025-467', Alessio Rovere, 30 Oct 2025 reply

Dear Editor,

I have now completed my assessment of the manuscript under consideration.

The PAGES CoralHydro2k Seawater $\delta^{18}O$ Database represents a new and comprehensive compilation of seawater stable isotope ($\delta^{18}O$ and $\delta^{2}H$) and salinity data, developed in accordance with the FAIR principles (Findable, Accessible, Interoperable, Reusable). The initiative successfully addresses the lack of a centralized archive and makes previously “hidden” or scattered data accessible—information that is essential to the Earth Science research community. Its main goal is to support the calibration of coral-derived paleoclimate proxies and to enhance understanding of tropical hydrological processes and the performance of isotope-enabled climate models. The database, which includes more than 18,600 measurements collected between 1972 and 2021, constitutes the most extensive synthesis of marine isotope observations currently available. However, the authors rightly note the need for further globally coordinated sampling efforts, as the spatial and temporal coverage of the data remains uneven and incomplete.

Overall, the topic is valuable and worthy of publication. The dataset is impressive, and the manuscript reads well. Nevertheless, I believe a few aspects could be strengthened.

First, the description of the data fields seems too focused on what has been standardized from previous works, rather than providing clear guidelines for how new data should be reported. The authors should make an additional effort to describe how new data submissions should be formatted—perhaps specifying character limits or input rules for certain fields (some entries in the current database appear excessively long).

The metadata fields are meant to do just this- provide a specific set of guidelines for new data reporting. We have added a brief discussion on this point to Section 2.3: Metadata description and quality control (now Section 2.4: Metadata, quality control, and best practices for future data reporting). Only 10 metadata fields have standardized vocabulary (provided in Table 2). The remainder are intended to be free form to provide flexibility and ease for the data submitters, thus lowering the energy barrier for submission.

Second, the database’s structure—as a single CSV file rather than a true relational database—is a limitation. This format constrains automated validation (e.g., checking date formats, field lengths, or required entries) and makes the file cumbersome to handle. I am not suggesting that the authors completely redesign the structure at this

stage, but these issues should at least be discussed, particularly regarding data quality control and verification.

We believe a simple CSV is the optimal format for this database, as the intent is to make the database interoperable with any software program and in a format that is familiar to the widest possible user base. However, we agree with the reviewer that automated validation would be a beneficial feature and we will consider future options for integrating a validation tool into EarthChem data submission portal.

Concerning usability, I reviewed the GitHub repository containing the example code. Currently, it provides only basic spatial plotting examples with minimal commenting and no use of Markdown cells to explain the workflow—one of the key advantages of Jupyter notebooks. Enhancing this documentation with better-annotated examples and more diverse use cases would greatly benefit potential users.

Thank you for this feedback. We improved the Jupyter notebook by adding explanatory Markdown cells, clearer code comments, and more annotated spatial plotting examples to better demonstrate how users can explore, filter, and visualize the database. In addition, we added a MATLAB script on Github that imports the $\delta^{18}O$ seawater dataset, and provides users with a clear interface to query, filter, and plot the data. We have also added a sentence to Section 5.2: Code availability in the manuscript to encourage users of the database to share their scripts on Github.

To further improve usability and encourage community contributions, I suggest providing a CSV template for new data submissions, along with a validation script to ensure that all fields are completed correctly and all mandatory information is present. This would enhance both the long-term sustainability of the resource and its adoption by the broader community.

An Excel template has been provided for new data submissions. These templates have been uploaded to the database Github page. While a validation script is not available at this time, we agree that it would be a useful tool. We will look into creating such a script to upload to the database Github page and the EarthChem data submission portal.

I hope these suggestions will help the authors strengthen their work.

Thank you for the very helpful suggestions!