Earth Syst. Sci. Data Discuss., community comment CC2
https://doi.org/10.5194/essd-2022-198-CC2, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on essd-2022-198

Nicole Khan

Community comment on "The World Atlas of Last Interglacial Shorelines (version 1.0)" by Alessio Rovere et al., Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2022-198-CC2, 2022

I enjoyed reading “The World Atlas of Last Interglacial Shorelines (Version 1.0)” by Rovere et al. The paper provides a nice overview of preceding database efforts and the WALIS database presented in the special issue.

Coming from the perspective of someone who works on Holocene databases, most of my comments and suggestions have to do with clear and consistent usage of terminology, especially for terms that were primarily developed with Holocene data in mind. I’ve made more specific comments on this point (and other additional minor issues) by line or figure number below:

Line 18-19: Not entirely clear what is the difference between ‘sea-level proxies’, ‘dated samples’, and ‘other time constraints’, but perhaps it will be explained later on in the paper.

Line 74: “work” not “works”

Line 75: “work” not “works”

Line 79: replace “to” with “from”

Line 83: Correct referencing format of Pedoja et al., 2014

Line 129: “… are not duplicated in the database” Does this mean that multiple interpretations of the same sample is not allowed?

Line 148: “a” not “an”

Line 161: “reporting on samples and their dating” - what aspect of samples?

Line 163 (Fig 2): Unclear what is meant by “sea-level datum”

Line 163: all samples need information about their geographic position (and elevation), not just index points, right? This is what the wording implies.

Line 164: what is specifically meant by “included” in this context?

Line 185: "This can happen when different sites containing sea-level index points correlate to the same sample dated, for example, with U-series." Sorry, does this sentence suggest that index points from multiple sites in the database can be generated from a single date present only at one site? To me, this would violate the traditional definition of an index point described in Shennan et al., 2015, Handbook for Sea-level Research, where the location, age, elevation and indicative meaning of a dated sample, not a dated stratigraphic unit is defined. I presume the assumption is that the age of that stratigraphic unit is known within some uncertainty from one location and then can be assumed elsewhere. What about instances where there is only one date from a stratigraphic unit, but a sample from another location/different sites in that same unit? That unit could have formed over hundreds of years to millennia, so is it likely that a single date is representative of the entire unit? Some further explanation, especially of caveats or assumptions, is needed here or elsewhere.

Line 188-189: "In these databases, the relationship between these two entities is always one-to-one (e.g., one radiocarbon sample corresponding to one peat layer indicating a former RSL)." This is not strictly true for the Holocene sea-level databases - there can be multiple radiocarbon samples from the same peat layer that indicate one or multiple positions of RSL depending on the depths of the radiocarbon samples and the data creator's interpretation. I think the examples of "many-to-many" and "one-to-one" relationships could be described more clearly in this paragraph. Furthermore, what's the benefit of allowing for one-to-many and many-to-many relationships from an administrator, data compiler or end user's perspective?

Line 192: "PHP" - should this abbreviation be defined?

Line 195: Define "varchar"

Line 240: "Two manuscripts for which data is available in WALIS 1.0 are only available as preprints." Which manuscripts and will they eventually be published in peer reviewed journals?

Line 275: add "chronological" before "constraints" for clarity.

Line 275: "There is no upper limit on the number of constraints associated with a single sea-level proxy." This makes sense, but some description of how age uncertainties are/should be considered should be provided.

Line 276: Can an example of a maximum or minimum limiting age be given? I assume it's from stratigraphic units above (younger) or below (older) the unit from which the sea-level proxy was derived?

Line 277: Wording of sentence starting with "Together…" is a bit unclear; can you please rephrase?

Line 284: Add "types of" before "sea-level indicators"

Line 311 (and Fig 9): Do you mean orthometric (instead of 'ordnance') datum? My understanding is that "Ordnance Datum" is a specific vertical datum used in the UK that measures heights relative to historic MSL recorded at specific tide gauges (Newlyn, Belfast, and Malin). Some clarity on use of terminology (i.e., providing a definition of each term in the context of the database) is important here when presenting fields of metadata in the database.

Line 319: "the WALIS interface requests as mandatory information the elevation of the proxy and the associated 2σ error" conflicts with what is written in Figure 10, which says

"Elevation errors in WALIS are insterted but the compiler as +/- 1σ"

Line 332: What's a "modal limit"? Odd phrasing to me.

Fig 1: A few questions/suggestions to improve clarity:

- From this flow diagram, it seems at present, an end user can only access data through zenodo as .csv, .xlsx or .geoJson files if they do not have SQL knowledge. There is a 'Visualisation via Shiny App' option on the diagram, but it's unclear how an end user accesses this app from this diagram.
- There seem to be some details lacking between the data compiler and PHP interface - can there be bulk uploads via csv or excel files or can data only be added one point at a time.
- There are some actions (e.g., "Mod/Delete request") and objects ("Jupyter notebooks for database queries") in the diagram, but it's somewhat unclear who can perform these actions, or what the objects relate to. It might be useful to colour-code the diagram to indicate which actions are possible for the different parties involved (admin, data compiler and creator, end user) and have different symbols for actions and different states of data.

Fig 2: Why do the number of records on the right hand side of the figure differ from the number on the left? They are also referred to as "records" in the text in the top right of the figure and "amount of data" in the figure caption. Some explanation and consistency in terminology would avoid confusion.

Fig 3: The names of tables shown here (and described in the text) do not correspond to their names in the WALIS documentation (found here: https://walis-help.readthedocs.io/en/latest/database.html#). As a a result, it was difficult to relate the contents of the paper to the support documentation. Perhaps the tables in the support documentation can be shown here in parentheses to help.

Fig 4: Is this the PHP interface described in Fig 1? Or is this the mySQL database? Or does this not appear on Fig 1? Should be referred to consistently (either PHP interface, MySQL database, or WALIS database interface) for clarity here and in the text.

Fig 7: Assuming these descriptions apply to the data visualisation tool, can the RSL and age error bars and boxes be defined in terms of the uncertainties they represent? For example, a "SLI" here shows 2 sigma RSL and age uncertainties? For the limiting type points, are the error bars plotted at the midpoint of the elevation or age uncertainty distribution of the data point or at the upper/lower end of the distribution? Also, should the y-axis be labelled as "Relative sea level" rather than "Paleo sea level"?