

Earth Syst. Sci. Data Discuss., referee comment RC1
<https://doi.org/10.5194/essd-2021-171-RC1>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

What an incredible solution! (Comment on essd-2021-171)

Todd OBrien (Referee)

Referee comment on "The Plankton Lifeform Extraction Tool: a digital tool to increase the discoverability and usability of plankton time-series data" by Clare Ostle et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2021-171-RC1>, 2021

Dear Authors,

Wow! What an excellent tool/dataset and a super clever solution to the challenge of comparing disparate time series programs and their data! By aggregating species into a cluster of "lifeform" groupings, you solve methodological intercomparison challenges, while also creating a product that is more readily understandable to ecosystem- and policy- level users. You have been very careful and thorough in your design, and I especially appreciate the confidence ratings that you applied to the different lifeform categories. The PLET trait lookup table by itself is super valuable and useful, and being able to apply it directly to time series via the BASSH/PLET tool is even better. Overall, everything about your manuscript was excellent, and it was a pleasure to read. You provided a really well written paper and methodology, and I had only a tiny few questions after reading through it.

First Question: You mention that each time series data set is preserved via a DOI. Is this doi/data the "original raw unaggregated data" (e.g., species counts by month by year **before being translated into lifeform categories**) or is it the data after it has been aggregated into lifeform categories (and individual species data is likely removed from that doi)? I ask because the reason IGMETS (and ICES WGZE/WGPME) only worked with totals (total copepods, total diatoms) was/is because some time series holders were hesitant to share their full raw species data. If you are sharing only the aggregated data, that would sooth most contributors (by not releasing the full raw data itself) and greatly increase/encourage more participation. That is an excellent solution to this ongoing challenge, and I think you should talk to ICES WGZE and ICES WGPME about getting more data sets into your tool.

Second Question: I really like Figures 5, 6, 7, and 8. Is it possible to get the BASSH/PLET tool to automatically generate those? (If it already does, I could not get figure it out.) Or perhaps you can pre-generate them, for the fixed site time series at least? This is very useful summary information about the time series, with or without the

interactive tool component.

Third Question: The PLET Trait Look-Up Table is probably most important part of the entire database/tool "ecosystem". With what frequency do you hope to maintain and expand that table? On a related note, while you say you are marine-focused, adding Baltic Sea species would greatly expand your area of coverage in Europe. Surely HELCOM has most of the Trait info you need to make this expansion in the Look-up Table?

Fourth Question: While you say (in the manuscript) you can't really compare different time series, you actually did .. in Figures 5,6,7,8, by using the Z-score. If you add these graphics to PLET somewhere (Question 2 above), multi-site comparisons or overviews should also be possible. Maybe not "live" (via the tool), but perhaps as pre-generated products elsewhere in the BASSH/PLET web page?

I do not have anything negative to say, but two suggestions:

Suggestion #1: For me, the BASSH/PLET tool will usually "timeout" on the CPR data unless I subset the geographic region and/or time period. (This is not a problem with the single site time series, as they are much much smaller.) Are you using raw, full-geographic-resolution CPR data (i.e., the original silk locations)? For performance, you may want to subset those into geographic average boxes, perhaps 0.5 x 0.5 or 1.0 x 1.0 degree boxes. It would greatly reduce the number of data points the tool would have to process "on the fly".

Suggestion #2: Table A1 is super long ... as in 29 pages in the review PDF. Since I am guessing that listing will change fairly regularly, why not make it an online file and only give a one page example of its content in the manuscript? I find the 29 pages distracting as I am trying to get down to Table A2 ...

I am really excited to see where this will go! Please reach out to ICES WGZE/WGPME to expand the coverage of this tool!

Todd O'Brien