We would like to thank the reviewer for their constructive comments. We copied the comments below and provide our proposed changes in red. We hope that these changes address the concerns and that a revised manuscript will meet the criteria for publication in Earth System Science Data.

Lukas Jonkers
(On behalf of all authors)

The authors present a compilation of paleoclimate data from marine sediment cores covering the past 130 kyr. They give a clear account of their data acquisition strategy, which focussed on cores with d18O measured on benthic foraminifera and radiocarbon dates so that a robust common chronology could be constructed for the entire compilation. Where other paleoclimate data were available for the same core, these data were included in the compilation. They pay close attention to including meta data required to analyse the data further.

Parameter and metadata names have been harmonised and the original naming is preserved so that these can be traced in the original publications. The data are all well referenced with DOIs and citations. New depth-age models have been constructed for all sites using BACON and the published chronologies preserved for reference.

In all this represents a very well researched and harmonised dataset with rich and useful metadata that does not exist elsewhere. The data are supplied in a variety of formats, as R data objects, NetCDF, and in the LiPD format which is itself a set of zipped plain text (csv) files containing the data in a highly structured JSON format (http://wiki.linked.earth/Linked_Paleo_Data).
We appreciate the reviewer's feedback on the value of the data product.

However, the structure of the data within the R objects makes it very difficult to search for and extract subsets of data. For example, all the records of variable "planktonic.d18O", or all the records in a certain geographical region. In the NetCDF and LiPD formats the data are also structured in a similar way, although there may be tools available to help work with the LiPD data.

Use of this data compilation would be greatly enhanced if the data were re- structured into a set of "partially normalized tables" in a "star schema" so that queries can be made in an SQL-like way by joining tables and using select and filter type statements. See Brian McGill's 3rd Commandment here (https://dynamicecology.wordpress.com/2016/08/22/ten-commandments- for-good-data-management/). A key table in this format would for example be the "ParameterListWithRefs.csv" table linked to in the Data availability statement of this manuscript but not found within the data objects.

No specific database software needs to be used; these could be plain text files that could be read in by many data analysis software. This is not a “big” dataset so the structure does not need to optimise storage or retrieval efficiency.

I’m not suggesting that reformatting the data in this way would be trivial for the authors, but the data in their current format are well structured and so it should be possible to write code to do it – and this should be much easier for the authors than for someone coming to it fresh.

We thank the reviewer for their feedback on the format of the database. Because of its structure, linking multiple records to common site information, complex metadata and chronological information, the data we provide are indeed not as easy to query as a simple text file would be. Yet, the referee is right to point out that the ability to extract user-defined sets of records, such as by regions or by proxy type, is a key functionality. We will therefore provide example scripts to query the RDS files and make them available on GitHub. We would like to highlight that similar tools are already available to query LiPD files (https://github.com/nickmckay/LiPD-utilities) and that the PaleoDataView software is specifically designed to interact with the netcdf files (https://www.marum.de/en/Stefan-Mulitza/PaleoDataView.html). Considering that for both the LiPD and netCDF formats querying tools are available, and that we provide similar options for R, we prefer to keep the format of the database as it is now, without shoehorning its highly interlinked structure into another format. This is because the structure of the database also allows for better quality control and updating of the (meta)data and age=depth models. The individual files also make unintentional mixing of information less likely.
The example scripts that we will provide will allow the user to build their own version of a table like ‘ParameterListWithRefs.csv’ and to query the database by:

- Parameter
- Parameter detail
- Sensor species
- Minimum age
- Maximum age
- Resolution
- Number of tie points

The example code returns the indices of the sites that meet the criteria, which the user can use to extract the desired time series data, metadata or chronology data. This is intentional as custom scripts are required to tailor the extraction to each analysis. We will update section 6 and incorporate the information on how to interact with the database.

Minor comment: l. 486 - In the section “recommendations for data archiving” “Include metadata” I would also recommend including information about the size of the sample on which the parameter was measured, e.g. number of foraminifera, mass of sample, total peak area for Alkenones. As this can be very useful when assessing the uncertainty of the value.

We agree and will include a statement encouraging researchers to include those data in the section with recommendations. Note that this also follows the recommendation of PaCTS (Khider et al., 2019).

Text errors:

l. 147 “were” -> “where”

l. 161 “more of data”

These will be addressed.

References:

Khider, D., Emile-Geay, J., McKay, N. P., Gil, Y., Garijo, D., Ratnakar, V., Alonso-Garcia, M., Bertrand, S., Bothe, O., Brewer, P., Bunn, A., Chevalier, M., Comas-Bru, L., Csank, A., Dassié, E., DeLong, K., Felis, T., Francus, P., Frappier, A., Gray, W., Goring, S., Jonkers, L., Kahle, M., Kaufman, D., Kehrwald, N. M., Martrat, B., McGregor, H., Richey, J., Schmittner, A., Scroxton, N., Sutherland, E., Thirumalai, K., Allen, K., Arnaud, F., Axford, Y., Barrows, T. T., Bazin, L., Pilaar Birch, S. E., Bradley, E., Bregy, J., Capron, E., Cartapanis, O., Chiang, H. W., Cobb, K., Debret, M., Dommain, R., Du, J., Dyez, K., Emerick, S., Erb, M. P., Falster, G., Finsinger, W., Fortier, D., Gauthier, N., George, S., Grimm, E., Hertzberg, J., Hibbert, F., Hillman, A., Hobbs, W., Huber, M., Hughes, A. L. C., Jaccard, S., Ruan, J., Kienast, M., Konecky, B., Le Roux, G., Lyubchich, V., Novello, V. F., Olaka, L., Partin, J. W., Pearce, C., Phipps, S. J., Pignol, C., Piotrowska, N., Poli, M. S., Prokopenko, A., Schwanck, F., Stepanek, C., Swann, G. E. A., Telford, R., Thomas, E., Thomas, Z., Truebe, S., von Gunten, L., Waite, A., Weitzel, N., Wilhelm, B., Williams, J., Williams, J. J., Winstrup, M., Zhao, N., and Zhou, Y.: PaCTS 1.0: A Crowdsourced Reporting Standard for Paleoclimate Data, Paleoceanography and Paleoclimatology, 10.1029/2019pa003632, 2019.