Clim. Past Discuss.,
doi:10.5194/cp-2015-165-AC2, 2016

Climate of the Past Discussions

Interactive comment on "An interactive tool for navigation within a database of water and carbon stable isotope records from natural archives" *by* T. Bolliet et al.

T. Bolliet et al.

timprobox@googlemail.com

Received and published: 10 May 2016

Scientific comments

Unclear goal

Thank you for stressing this issue, and we have done our best to clarify our goals in the revised version. This paper had indeed two main original goals. The first one was to provide a simplified and homogenous format for the isotopes data published without any common format or protocol during the last decades, and to compile all these records in a common database. Secondly, as this standard format offers new possibilities in terms of interaction with the data, we wanted to link this database to

an online portal. We estimated that the existing repositories were somehow limited in terms of interactivity, and that additionally to the new frame for datasets, it would be useful to provide new features for the sometimes fastidious online data browsing process, such as dynamic data browsing and visualization. This simultaneous effort to simplify the datasets format and improve data browsing makes the database and online portal highly interlinked. We estimated that this comment could be a good opportunity to clarify the original aim of this project in the article and therefore add some modifications to the abstract and introduction. We also agree that the original title of the article may have been somehow inappropriate and thus modified it as "Water and carbon stable isotope records from natural archives : a new database and interactive online platform for data browsing, visualizing and downloading".

Generalizability

We are favorable to the idea of sharing code and published data from our project. As the construction of the online portal is still ongoing, we cannot yet contribute to code repository, but will include it as soon as the version described in the article is fully functional and open to the community. As mentioned to Reviewer #1, we are also highly open to suggestions and collaboration with other institutions to upgrade our database and online portal.

Data standardization.

We think that LiPD data format might be a helpful improvement for data storage. However, we also think that a standard should be accepted by the whole community, and that all the data repositories should adopt this standard or a similar one. As we explain in a new paragraph of the manuscript, this format is a great opportunity to store and organize data from future publications. We however think that the fastest and less fastidious way to compile hundreds of previously published heterogeneous datasets is a single spreadsheet. This work was done in the frame of a post-doctoral fellowship so time and manpower were very limited to compile data and by the time of this data

compilation (2013 to late 2014), we were not aware of the existence of this LiPD format. Nevertheless, we encourage the community to adopt the LiPD as a new standard for metadata disposition, and we think that, if this format if finally accepted, our database could be relatively easily converted into such a structured standard. In the revised manuscript, we make explicit reference to this issue.

Data synthesis :

The LiPD standard and LinkedEarth project are indeed particularly promising. At the beginning of our work (2013), we were not aware of the existence of these projects and we adopted our own structure following what had been performed earlier for the MARGO project. The LiPD standard will highly facilitate future data storage with a comprehensive structure. We think that the original idea between the two concepts (metadata storage in a single spreadsheet versus individual files or tabs) may be different. LiPD is a great standard for future publications, as each author will be able to directly associate data and metadata in a hierarchical structure. However, considering the limited time and manpower, and as we had to compile and harmonize the existing datasets published along the last forty years, the single spreadsheet remained the fastest and most convenient solution for compiling metadata from these hundreds of datasets.

We also think that it would be feasible, with the participation of the community, to extract the information contained in our metadata spreadsheet and convert it to a structured format, and the LiPD seems particularly appropriate for that. Note that the time required to fill in the LiPD is approximately 10 minute per entry, and therefore the conversion is a workload way beyond our internal capabilities, and will require a community effort, as currently ongoing with the PAGES2k project. As we have also exchanged metadata information with the ISO2k project, we expect that the most recent records from our data base will be soon also available in the LiPD format. We therefore added a paragraph in the manuscript to highlight the need for the community to validate a definitive and interoperable format for metadata before starting to convert all the information from the

compiled datasets.

Age modelling

The original aim of this work performed in the frame of a post-doctoral position was the conversion of hundreds of heterogeneously formatted (age format, file disposition,...) datasets into homogenous records, so we had a very limited time and manpower to additionally gather and compile age-model information, as these data are particularly fragmented. We however agree that the next step in paleoclimate data formatting should be focused on age model information. Similarly to what we mentioned for the metadata, an agreement on the contents and format disposition has to be found within the paleoclimate community, before starting compiling data, otherwise numerous new "standards" types of containers will emerge and the problem of homogeneity will persist. Also, as mentioned by reviewer #2, age model precision and uncertainties became more and more crucial during last years, particularly for the study of fast and abrupt climatic events and transitions. Consequently, it is more and more important to gather all parameters used for the establishment of age models, as well as their associated uncertainties. Unfortunately, many of the records we collected were published more than twenty years ago, and the associated information concerning age model establishment is very limited and incomplete, notably concerning uncertainties.

Chronology ratings

Rating age models involves both qualitative and quantitative factors, and although we did our best to find a clear rating procedure, we thus agree that the expert judgment associated with qualitative aspects might inevitably be challenged. We believe that evaluating qualitative information on measurements such as the posterior distribution of ages is a constructive idea. We also think that asking the authors to provide this information with their future publications will contribute to enhance the evaluation of age models, but we also consider that calculating this distribution for the hundreds of published datasets might necessitate considerable time and manpower. We would be

favorable to include this information in our database and evaluation of the age models, but we would definitively need a collective effort to perform this task.

Statistical analysis

The significance of the difference between two different PMIP time slices was assessed by simply comparing the offset between the average isotopic value of these two periods, to the average value of the standard deviations of the isotopic record for each of the two periods.

We consider that the isotopic offset is (not) significant if the absolute value of the offset is greater (smaller) than the average standard deviation along the two periods. This information is now provided in the appendix of the revised manuscript.

Editorial Comments

We modified the manuscript and figures according to the constructive comments of reviewer #2. We however tried to modify the figures in A1 to A4 by making classes of number of datapoints but we estimate that this treatment leads to a loss of information and thus decided to keep the original figures as they were submitted although we agree that some of them might look a bit spiky.

Interactive comment on Clim. Past Discuss., doi:10.5194/cp-2015-165, 2016.