

Comment on essd-2021-212

Anonymous Referee #3

Referee comment on "Harmonized chronologies of a global late Quaternary pollen dataset (LegacyAge 1.0)" by Chenzhi Li et al., Earth Syst. Sci. Data Discuss.,
<https://doi.org/10.5194/essd-2021-212-RC3>, 2021

Review of ESSD-2021-212 entitled: 'Harmonized chronologies of a global late Quaternary pollen dataset (LegacyAge 1.0)' by Chenzhi Li et al.

In this work the authors present their attempt to harmonize mainly radiocarbon-based chronologies of continental climate records. The harmonisation is with respect of age-model software usage, calibration curve usage, which is a very valuable task. Furthermore, harmonisation is performed with respect to parameters used for the age-depth modelling software. As far as I understand, the authors use the age-modelling software Bacon for age-depth modelling of a huge quantity of records. Before modelling, the cores were manually evaluated in terms of complications, such as radiocarbon reservoir effects, water lines, etc.

While I appreciate their approach, I think there are some things to be improved before suggesting this piece of work for publication. Usually, such a data set and code is generated to be used. Unfortunately, I cannot find any description or manual, how to access the age-depth models. Nor is it possible for me to run the R-script. I admit, I am a R-noob, but I think, application should be properly described with at least a short manual for users with some R-experience (or even noobs). This does not have to come with this publication, but it should at least appear on their github space next to the R-file.

Furthermore, I am not able to find age-depth profiles on their provided Pangaea-page. I thought the authors did all their work (handling reservoir effects, water lines, deciding for the best thicknesses to be applied, ...) in order to provide a homogeneous age-depth data set. And according to their paper, they spend a lot of efforts to evaluate the datings etc of all records. It would be a pity, if they would not share this. Or is the user supposed to start from scratch again? Even if it 'only' means to run their script – if I understand the code structure correctly, the user has to run all of their thousands of records, even if the user is only interested in one or two records. Especially, as this means to run 'millions of MCMC iterations' (line 120) which cannot be that cheap as even admitted by the authors: "... it needs much supervision and computing power" (line 122). Why not provide all age

depth models (including uncertainties) in addition to all meta data and code? Or at least enable the user to only calculate the age-depth models of the records they are interested in?

Another critical question is about the final age models. As I cannot find them, nor am able to run the R script, I have to ask: Which depths intervals do you choose to save for the homogenised age-depth models? In the paper you mention the effect of choosing different levels or depth intervals on the goodness of the model data and that some are better suited than others. However, I even wonder, why a user should care about having the age-depth relationship on a fixed sampling interval? If I want to work with other paleoclimate data, I am interested in an age-depth model, which provides dates at depth, where the proxies were measured. Is the output of your script arranged in a way, that this could be easily accessed? Unfortunately, this is not mentioned in the paper. Or do you expect the user to apply some (more or less) fancy interpolation algorithm to assign ages for the proxy depths?

In addition I have some more issues with respect of minor and technical aspects:

L16 and 46: Please elaborate a bit more on what you understand by 'harmonized chronology' already this early in the manuscript. I am pretty sure, that different people understand different things under this term. I mean later in the paper it becomes clear, what you understand by this term, but I think it is worth to highlight this already in the beginning of your work.

L27-28: This sentence needs more explanations. Maybe not here in the abstract, but below in the according text passages. Please find a more detailed comment below.

L69-74: You provide quite some detailed information on metadata, which I appreciate a lot. However, I doubt that putting those data all in one file is the best option. I agree with referee 2 to splitting this file up in several is maybe more appropriate and easier to handle. At least keep this in mind for any potential future improvements.

L155: 'acc.mean' is possibly 'acc.rate'?

L158: 'We tested six thicknesses (2.5 cm, 5 cm, 10 cm, 30 sections, 60 sections, and 120 sections) ...'. I am not very familiar with Bacon. But, why would you want to test those 6 sampling intervals? I mean, the proxies of the cores were measured at specific depths - wouldn't it be more suitable to only interpolate to those depths, where proxy data exist?

Actually, this is the data, I would be interested in. But it seems, that this is missing completely. What do you suggest to finally obtain the ages at those depths?

L159: 'artificial surface age', Why would it be necessary to add an artificial date? I don't know if I understand the concept of adding an artificial date correctly. Stating things like this sounds very arbitrary. Or do you mean you added another age-constraint due to the assumption that the core sedimentation was active until core recovery? And that the additional age constraint is the year of core recovery? If yes, please consider to specify accordingly.

L159: 'generating 12 age models for each core'. Just to make sure I understand correctly. Your code provides 12 age-depth models for one core. Are all provided in output files?

L170: I think, C exchange between dissolved C-species in water and atmospheric CO₂ is not responsible for 'too old radiocarbon dates'. Instead, this process counter balances to some degree the effect of the arguments listed earlier in this sentence.

L171-173: For some records you added your evaluation of reservoir effects. I appreciate this a lot, but I think it is worth to add a column in your metadata file and mark those records. This would allow a better transparency about what is your evaluation and which information came from the original studies.

L184: For the use of radiocarbon dates for modelling purposes, you followed 'in most cases the suggestions in the original publications'. Please consider – again for a better transparency - to provide information (maybe in your metadata file), for which records you did not follow the suggestions of the original publications.

L189-191: 'For each record, 12 age models were visually assessed. Preference was given to models that fitted the dates well and with small uncertainties when choosing the 'best' model for each record (Blaauw and Christen, 2011; Blaauw et al., 2018)'. This is a lot of work for thousands of records. You are sure, that you did this all correctly for this large amount of records? I wonder if it would have been more objective to apply a short statistical test on this. I mean, most likely a simple least square test between age model and ages of dated depths would do a better and faster job. Also the 'small uncertainty' argument would be most likely more precise and faster to obtain, when calculating the mean uncertainty instead relying on visual assessment.

L203: Who did the evaluation about what a reliable date is? You or the original authors? I can imagine, that this is a difficult task, especially for cores from others.

L247-248: 'where original chronologies outperformed LegacyAge 1.0, ...' How do you know, which model approach outperforms the other? How can you measure or evaluate this? Do you have knowledge of the 'true sedimentation history' of all those records to be able to judge this? Which one do you choose from your 12 ones/core? I think it is very crucial to provide more details on this issue. Or, in case you wanted to express a different thing with this expression, please consider to rephrase this sentence.

Fig. 7: Please provide information about which of the twelve generated age-depth models for each record you show here! Would it be possible to show one additional age-depth realisation, which fits less good with the measured ages. Only to give the reader an idea about the effects of the choice of depths intervals.